Empirical least squares regression models for employment, in- and out-migration, and income distribution in the Northern Great Plains region of the United States
by Eugene Patrick Lewis

A thesis submitted in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE in Applied Economics
Montana State University
© Copyright by Eugene Patrick Lewis (1974)

Abstract:
This research effort is aimed at determining empirical least squares regression models for employment, in- and out-migration, and income distribution. Secondary data are used exclusively. The observation units are 181 non-metro counties in the Northern Great Plains Region of the United States.

The statistical results show that all four models are directly linked to variations in the economic bases of counties. To some extent, this allows the models to be used concurrently in determining impacts.

It was hypothesized and shown that the multiplier effect for employment varies with industry, scale of operation of the various industries, and location in economic space. This conclusion along with the successful inclusion of migration and income distribution suggests that the approach taken in this study is a possible alternative to standard aggregate economic base and input-output studies.

STATEMENT OF PERMISSION TO COPY

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Montana State University, I agree that the Library shall make it freely available for inspection. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by my major professor, or, in his absence, by the Director of Libraries. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature Eugene P. Lewis

Date Nov. 26, 1974

EMPIRICAL LEAST SQUARES REGRESSION MODELS FOR EMPLOYMENT, IN- AND OUT-MIGRATION, AND INCOME DISTRIBUTION IN THE NORTHERN GREAT PLAINS REGION OF THE UNITED STATES

by

EUGENE PATRICK LEWIS

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Applied Economics

Approved:

Chairman, Examining Committee

Head, Major Department

Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

November, 1974

ACKNOWLEDGMENTS

I wish to express my appreciation to the Chairman of my Graduate Committee, Dr. Lloyd D. Bender, for his extensive assistance in the conceptualization and preparation of this thesis. Gratitude is also extended to Dr. Richard McConnen, Dr. M. S. Stauber, Dr. Ancel Haroldsen, and Dr. P. J. Hill, who were members of my committee.

I would like to thank the Economic Research Service of the United States Department of Agriculture for providing funding for this study.

Finally, special thanks go to Stuart B. Townsend for his assistance with the computer manipulations and to Peggy Humphrey who cheerfully and expertly typed the final draft.

TABLE OF CONTENTS

	Page
VITA	ii
ACKNOWLEDGMENTS	iii
LIST OF TABLES	
LIST OF FIGURES	
ABSTRACT	
CHAPTER I: INTRODUCTION	1
Statement of the Problem	1
Justification for the Study	2
Objectives	3
Description of the Study Area, Procedures, and Data	4
Organization of the Study	9
CHAPTER II: INCOME DISTRIBUTION MODEL	11
Description of Variables	12
Regression Results	15
Interpretation of the Model	19
Limitations of the Model	20
Summary	21
CHAPTER III: MIGRATION MODELS	23
Out-Migration Model	25
Description of Variables	25
Employment Opportunities and Income Expectations	28
Amenities	29
Barriers to Migration	30
Regression Results	30
Employment Opportunities and Income Expectations	30
Amenities	32
Barriers to Migration	33
In-Migration Model	33
Description of Variables	33
Employment Opportunities and Income Expectations	34
Amenities	34
Barriers to Migration	35
Regression Results	35
Employment Opportunities and Income Expectations	35
Amenities and Barriers to Migration	37

TABLE OF CONTENTS (cont'd.)

	Page
Interpretiation of Migration Models	37
Limitations of the Models	39
Summary	40
CHAPTER IV: EMPLOYMENT MULTIPLIER MODEL	42
Input-Output Models	43
Economic Base Models	46
Aggregate and Disaggregate Models	49
Description of the Model	50
Dependent Variable	52
Independent Variables	52
Regression Results	58
Interpretation of the Employment Multipliers	60
Summary	68
CHAPTER V: ILLUSTRATION, SUMMARY, AND CONCLUSIONS	70
Illustration	70
Summary and Conclusions	79
LITERATURE CITED	82

LIST OF TABLES

Table		Page
I	PLAINS AREA: 191 COUNTIES INDEXED BY STATE AND STATE ECONOMIC AREA	6
II	PLAINS AREA: RAND McNALLY DESIGNATED REGIONAL TRADE CENTERS	8
III	LABEL, DESCRIPTION, AND SOURCE OF VARIABLES FOR THE INCOME DISTRIBUTION MODEL	13
IV	SUMMARY REGRESSION RESULTS FOR INCOME DISTRIBUTION MODEL	16
V	LABEL, DESCRIPTION, AND SOURCE OF VARIABLES FOR GROSS IN- AND OUT-MIGRATION MODELS	26
VI	RESULTS OF THE GROSS OUT-MIGRATION REGRESSION MODEL	31
VII	RESULTS OF THE GROSS IN-MIGRATION MODEL	36
VIII	DESCRIPTION AND NAMES OF THE INDEPENDENT VARIABLES IN THE EMPLOYMENT MODEL	53
IX	SUMMARY REGRESSION RESULTS FOR EMPLOYMENT MODEL	59
X	SIMPLE R'S GREATER THAN 0.5 FOR AGRICULTURE, MINING, MANUFACTURING, AND TRANSPORTATION TERMS	63
XI	PLAINS AREA: MULTIPLIERS AT MEAN VALUES, 1970	66
XII	LEVELS OF AW_i AND UE_i WHICH INDUCE A 1,390 PERSON POPULATION INCREASE IN MONTANA SEA#4	77

LIST OF FIGURES

Figure		Page
1	Plains Study Area	5
2	Possible Shapes of the Marginal Multiplier Function Given that Only the Scale of the Activity Varies; Dotted Lines Illustrate the Effect of the Distance Variable	61

ABSTRACT

This research effort is aimed at determining empirical least squares regression models for employment, in- and out-migration, and income distribution. Secondary data are used exclusively. The observation units are 181 non-metro counties in the Northern Great Plains Region of the United States.

The statistical results show that all four models are directly linked to variations in the economic bases of counties. To some extent, this allows the models to be used concurrently in determining impacts.

It was hypothesized and shown that the multiplier effect for employment varies with industry, scale of operation of the various industries, and location in economic space. This conclusion along with the successful inclusion of migration and income distribution suggests that the approach taken in this study is a possible alternative to standard aggregate economic base and input-output studies.

CHAPTER I

INTRODUCTION

Statement of the Problem

Economic conditions in the nation and world are imposing external pressures on natural resources development in the Northern Great Plains. The advent of the energy crisis is a salient example. Development of the coal resource in the area could include on-site generation and/or gasification plants as well as strip mining. The mining activity itself requires limited labor but direct labor requirements for electric power plants range from 300 to 500, and for gasification the primary labor projections approach 3,000.

The implications of such development, although much of it is still speculative, are more striking when considered in the context of the existing state of economic development in the rural Great Plains region. The industrial base is almost exclusively agricultural. The average county has a total population of only 9,765. This translates to a population density of about one person per square mile. The local communities (few towns larger than 10,000 population) exist primarily to serve the needs of the dominant agricultural sector. The economies and populations of these communities are stable and they might be considered fragile. Whereas the impacts of a major development can be assimulated in a large well-developed region and the effects muted, the local area immediately surrounding the activity might suffer from disproportionate growth. Given the configuration of the average county

even minimal growth will cause drastic changes in community needs for schools, roads, sewers, and other components of the local infrastructure. Local planners will be confronted with decisions on the level of expansion needed, when to undertake the expansion, and how to pay the bills given the existing tax structure. The problem is compounded by the lack of impact estimates for the local area specifically, and the general uncertainty surrounding the development.

Justification for the Study

Two methods typically are used to determine employment impacts due to changes in an economic structure. The Input-Output approach is well developed and desirable because it demonstrates the interactions in an economy. The high cost of obtaining data to construct such a study usually limits its applicability to state or regional areas. The other common technique is economic base analysis. This approach is inexpensive due to the use of readily available secondary data, but does not provide detailed information about the relationships between sectors in the economy. These short-comings in the available tools have in the past forced planners to use state or regional models to calculate the impacts. Then they rely on informed guesses as to what portion of the total impact accrued to the local area.

The present techniques for examining economic impacts generally are limited to estimating net employment and/or income changes. Two other

important components of a local economy usually overlooked are migration, and income distribution patterns. The common link between these two elements and the multiplier approach is that, in theory, all three should be affected by changes in the local economic base. Employment multipliers estimate the total number of new jobs that will be available due to development. In theory then, a migration model should provide some estimates as to the numbers of people who will move in and out of the area, where they will come from and go, and what relates to their choice of a particular destination. This element is important because incoming and outgoing persons have different implications for the local infrastructure than does, for instance, a reduction in the natural fertility rate. Income distribution also has far-reaching implications for the local community. Will the development reduce or increase the number of low-income families and what are the characteristics of this segment of the population? Estimates of this sort would be useful in determining tax needs versus the existing tax base.

Objectives

The objective of this study is to provide some empirical estimates of local impacts using techniques other than I/O or standard economic base. This overall objective will be obtained by specifying and testing three models:

1) An employment multiplier model which is appropriate to local areas which have different characteristics.

2) Gross in- and out-migration models which relate economic conditions in local communities to population movements.

3) An income distribution model which links income variations to differing land use patterns and other socioeconomic conditions.

Description of the Study Area, Procedures, and Data

The area used for the study, hereafter referred to as the Plains Area, includes the eastern part of Montana, all of North and South Dakota, north-eastern Colorado, and roughly the eastern half of Wyoming. This division was made by the Economic Research Service and closely parallels the Census Boundaries for State Economic Areas. Figure 1 illustrates the geography of the study area in relation to the total United States. Table I lists the 191 counties in the Plains study area by state and State Economic Areas.

Procedures are to estimate each of the four models using the ordinary least squares regression technique. The regression coefficients in the estimating equations are then formatted as impact multipliers. This greatly reduces the complexity of implementation of the models by planners.

Secondary data are used exclusively. Most of the data are either from Census or Office of Business Economics compilations. All data used in the employment multiplier model and much of the data for the income



Figure 1. Plains Study Area.

TABLE I. PLAINS AREA: 191 COUNTIES INDEXED BY STATE AND STATE ECONOMIC AREA.

State	SEA #	Counties
North Dakota	1	Adams, Billings, Bowman, Dunn, Golden Valley, Grant, Hettinger, McKenzie, Mercer, Morton, Oliver, Sioux, Slope, Stark
	2	Burke, Burleigh, Divide, Emmons, Kidder, Logan, McIntosh, McLean, Mountrail, Sheridan, Williams
	3	Barnes, Benson, Bottineau, Cavalier, Eddy, Foster, Griggs, La Moure, McHenry, Nelson, Pierce, Ramsey, Renville, Rolette, Steele, Stutsman, Towner, Ward, Wells
	4	Cass*, Grand Forks, Pembina, Traill, Walsh
	5	Dickey, Ransom, Richland, Sargent
South Dakota	1	Bennett, Butte, Corson, Custer, Dewey, Fall River, Haakon, Harding, Jackson, Jones, Lawrence, Lyman, Meade, Mellette, Pennington, Perkins, Shannon, Stanley, Todd, Washabaugh, Ziebach
	2	Beadle, Brown, Campbell, Clark, Day, Edmunds, Faulk, Hand, Hughes, Hyde, McPherson, Marshall, Potter, Spink, Sully, Walworth
	3	Aurora, Bon Homme, Brule, Buffalo, Charles Mix, Davidson, Douglas, Gregory, Hanson, Hutchinson, Jerauld, McCook, Miner, Sanborn, Tripp
	4	Brookings, Codington, Deuel, Grant, Hamlin, Kingsbury, Roberts
	5	Clay, Lake, Lincoln, Minnehaha*, Moody, Turner, Union, Yankton
Montana	2	Blaine, Cascade*, Chouteau, Daniels, Dawson, Fallon, Fergus, Glacier, Hill, Judith Basin, Liberty, McCone, Phillips, Pondera, Prairie, Richland, Roosevelt, Sheridan, Teton, Toole, Valley, Wibaux

TABLE I. (continued).

State	SEA #	Counties
	3	Big Horn, Carbon, Stillwater, Yellowstone*
	4	Carter, Custer, Garfield, Golden Valley, Musselshell, Petroleum, Powder River, Rosebud, Sweet Grass, Treasure, Wheatland
Wyoming	2	Big Horn, Campbell, Converse, Crook, Fremont, Goshen, Hot Springs, Johnson, Laramie, Niobrara, Park, Platte, Sheridan, Washakie, Weston
Colorado	3	Larimer, Logan, Morgan, Sedgwick, Weld
	4	Cheyenne, Douglas*, Elbert, Kiowa, Kit Carson, Lincoln, Phillips, Washington, Yuma
	A	Adams*,Arapahoe*, Denver*, Jefferson*
	D	Boulder*

*These are the ten Standard Metropolitan Statistical Areas (SMSA's) and fringe counties of SMSA's with a population less than or equal to 1,000,000.

TABLE II. PLAINS AREA: RAND McNALLY DESIGNATED REGIONAL TRADE CENTERS.

State	Trade Center
Wyoming	Cheyenne (Scotts Bluff, Nebr.)
	Casper
	Sheridan
Colorado	Greeley
	Colorado Springs
	Grand Junction
	Denver
	Pueblo
	Durango
Montana	Miles City
	Billings
	Great Falls
	Havre
	Bozeman
	Helena
	Kalispell
	Missoula
	Butte
North Dakota	Williston
	Dickinson
	Bismark
	Minot
	Grand Forks
	Fargo
South Dakota	Rapid City
	Aberdeen
	Huron
	Mitchell (Sioux City, Iowa)
	Watertown
	Sioux Falls

distribution section was received via magnetic tape. Remaining data were gathered at the Montana State University library and computerized.

The observation matrix for the migration models is based on State Economic Areas, 1/ and the 48 contiguous states while the income distribution and employment multiplier equations are calculated using non-metro counties as the sample observation unit. 2/

Organization of the Study

The balance of this study consists of four chapters. Chapter II, "Income Distribution Model," is devoted to testing hypotheses which link low-income families to variations in the economic base of counties (land use patterns) and several traditional variables.

In Chapter III, "Migration Models," estimating equations for gross in- and out-migration patterns are determined. The significance of the various hypotheses and illustrations of the model are examined.

Chapter IV, "Employment Multiplier Model," provides least squares estimates of partially disaggregate employment multipliers for agriculture, mining, manufacturing, and transportation. These estimates

---

1/ The logic behind the boundaries for State Economic Areas (SEA's) is outlined in the Bureau of the Census publication, State Economic Areas, P(2)-10B.

2/ See footnote to Table I.

are the crux of the study as the employment estimates calculated from them are the inputs for the two previously mentioned models.

Chapter V, "Illustration, Summary, and Conclusions," is an illustration of the three models and summarizes the findings of this study.

CHAPTER II

INCOME DISTRIBUTION MODEL

The model developed in this chapter is intended to explain the socioeconomic structure associated with differences in the distribution of income between the 181 Plains counties. Multiple regression is used to test hypotheses drawn from various theories of the size distribution of income.

Normally income distribution is calibrated and discussed in terms of the Lorenze Curve and its associated Gini Ratio [15]. The Gini Ratio is descriptive rather than predictive. It measures non-uniformity in the distribution of income, but fails to indicate the source. The uneven distribution may be due to a disportionately large number of wealthy families or low-income families. The Gini Ratio treats both extremes with the same index, although the two have vastly divergent social and economic implications. The approach taken here will scrutinize uneven income distribution by examining the lower tail of the Lorenze Curve as the origin of existing non-uniformity. The population which characterizes this section of the curve has welfare, tax base, productivity, and other possibly negative consequences for the local community.

Explanations for skewness in the income curve fall in one of two categories: (1) the natural phenomena theory or (2) the institutional theory. The "natural" theorists rely on probability distributions to explain the skewness. The income curve is viewed as a function of the

distribution of abilities and chance which are exogenous of the political, economic, and social structure [16]. Alternatively, it is explained in terms of statistical games of chance [17,18].

The "institutional" approach will be applied in this study. Under this theory, income inequity is explained in terms of many interwoven social and economic conditions at a given time and place [19,20,21]. As such, it is assumed that no law of income distribution can be postulated to fit all situations. Each case must be examined and tested empirically. Relative investment in human capital falls under this theory. The basic premise is that some persons or groups or persons have the ability, opportunity, or foresight to invest in themselves greater amounts of education and training. Each individual's investment then pays returns relative to every other "personal" investment and since the propensities to invest differ, so do incomes. The institutional theory also includes varying economic bases as a determinant of income distribution. The specific hypotheses are outlined in the following section.

Description of Variables

The intention is to relate the variation in the number of low-income families (those with incomes below $5,000 (F)) 1/ to inde-

---

1/ The percent of families below the poverty level was also tested as the dependent variable. This is a complex variable which includes family size, location, and other factors as well as income. When

TABLE III. LABEL, DESCRIPTION, AND SOURCE OF VARIABLES FOR THE INCOME DISTRIBUTION MODEL.

	Variable Name	Source*
(F)	No. families with incomes < $4,999	USDC, Bureau of the Census, Census of Population, Characteristics of the Population, Part I, 1970
(P)	County population	"
(AE)	Agricultural employment	"
(ME)	Manufacturing employment	"
(O)	No. persons ≥ 65 years	"
(FA)	No. farms with ≥ 2,000 acres	USDC, Bureau of the Census, Census of Agriculture, 1969
(S)	No. persons with ≤ 9 years school	USDC, Bureau of the Census, Census of Population, Characteristics of the Population, Part I, 1970
(W)	No. families with female heads	"
(I)	Percent of population non-white	"
(SE)	Mining employment	"

*All data are collected by county.

pendent variables which: (1) measure variations in the economic base of counties, (2) measure the presence of institutions, for example, discrimination, and (3) reflect differences in personal characteristics and human capital investments. The specific hypotheses to be tested are:

1) Education is inversely related to the number of low-income families. The greater the number of persons with less than nine years of school (S), the greater the number of low-income families (F).
2) The percentage of the population that is non-white (I), increases the probability of a greater number of low-income families.
3) The greater the number of families headed by females (W), the greater the expected value of F.
4) The number of persons 65 and older (O) is expected to be directly related to F.
5) The more persons employed in agriculture (FE), the greater the number of low-income families.
6) Extensive use of agricultural lands are inversely related to F. The more farms with areas equal to or greater than 2,000

---

regressed against the independent variables, the resulting equation had numerous sign errors and the R^2 was only 0.50. Because of these problems, the percent of low-income families was not used as the dependent variable.

acres (FA) the fewer low-income families.

7) Manufacturing employment (ME) is directly related to F.

8) Mining employment (SE) and the number of low-income families are inversely related.

The county population (P) is entered to control for the population at risk.

The data matrix for the analysis consists of the 181 non-metro counties in the Plains area. The regression equation is the simple linear formulation where F is the dependent variable. Standard ordinary least squares was used to estimate the relationships.

Regression Results

The independent variables account for 98 percent of the variation in the number of low-income families as measured by variable F. Six of the nine independent variables are significant at the 95 percent level and the signs on the coefficients that are statistically significant are as hypothesized.

One of the three insignificant variables is the education variable. It was entered to test the hypothesis that increased investment in human capital will reduce the number of low-income families but failed to be significant at the 95 percent level, although the sign is correct. Also, the percent of the population which is non-white (I) was not significant. This variable was entered to test the hypothesis that

TABLE IV. SUMMARY REGRESSION RESULTS FOR INCOME DISTRIBUTION MODEL.

Independent Variables	Mean	Regression Coefficient	Computed t-value
O	1127.0	.0833	2.966
S	1197.0	.0297	1.066*
W	58.16	2.595	10.90
I	5.707	.2927	.4821*
AE	749.5	.3268	11.27
FA	96.07	-.6121	-6.018
ME	211.2	.2599	4.810
SE	61.42	-.0211	-.5596*
P	10120.0	.0123	3.731

Dependent Variable = F

Intercept = 47.54 R^2 = .9837 SY_x = 93.63 F-value = 1151.79 N=181

*Not significant at the 95 percent level.

racial discrimination leads to more low-income families and that the predominantly Indian non-white population, in general, does not have the opportunity to invest in education or training, thus leading to lower incomes. The failure of this variable to be significant likely is due to lack of variation in the observation matrix rather than to misspecification of the hypothesis.

The number of families with female heads (W) is significant and directly related to the number of families receiving incomes less than $5,000. This supports the hypothesis that women work at generally low-paying occupations. Also, they typically have invested far less in on-the-job training than their male counterparts and are less mobile.

The age variable (O) predicts as expected: The more persons over 65, the more families with incomes below $5,000. This again is due to discrimination in the job market (which derives, in part, from the rigidity of the skills which characterize this group) and institutional arrangements such as social security and other retirement arrangements, established to provide this segment of the population with compensation.

The productive use to which the resources of an area are put, in part, should determine the occupational mix of the population, the relative wage level, and other socioeconomic characteristics. As such, it is hypothesized that various economic bases have divergent

effects on the distribution of income. As has been demonstrated in other studies [27,31], an agricultural economy will tend to have a low-income population. This hypothesis is tested and born out in variable AE, number of agricultural employees. The relationship is significant and direct.

Agriculture can be divided into intensive and extensive land use operations. It is hypothesized that extensive use of the land, grazing, and large wheat farms, for example, are operated by one or two families using machinery and very little hired labor. The families in these types of operations have relatively high incomes. The measurement used to distinguish extensive from intensive farms is the number of farms with more than 2,000 acres (FA). These farms are thought to be extensive. The sign on variable FA is negative as expected and FA is significant.

The number of manufacturing employees (ME) in the Plains counties is significant and directly related to F. These counties can be considered developing areas in terms of industrialization because firms tend to locate in these areas to take advantage of the existing pool of cheap labor. 2/ Over time an area reaches a degree of industrialization beyond which the indirect and induced effects of manufacturing tend to reduce the number of low-income families.

2/ These are firms which are not tied to a necessary productive resource or to a specific market (transportation costs).

The third insignificant variable is mining employment (SE). Although not significant at the 95 percent level the sign is as anticipated. Mining locates in an area according to the availability of the desired mineral, not to exploit a cheap labor market. The jobs mining provides generally are at high wages, and some of these could be filled by the local labor force thus reducing the number of low-income families. The failure of the mining variable to be significant possibly is due to the low level of the activity in the sample data. The mean number of mining jobs is only 61 while manufacturing is three times larger and agriculture employes more than ten times as many persons in the average Plains county.

Interpretation of the Model

Using the income distribution model to calculate impacts is straightforward due to the reduced form of the equation. As an example, assume the number of extensive farming units in an average county doubles:

$$F = (-.6121)(96)$$

$$F \approx -59.$$

The number of low-income families in that county would be reduced by approximately 59. Supposedly these families would now be involved in extensive farming and hence receive greater income or more likely, they would migrate out of the area.

Agricultural employment is directly related to the number of low-income families in a county. A reasonable assumption, given the historical trend, is for jobs in agriculture to continue declining thus reducing the value of F, ceteris paribus. From the migration model in Chapter III, it is expected that these people will move to nearby urban areas where probability of employment is higher.

Expanded manufacturing in a county accounting for 100 jobs will increase the number of low-income families by about 36. Theoretically, manufacturing locates in pockets of cheap available labor. As such the increase in low-income families is expected to come from surrounding counties. The in-migration model adds credence to this suspicion. In-migrants move to jobs and they move as short a distance as possible. Unfortunately the migration data are not disaggregated by income or other social characteristics. This shortcoming creates difficulty in making statements about low-income families as a specific class in the migration stream. The impact of other significant variables can be calculated and interpreted in the same manner.

Limitations of the Model

The accuracy of the income distribution estimators is decreased because of the measurement errors incurred through the use of secondary data. As in the migration equations, some variables are reflecting more than the specific relationship which is intended. There is

overlapping in the traditional variables. An example might be the education and sex variables. With a minimum of measurement error, this or other interrelationships do not cause serious statistical problems or severely bias the estimators.

Impacts on the number of low-income families resulting from changes in manufacturing employment are estimated to be equal for all types of manufacturing activities. Probably this is not the case. The small fraction of the economic base involved in manufacturing in the region forced the aggregation of manufacturing into a single variable in order to raise the sector to a statistically measurable level.

Defining extensive land use in terms of farm size is not ideal. The data on acreage was available, however, while capital-land-labor ratios were not readily accessible.

Care must be taken not to extrapolate beyond the range of those variables with small statistical means such as manufacturing employment. Projections beyond the range of the data should be scrutinized closely and the confidence interval surrounding these projections expanded.

Summary

The distribution of income in the Plains study area as measured by the lower tail of the Lorenze Curve definitely is related to the economic base as well as several traditional variables, such as age

and sex, which relate to human capital investment differences and institutional anomalies. The economic base in this area is predominantly agricultural with little else in the way of manufacturing or resource exploitation activities. This presents difficulties in predicting changes in the number of low-income families given an increase or decrease of very great proportions in manufacturing or mining employment. The problem is intrinsic in the sample data and cannot be eradicated.

CHAPTER III

MIGRATION MODELS

Two single equation regression models will be used to estimate population changes due to migration in a local area (county): (1) a gross out-migration model and (2) a gross in-migration model. Both equations are in double-log and as such the coefficients are interpreted as elasticities. This functional form was chosen based on a priori knowledge of other studies using similar techniques [1,5]. 1/

It is assumed in both models that participants are utility maximizers and they expect increased utility from migration. Utility functions are of the form $U = f(X_I, X_A)$ where X_I is the individual's income and X_A represents amenities (climate, scenery, and social ties are examples) from which the individual would derive utility. The secondary data used in this study do not provide measurement of amenities, with the possible exceptions of the kinship and urbanization variables. As such, the equations are comprised mostly of variables which may affect the income of a migrant. Two migration theories are combined to determine the sets of independent variables hypothesized to influence the participants' utility from migration: (1) the classical labor mobility theory [1] and (2) the human capital theory [7]. The classical labor mobility theory asserts that workers migrate if the expected discounted

---

1/ The R^2 on both equations is approximately halved when linear forms are used.

present value of their long-run income stream can be increased by such action. Their income is assumed to be functionally related to wages and employment opportunities. As such both these considerations affect the utility of the migrant. By assumption:

$$U = f(X_I, X_A)$$

but

$$X_I = f(W_t, J)$$

where: W_t = wage discounted to present value, and

J = employment opportunities.

So

$$U = f(W_t, J, X_A).$$

This theory further assumes that workers have identical preference functions for all labor markets. If this assumption is relaxed the role of amenities can be examined. A worker's non-market tastes and preference then become part of his utility function:

$$X_A = f(C, K, P, \ldots, X_{An})$$

where: C = climate;

K = presence of friends and relatives;

P = pollution; and

X_{An} = other amenities.

So,

$$U = f(W_t, J, C, K, P, \ldots, X_{An}).$$

The human capital theory of migration states that a person's income is directly influenced by their investment in education and training. As such wage can be expressed as a function of these variables also.

Using this utility function and a priori knowledge gathered from other studies, two migration models were specified and tested.

Out-Migration Model

Description of Variables

The aim of this model is to relate gross out-migration movements from the 17 Plains State Economic Areas (SEA's) between 1965 and 1970 to factors affecting the migrant's utility. The sources and names of the dependent and independent variables are listed in Table V.

The first variable listed, number of out-migrants (M_{ij}), is the dependent variable. Specifically it is the total number of people five years of age and older who lived in an SEA in the study area in 1965 but resided in any one of the other 48 contiguous states or another SEA in the region in 1970. 2/

2/ There are (17 x 48 = 816) possible observations on M_{ij}. Where M_{ij} = 0, the observation is rejected (due to the log transformation) leaving the total number of observations equal to 762.

TABLE V. LABEL, DESCRIPTION, AND SOURCE OF VARIABLES FOR GROSS IN- AND OUT-MIGRATION MODELS.

Label and Description		Source
Dependent Variables:		
M_{ij}	No. of out-migrants	USDC, Bureau of the Census, U. S. Census of Population, Migration Between State Economic Areas, 1970.
M_{ji}	No. of in-migrants	"
Independent Variables:		
Income & Employment:		
E_i	Median years education in a Plains SEA, 1970	USDC, Bureau of the Census, U. S. Census of Population, State Economic Areas, 1970.
I_i	Median income of families and unrelated individuals in a Plains SEA, 1970	"
UE_i	Unemployment expressed as a percentage of the civilian labor force in a Plains SEA, 1970	"
AE_i	Absolute change in non-agricultural employment in a Plains SEA from 1964-1970	USDC, Bureau of the Census, County Business Patterns, 1964 & 1970.
AW_i	Average yearly non-farm wage in a Plains SEA from 1964-1970	"
UE_j	Unemployment expressed as a percentage of the civilian force in the 48 contiguous states, 1970	USDC, Bureau of the Census, U. S. Census of Population, Characteristics of the Population, Part I, 1970
AE_j	Absolute change in non-agricultural employment in 48 contiguous states from 1964-1970	USDL, Bureau of Labor Statistics, Employment and Earnings, 1932-1972

TABLE V. (continued).

Label and Description		Source
AW_j	Average yearly non-farm wage in the 48 contiguous states from 1964-1970	USDL, Bureau of Labor Statistics, Employment and Earnings, 1932-1972
Amenities:		
U_j	Percent of population of 48 contiguous states classified as urban 1970	USDC, Bureau of the Census, U. S. Census of Population, Characteristics of the Population, Part I, 1970
K_i	No. of persons living in a Plains SEA in 1970 who were born in a different SEA	USDC, Bureau of the Census, U. S. Census of Population, State of Birth, 1970
K_j	No. of persons living in one of the 48 contiguous states in 1970 who were born in a Plains SEA	"
Barriers to Migration:		
D_{ij} or D_{ji}	Distance between Plains SEA and 48 contiguous states using state capitols as base points	Rand McNally Road Atlas, 1970
Correction Factors:		
P_i	Population of a Plains SEA	USDC, Bureau of the Census, U. S. Population, State Economic Areas, 1970

The independent variables represent (1) employment opportunities and income expectations in both the origin and the destination, (2) personal considerations (amenities) which do not directly influence the migrant's wage or money income, and (3) physical barriers and costs of movement.

Employment Opportunities and Income Expectations

According to the classical labor mobility theory, people tend to migrate to gain higher wages and hence, greater incomes. As such, the sending areas of low median family income (I_i) should have a larger volume of out-migrants than areas of higher median family income (an inverse relationship). Further, the states with high non-farm wage (AW_j) could expect to receive a greater volume of migrants, i.e., a direct relationship.

Migrants also consider job availability and probability of employment in making their decision to change locales. Following this idea, it is hypothesized that migrants will leave areas where job opportunities are relatively fewer and will prefer destinations where employment opportunities are greater. Migrants will then leave areas of high unemployment (UE_i) and move to areas where the unemployment rate (UE_j) is lower and more jobs are available. The variable measuring job availability is the change in non-agricultural employment (AE_j). Variables AE_j and UE_i are expected to have positive signs on the regression

coefficients (a direct relationship) while UE_j should be inversely related.

It is expected that better educated persons will have a greater propensity to migrate. They stand to gain more than people with less education because of the wider spectrum of job opportunities available to them and the potentially higher wages, i.e., discounted earnings stream.

Amenities

The urbanization variable (U_j) is entered to measure the preference of migrants from the predominantly rural Plains study area for urban destinations. 3/ A direct relationship is expected because urban areas provide cultural, educational, and other amenities due to scale economies in urban centers which are not available in rural areas.

A direct relationship between K_j (the presence of friends and relatives) and migration flows is hypothesized for two reasons: (1) there would be a much larger flow of job information from these areas back to the sending SEA's and (2) the presence of acquaintances with

---

3/ Alternatively this variable's inclusion can be explained by gravitational migration models. These models treat migration as a statistical phenomenon based on probability distributions. These models are discussed in depth by Walter Isard in Methods of Regional Analysis: An Introduction to Regional Science (Cambridge, Mass.: MIT Press), pp. 493-566. Petto [1] discusses them in summary form.

similar backgrounds to sponsor newcomers would tend to lessen the difficulties of transition to new surroundings.

Barriers to Migration

Because moving costs increase with distance and because there will be a greater probability of intervening opportunities with greater distance, it is hypothesized that the distance variable (D_{ij}) will be inversely related to migration.

The population variable (P_i) is entered to control for population size differences between sending areas. More migrants would be expected from areas with larger populations ceteris paribus.

Regression Results

The results of the out-migration regression model are presented in Table VI. The independent variables in the equation are associated with 80 percent of the variation in the dependent variable. Seven of ten independent variables are significant at the 95 percent level and the signs of these coefficients are as hypothesized.

Employment Opportunities and Income Expectations

Two of the three variables which failed to be significant are: (1) average yearly non-agricultural wage (AW_j) in the receiving area and (2) the income in the sending SEA (I_i). The finding that workers do not consider wages and income in an area as the primary motivating

TABLE VI. RESULTS OF THE GROSS OUT-MIGRATION REGRESSION MODEL. 1/

Independent Variables	Mean	Regression Coefficient	Computed t-value
Employment Opportunities & Income Expectations:			
E_i	11.90	1.580	3.102
I_i	6343.0	-.2617	-.8921*
UE_i	4.60	.9129	7.161
UE_j	4.50	.2034	1.470*
AE_j	181000.0	.3335	10.42
AW_j	5980.0	-.1497	-.5862*
Amenities:			
U_j	74.25	.3718	2.276
K_j	100800.0	.4940	19.92
Barriers to Migration:			
D_{ij}	1224.0	-.6303	-13.74
P_i	204300.0	1.127	24.85

Dependent Variable = M_{ij}

Intercept = -6.263 R^2 = .7956 S_{yx} = .3328 F-value = 292.7 N=762

*Not significant at the 95 percent level.

1/ The equation is in double-log form so the coefficients are interpreted as elasticities.

factors (the signs are correct, but the confidence intervals are very large) in their migration decision is not in accordance with the hypothesis outlined in the previous section. However, other researchers [12, 49] examining this same theory also have failed to produce statistical results enabling them to accept this hypothesis. This should not be interpreted to mean that migrants do not consider income as an important determinant in their utility function, but rather that other variables also functionally related to income are more prominent.

Job opportunities were important in the migrant's decision. A high unemployment rate at the origin (UE_i) led to larger numbers of out-migrants. Growth in employment opportunities at the destination, measured by the change in non-agricultural employment (AW_j), directly influenced migration flows. Migrants, however, did not make the decision to locate on the unemployment rate in a destination area (UE_j). UE_j is the third statistically insignificant variable.

The education variable was positively significant as expected indicating that those areas with higher educational levels experienced a larger volume of out-migration.

Amenities

The urbanization variable was significant and positive. Migrants from the rural Plains region tended to locate in urban areas. This may be because of the amenities mentioned previously or because information

is more readily available about urban than rural destinations. Also, migrants may anticipate greater job opportunities in urban centers. Another possible explanation is the gravitational migration theory.

The presence of friends and relatives (K_j) exerted a strong positive influence on migration flows as was anticipated. Present migration streams tend to follow previously established patterns.

Barriers to Migration

Distance (D_{ij}) proved to be a significant barrier to migration. The negative sign on the distance variable shows that nearby areas tended to be recipients of out-migrants. Also, the population control variable (P_i) reacted as expected.

In-Migration Model

Description of Variables

The in-migration model is formulated to explain movements from the 48 contiguous states into a Plains study area SEA from 1965-1970. Again the sources and identities of the variables are listed in Table V.

The second variable in Table V is the dependent variable for the equation. Specifically, it is the total number of people five years of age and older who lived in one of the 48 contiguous states in 1965 but in 1970 resided in a Plains SEA.

The theory justifying the independent variables in this model is the same as for the out-migration equation except there is no interest in the economic conditions in the sending area. This model is concerned with the migrant's choice of locations given that he has decided to move. What factors make one Plains SEA preferable over another?

Employment Opportunities and Income Expectations

It is expected that migrants will gain greater utility by moving to SEA's which offer superior income opportunities and more abundant employment opportunities. It is then hypothesized that volume of in-migrants will vary directly with the non-agricultural wage (AW_i) and migrants will tend to SEA's with a relative abundance of jobs as measured by the change in non-agricultural employment (AE_i).

Population (P_i) 4/ was entered to test the hypothesis that migrants move to population centers because of expected broader job opportunities and higher wages, and, to some extent, amenities [14].

Amenities

The educational level (E_i) in the respective receiving areas is entered as a surrogate variable to measure the relative social status of an area. In-migrants are expected to prefer more status to less and so a direct relationship is anticipated.

4/ Inclusion of this variable can also be rationalized by gravitational migration theory.

Barriers to Migration

Distance (D_{ji}) should pose a barrier to migration as it did in the out-migration model.

Finally, U_j, the percent of state populations which are urban was entered to control the population at risk in the sending areas.

Regression Results

The statistical results of the gross in-migration model are presented in Table VII. The independent variables are associated with 73 percent of the variation in gross in-migration. Six of the seven independent variables are significant at at least the 95 percent level. Also, all regression coefficients exhibit the expected signs.

Employment Opportunities and Income Expectations

Average non-agricultural wage (AW_i), measured directly, is not an important consideration in this case (nor was it significant in the gross out-migration model).

Employment opportunities (AE_i) are important in the migrant's location decision. Again, the population variable (P_i) is also entered as a surrogate to reflect job opportunities. It is significant and positive as anticipated.

TABLE VII. RESULTS OF THE GROSS IN-MIGRATION MODEL. 1/

Independent Variables	Mean	Regression Coefficient	Computed t-value
Employment Opportunities & Income Expectations:			
AE_i	11460.0	.2065	2.570
AW_i	4668.0	.2346	.4941*
P_i	204300.0	.9829	10.64
Amenities:			
K_i	45800.0	.4395	14.28
E_i	11.90	6.379	10.12
Barriers to Migration:			
D_{ji}	1224.0	-.5827	-11.00
U_j	74.25	1.186	7.652

Dependent Variable = M_{ji}

Intercept = -13.41 R^2 = .7332 S_{yx} = .4052 F-value = 287.4 N=740

*Not significant at the 95 percent level.

1/ The equation is in double-log form, so the coefficients are interpreted as elasticities.

Amenities and Barriers to Migration

In-migrants mostly come from nearby areas (D_{ji}). They follow established patterns (K_i) of other in-migrants from the same area and prefer destinations where the populace is relatively more educated (E_i).

Interpretation of Migration Models

As previously mentioned, both equations are in double-log form and the regression coefficients associated with each variable are elasticities. For example, the regression coefficients for the distance variable in the in- and out-migration models, respectively, are -.5827 and -.6303. This means that a 1 percent increase in the distance from the origin, all other variables held constant, will reduce the number of in-migrants by .58 percent and the number of out-migrants will decline by .63 percent. It is notable that the coefficient for the two groups are nearly equal in magnitude. Persons moving either in or out of a Plains SEA are equally affected by intervening opportunities and increasing costs of moving greater distances.

The kinship variables, K_j and K_i, also have nearly equal size, .4940 and .4395, respectively. Migrants either entering or leaving a Plains SEA in general, ceteris paribus, show preferences to follow established routes.

To a lesser degree both the change in non-agricultural employment (AE_j or AE_i) and the average yearly non-farm wage (AW_j or AW_i) demon-

strate the comparability of the two equations. These two similarities along with the two mentioned above are important because in the specification of the models the same theory is assumed to be appropriate for both. The major difference is that the out-migration model considers conditions in both sending and receiving areas and the in-migration model includes only the second step in the decision making process.

Predicting the impact on migration using the in- and out-migration models is a straightforward task because both are reduced form equations: That is, the dependent variable is the only endogenous variable, i.e., the independent variables are all exogenous. As such, the reduced form regression coefficients can be used as short-run impact multipliers and they are additive. As an example, imagine a fictional SEA in the Plains area with established migration flows and a given level of non-agricultural employment. What would be the net population change due to migration given an increase in non-agricultural employment of 400 percent including direct, indirect, and induced effects, all other things held constant? A 400-percent increase might be from 100 to 500 jobs which is not unreasonable considering the number of positions in a manufacturing establishment. The number of in-migrants can be calculated by applying the regression coefficient for the change in non-agricultural employment (AE_i) to the increase:

M_{ji} = (.2065%)(400%)

M_{ji} = 82.6% increase.

If for the previous data period, 1965-1970, there were 100 in-migrants, 183 in-migrants could be expected in the projection period, 1971-1975. This same change in employment opportunities can be examined in the out-migration model by assuming the increase in non-agricultural jobs will drop the rate of unemployment (UE_i). Assume the unemployment rate is at 6 percent and the development of new jobs, both direct and indirect, causes unemployment to decline to 3 percent, a 50 percent drop in the rate. The change in out-migration can be calculated by using the reduced form regression coefficient for UE_i:

M_{ij} = (.9129%)(50%)

M_{ij} = 45.6% decrease.

If there were 100 out-migrants (in 1965-1970), there would now (1971-1975) be about 54. Netting both effects produces a combined population increase of 130 persons. The other significant variables can be projected and analyzed in the same fashion.

Limitations of the Models

It is explicitly assumed in the double-log form that the elasticities are constant throughout the range of the variables. For most this probably is not the case and judgment must be exercised in deter-

mining a reasonable latitude of variability on the various independent variables.

These models are constructed using available secondary data to express the hypothesized connections. However, the data may not provide a good representation of the desired relationship. An example is the urbanization variable in the out-migration equation. U_j is entered to measure the rural migrants' preference for urban destinations. But the urbanization variable may very well indicate income levels, job opportunities, and other factors as well as amenities. There is no way, given the available data, to separate one effect from the other. Hence, the interpretation of the variable is hampered. This is the trade-off between using inexpensive secondary data sources as opposed to collecting primary data which accurately reflect the intended correlation.

Summary

Out-migration and in-migration models were tested in this chapter. The out-migration model is derived from two migration theories and *a priori* knowledge of other studies. The in-migration model was then designed on the assumption that the average migrant, whether leaving or entering the area, responds to similar conditions. The observation matrix for both models consisted of combinations of the 17 Plains SEA's with the 48 contiguous states.

The regression results showed that migration patterns are influenced by both market (income affecting) and non-market (amenities) conditions. The presence of friends and relatives, distance, and probability of employment are considerations which significantly influence location preference and the utility gained of both in- and out-migrants. However, neither of these groups tended to respond to wage differentials in the various receiving areas.

CHAPTER IV

EMPLOYMENT MULTIPLIER MODEL

The idea of a multiplier was first advanced by R. S. Kahn and later was refined and introduced as a useful analytical device by J. M. Keynes. A multiplier is a measure of the total 1/ impact on an economy due to a change in an autonomous element of the economy. It is the ultimate cumulative effect on a rate of aggregate flow per unit of time arising out of an initial unit change in one exogenous element of that flow. An example of a cumulative multiplier effect is the spending and respending of money where only a part of one person's expenditure is someone else's income, i.e., a portion of each person's income is not retained within the stream. This concept applies to an economic area which has well-defined boundaries (such as a nation, state, or county). Employment multipliers are formulated to measure the total cumulative effect on aggregate employment in a defined economy given a unit change in the number of employees in an exogenous sector of that economy.

As previously mentioned, this study is aimed at determining local (county) versus regional impacts. Local impacts are desired because they better pinpoint the area most intensely affected by changes in the economic base. By assumption, the multiplier is the same for any

---

1/ Total as used here is defined to include direct, indirect, and induced effects.

point in a given study area. If the area under study is a large region, then the multipliers for any given activity will be aggregates of the entire region and as such might distort the magnitude of the impact on locations where the specified activity is concentrated. It is this shortcoming of applying regional studies to local areas which this study attempts to overcome.

Input-Output Models

There are two common methods used to calculate employment multipliers. One of these is the static Input-Output (I/O) technique. Static I/O models are based on several explicit assumptions. The assumptions depart from reality, but by simplifying the model, they render it empirically implementable. The first assumption is that each industry is characterized by a linear homogeneous production function, i.e., there are no economies or diseconomies of scale. Secondly, it is assumed that the total effect of carrying on several types of production is the sum of the separate effects. In other words, two firms producing two separate products incur the same costs and gain the same revenues as one firm producing the same two products in a single plant. There is no overlap. The third explicit assumption is that factors required to produce any product, final or intermediate, must be used in fixed proportions (technology is fixed). Fourth, each industry produces a single product or product mix, i.e., there

are no joint products. Lastly, it is assumed, at given prices, the economic system is in equilibrium.

It is implicitly assumed in I/O models that supply is infinitely elastic, i.e., these models are demand-pull. Any fluctuation in the economy resulting in a positive or negative impact is initially caused by changes in demand. Usually these demands for goods are thought to be exogenous or demands for exports. 2/ As such, export industries, those which sell some proportion of their output outside of the economy are credited with the direction of economic change.

The foundation of an I/O model is the transactions table. It is a simple double entry accounting matrix which tables sales and purchases between the various sectors of the economy. The transactions table very clearly illustrates the interdependencies in the economy. Any transactions can be traced through the system to provide information on patterns and degrees of interaction between sectors. Thus, the total effect expressed by an I/O multiplier can be disaggregated to demonstrate the effect on each sector separately. This along with a formal theoretical and mathematical base is the major strength of the I/O method. The interworkings of the economy can be studied as

---

2/ The final demand vector in most I/O tables usually includes: (a) gross inventory accumulation, (b) exports, (c) government purchases, (d) gross private capital formation, and (e) households.

minutely as the researcher wishes. He need only choose the degree of disaggregation desired and define the sectors of the model accordingly.

The shortcomings of the I/O method are encountered in building the transactions table for the local economy. 3/ Primary data are necessary for the most part. It is very expensive and time consuming to gather and must be verified with the greatest of care. If a significant error occurs in a sector of the transactions table it will spread throughout the model during the matrix operations used to solve the system. This will reduce the validity of all sector multipliers.

I/O models are constructed for a point in time. There is no satisfactory method for keeping tables current. 4/ Using an outdated I/O model requires more limiting assumptions, some of which are difficult to defend. 5/

3/ The assumptions are limitations of varying import and severity under certain circumstances.

4/ Some research has been done to resolve this problem. Most notable are: (1) H. B. Cheaney and P. G. Clark, Inter-Industry Economics (New York: John Wiley & Sons, 1959), and (2) Chaiw-shang Yan, Introduction to Input-Output Economics (New York: Holt, Rinehart and Winston, Inc., 1969).

5/ For example it must be assumed that the technical coefficients (defined as the i-th column divided by the total from the j-th row where i = j) are constant over time. The relationships between sectors have remained stable.

To summarize, input-output is a rigorously defined static equilibrium model. The models are desirable for impact analysis because they demonstrate the interactions in the economy. The major weakness of the I/O model is the stringent data requirements needed to construct the transaction table.

Economic Base Models

As an alternative to I/O, employment multipliers can be calculated from an economic base model. These models are not as rigidly specified as input-output and are constructed from secondary data. Only one explicit assumption is made: growth in a given economy emerges exclusively from the export sectors of each industry (this implies that these models are demand-pull). 6/ In theory this assumption requires that exporting sectors be segregated for the analysis. Normally, exporting industries are labeled basic or primary while service or support sectors are referred to as non-basic or secondary. A sector is classified as basic or non-basic using one of three methods: (1) the assumption method, (2) the location quotient method, or (3) the minimum requirements technique.

6/ Economic base models do implicitly make the same assumptions which are explicit in the I/O models. In basic philosophy the two are virtually the same. It is in rigor of actual application where they diverge.

The assumption approach is the easiest. A subjective judgment is made by the researcher as to what sectors are export; the total employment of those sectors is used to calculate the multipliers. The error arising from this method is that most likely no industry exports all of its production. Some of the employment in any given sector should be in the non-basic classification. An exception might be specialized manufacturing activities in rural areas. This error is minimized as the economy under study becomes smaller.

The location quotient approach divides employment in a given sector into the basic or non-basic category according to the national average for that sector. An industry's employment is compared to the national average through a ratio. Any employment above the national ratio is considered export employment and is used to calculate the economic or export base multipliers. The danger in using the method originates with the necessary underlying assumption of uniform demand throughout the nation. In fact, people's tastes and preferences, incomes, and/or other determinants of demand may vary from region to region. 7/

7/ The location quotient method deals with net exports rather than the gross export formulation used in the assumption method. Location quotients assume that if all of a good is exported, then some will necessarily have to be imported to meet local needs. Hence, an industry may export all of its output, but all industry employment will not be classified as export.

An alternative to the location quotient method is the minimum requirements technique. This approach involves calculating and ranking in ascending order the percentages of the total labor force employed in a given industry for some sample of similar economic units (cities of similar size, for example). The smallest percentage is assumed to be the minimum required to serve the local economy. All employment exceeding this lower limit in other units is classified as export employment. In actual application, some number of the bottom observations are thrown out, for instance the bottom 3 percent, to correct for unusual cases in the sample. Difficulties arise in determining where the cut-off point should be. Again, the analyst is obliged to make a subjective determination as to where non-basic ends and export begins.

The necessity of distinguishing basic from non-basic employment in export base models is a deficiency which arises from using secondary data. A significant error in determining which workers are actually employed in the production of goods for export causes calculation of a misleading multiplier rendering the results of the study erroneous and possibly harmful. Another defect of these models is their application in illustrating interaction in the economy. The multipliers are aggregates which do not show changes in separate sectors. On the other hand, export base multipliers are inexpensive to construct and maintain because secondary data are used in the modeling procedure.

This is the advantage of the economic base model and because of this information-cost tradeoff, researchers frequently resort to this method rather than the more informative and expensive input-output approach.

Aggregate and Disaggregate Models

Export base multipliers can be calculated in one of two ways: (1) aggregate formulation or (2) disaggregate formulation. These methods follow the general specifications already discussed in the previous section but require one further assumption.

The aggregate employment multiplier is a weighted average over all export employment in a given economy. A change in employment in one export industry affects total employment the same as a change in any other exporting industry; in other words, there is a single multiplier for all exporting industries in the study area. If mining and agriculture were both defined as export industries, then both are assumed to require the same proportion of local service employment. This assumption is questionable because of the structural differences between industries such as mining and agriculture.

Disaggregate employment multipliers are calculated for each export industry over a sample of economic units (counties, for example) which are assumed to be structurally similar. 8/ Industry i is assumed

---

8/ Disaggregate models assume each industry has a different propensity to export, i.e., to sell goods outside of the defined economy base studied.

to have the same relationship to the non-basic sectors of the economy over the range of geographic observations. So a change in employment in industry i in county x will be the same in county z and one multiplier can be applied to both. This assumption is questionable depending upon the care taken to insure homogeneity in the observations.

Disaggregate multipliers are preferable to the aggregate multipliers. Billings [50] has shown that the disaggregate multipliers are mathematically identical to Type II I/O multipliers (induced effects included) when the two models are constructed with identical assumptions. I/O Type II multipliers are superior on a theoretical, not cost basis, to any economic basic multiplier because of the internal consistency of the models and their ability to indicate interactions between sectors. On the other hand, multiple regression techniques can be used to estimate disaggregate multipliers, thus adding the powers of statistical inference to the analysis.

Description of the Model

The aim of this model is to estimate partially disaggregate employment multipliers for the agriculture, mining, manufacturing, and transport sectors of a county economy. It is built on standard economic base and spatial economic theory. However, the model tested in this study deviates from classic export base models by avoiding the strict division between basic and non-basic sectors. Data permitting, the

variables entered to correct for anomalies in the economic bases of the sample observations are in units different than employment to avoid the statistical problem of multicollinearity.

The general model begins with the basic identity:

$$TE_j = A_j + \sum_i X_{ij} \tag{1}$$

where: TE_j = total employment in region j;

A_j = ancillary employment in region j;

X_{ij} = employment in autonomous sectors i, in region j.

Ancillary employment is assumed to be functionally related to the autonomous basic employment sectors. Equation (1) can be reformulated:

$$TE_j = \sum_i M_{ij} X_{ij} + \sum X_{ij}. \tag{2}$$

Subtracting $\sum X_{ij}$ from both sides:

$$A_j = \sum M_{ij} X_{ij}. \tag{3}$$

Expanding equation (3) to regression form:

$$A_j = a + M_{11}X_{11} + \ldots + M_{nn}X_{nn} + e. \tag{4}$$

The M_{ij}'s are then multipliers of the various autonomous sectors. 9/

9/ The typical multiplier would be reported as $1 + M_{ij}$.

The model was generalized and applied to the 181 Plains counties in a cross-sectional regression analysis using 1970 Census data. Improvements in the model (equation (4)) that are treated in more detail below allow multipliers to vary with locality and the scale of activities without partitioning employment *a priori* into autonomous and endogenous sectors.

Dependent Variable

The dependent variable is ancillary employment. It is the summation of all employment in a county minus total agriculture, mining, and manufacturing employment and minus the export portion of employment in the transportation and construction industries. It includes employment in wholesale and retail trade; finance, insurance and real estate; business and personal services; entertainment; professional and related services; public administration; industry not reported; and that part of autonomous basic employment not partitioned out by the exclusion of the previously mentioned export sectors. Defining the dependent variables in this manner, as ancillary rather than non-basic or secondary employment, makes no *a priori* judgment about the amount of export activities in each sector.

Independent Variables

The independent variables and their associated names are listed in Table VIII. The variables for which employment multipliers are to

TABLE VIII. DESCRIPTION AND NAMES OF THE INDEPENDENT VARIABLES IN THE EMPLOYMENT MODEL.

Name	Description
A	Total agricultural employment by county in 1970.
DA	Cross-product of the distance from the nearest Rand McNally trade center (D) and 1970 agricultural employment.
D^2A	Cross-product of the distance from the nearest Rand McNally trade center squared and 1970 agricultural employment.
A^2	The square of 1970 agricultural employment.
A^3	The cube of 1970 agricultural employment
M	Total mining employment by county in 1970.
DM	Cross-product of the distance from the nearest Rand McNally trade center (D) and 1970 mining employment.
D^2M	Cross-product of the distance from the nearest Rand McNally trade center squared and the 1970 mining employment.
M^2	The square of 1970 mining employment.
M^3	The cube of 1970 mining employment.
F	Total manufacturing employment by county in 1970.
DF	Cross-product of the distance from the nearest Rand McNally trade center (D) and 1970 manufacturing employment.
D^2F	Cross-product of the distance from the nearest Rand McNally trade center squared and the 1970 manufacturing employment.
F^2	The square of 1970 manufacturing employment.
F^3	The cube of 1970 manufacturing employment.
T	Total basic transport employment by county in 1970 calculated by: $T = T_j - (TE_j K) \geq 0$, where K = transport employment in the region divided by total employment in the region where the region is the sum of counties under observation; TE_j is total employment in 1970 in countyj; and T_j is transport employment in 1970 in county j.

TABLE VIII. (continued).

Name	Description
DT	Cross-product of the distance from the nearest Rand McNally trade center (D) and 1970 transportation employment.
D^2T	Cross-product of the distance from the nearest Rand McNally trade center squared and the 1970 transportation employment.
T^2	The square of 1970 transportation employment.
T^3	The cube of 1970 transportation employment.
V	Total basic local government employment in 1970 calculated in the same manner as the variable T above.
I	Total institutional population by county in 1970.
C	Total number of college students in group quarters in 1970.
Cd	Dummy variable for location of colleges where 1 = no students.
H	Total number of motel and hotel employees in 1970.
$\bar{Y}$	Median family income in 1970.

be directly calculated are agriculture (A), mining (M), manufacturing (F), and transportation (T). It is hypothesized that the multiplier for each of these industries is a non-linear function of (1) distance from a major trade center and (2) scale of the activity. The non-linear distance variable will allow the value of the multiplier to vary, corresponding to the location of a given activity relative to major trade centers (as defined by Rand McNally). No a priori hypothesis is made concerning the shape of the function as distance changes. According to spatial economic or central place theory the value of the multiplier should at first increase, reach a maximum, and then decrease (an inverted U-shaped curve).

It is frequently assumed that scale of basic activities in rural areas result in economies of size in support or service sectors [45, 51]; hence, the hypothesized non-linearity in the agriculture, mining, manufacturing, and transportation sectors of this model. This improves on the I/O model specification where the same employment multiplier is used regardless of scale or location within an area.

The multiplier for each of the four sectors, A, M, F, and T, is the partial derivative of a reduced form of the standard cubic function in two variables. For example, agriculture is expressed as:

$$\text{Ancillary}_A = \beta_0 + \beta_1 A + \beta_2 DA + \beta_3 D^2 A + \beta_4 A^2 + \beta_5 A^3, \qquad (5)$$

and the employment multiplier for agriculture is:

$$\frac{\partial \text{ Ancillary}}{\partial \text{ Agriculture}} = \beta_1 + \beta_2 D + \beta_3 D^2 + 2\beta_4 A + 3\beta_5 A^2. \qquad (6)$$

The calculation for the other multipliers, M, F, and T, is analogous. I/O and standard export base models implicitly assume equality of the marginal and average multipliers. This restrictive assumption is avoided here through the use of partial differentiation. $\frac{\partial \text{ ancillary}}{\partial \text{ agriculture}}$ is the incremental or marginal addition to ancillary employment given a one unit change in agricultural employment. The average multiplier is reported in I/O and economic base models. These are not comparable with the above marginal configuration.

Total jobs in agriculture, mining, and manufacturing are surrogates for levels of export employment in each industry, i.e., these industries were partitioned using the assumption method. Misspecification arises if there is wide variation in the percent of export employment in these industries in the Plains area.

Transportation employment was divided into export or ancillary through the use of location quotients. 10/ This assumes that local demands for transportation in the Plains county is the same as the overall proportion for the region. The objective was to isolate and assign a surrogate measure to the influence of transcontinental railroad maintenance and operation centers within the Plains region.

10/ The exact formula used is listed and explained in Table VIII.

The remaining seven independent variables in Table VIII are entered to correct for employment in sectors which are usually considered as local service industries, but, because of their size or other peculiarities, are autonomous to the local economy. The specification of these variables deviates from previous studies in at least one important way. They are entered as surrogates rather than levels of export employment. This is done to avoid possible collinearity from becoming a statistical estimation problem.

Variables C and Cd are designed to account for counties in which a school serving a state, region, or multicounty area is located. Variable H is intended to correct for those employees providing services along major transportation routes and at heavily visited tourist sites. Institutional population (I) is entered to reflect institutional establishments which serve greater areas than the county in which they are located. Montana State Prison in Powell County and the Veteran's Hospital at Fort Harrison in Lewis and Clark County are examples of this circumstance. Counties with greater than average local service government sectors are isolated by inclusion of variable V. As indicated in Table VIII, basic local government employment (V) was calculated using the location quotient method. Lastly, median family income ($\bar{Y}$) is entered to neutralize the effect of income variations on the demand for local services. Higher incomes might have off-

setting influences, They increase the demand for local services and and their convenience but also increase the consumer's shopping range.

Regression Results

The summary regression results for the model are listed in Table IX. The independent variables are associated with 97 percent of the variation in ancillary employment. Twelve of the 26 independent variables are below the 95 percent level of significance. The signs on the terms comprising the multipliers will be discussed in the following section.

Four of the six variables entered to correct for abnormally large service employments in various counties are significant. I, C, Cd, and H are all directly linked to increased ancillary employment supporting the theory that these industries service autonomous demands.

The income variable ($\bar{Y}$) is not significant. This may be due to the conceptually offsetting effects of increased incomes or to a lack of variation in median income in the 181 observations. This is plausible because of the rural nature of the counties under study. There is little deviation from the mean value of $\bar{Y}$ throughout the cross-sectional data. The other insignificant correction variable is the measure of local basic government employment (V). This indicates that local government sectors are similar throughout the Plains area.

TABLE IX. SUMMARY REGRESSION RESULTS FOR EMPLOYMENT MODEL.

Independent Variables	Mean	Regression Coefficient	Computed t-value
A	744.3	-1.040	-1.537*
DA	53,410.0	.0174	3.813
D^2A	6,959,000.0	-.0001	-3.258
A^2	780,100.0	.0010	1.952
A^3	13,340.0	-0.0263×10^{-5}	-2.280
M	61.73	5.057	2.606
DM	4,872.0	-.0315	-1.500*
D^2M	594,800.0	.0002	1.861*
M^2	40,690.0	-.0077	-1.414*
M^3	498.4	$.3602 \times 10^{-5}$	1.129*
F	212.1	6.719	5.447
DF	9,473.0	-.0718	-3.920
D^2F	1,116,000.0	.0003	3.629
F^2	362,100.0	-.0011	-1.181*
F^3	14,120.0	$.0196 \times 10^{-5}$	1.224*
T	35.57	13.27	3.743
DT	1,203.0	-.0503	-12.49*
D^2T	152,300.0	-.0004	-1.286*
T^2	18,870.0	-.0172	-2.128
T^3	187.9	$.7234 \times 10^{-5}$	1.703*
V	29.44	1.157	.7746*
I	134.5	1.493	4.920
C	152.9	.8045	4.572
Cd	.8343	-762.5	-3.734
H	13.52	7.11	3.011
$\bar{Y}$	7,200.0	.0744	1.246*

Dependent Variable = Ancillary Employment

Intercept = 538.8 R^2 = .9698 S_{yx} = 693.7 F-value = 190.3 N=181

*Not significant at the 95 percent level.

Interpretation of the Employment Multipliers

The extent of ancillary employment is thought to be a scale function of employment in the agriculture, mining, manufacturing, and transportation industries. Simultaneously, ancillary employment is hypothesized to be systematically related to the distance of the activity from a major trade center. Specifically it is a second degree quadratic function of distance combined with cross-products of agriculture, mining, manufacturing, and transport employment. The functions of the four actual industries are analogous. Each industry is considered separately. In equation form = (fictional industry Q),

$$\text{ancillary}_Q = \beta_1 Q + \beta_2 DQ + \beta_3 D^2 Q + \beta_4 Q^2 + \beta_5 Q^3. \qquad (7)$$

The graph of equation (7) is a surface in three-dimensional space with combinations of ancillary employment, industry Q employment, and distance representing two-dimensional planes in the surface. The graphic representation of the function is extremely complex. For this reason, equation (7) is differentiated with respect to variable Q holding variable D constant, giving:

$$\frac{\partial \text{ ancillary}}{\partial Q} = \beta_1 + \beta_2 D + \beta_3 D^2 + 2\beta_4 Q + 3\beta_5 Q^2. \qquad (8)$$

Having taken the partial derivative, the presentation is in two-dimensional planes and the function is interpreted in terms of additions and subtractions at the margin.

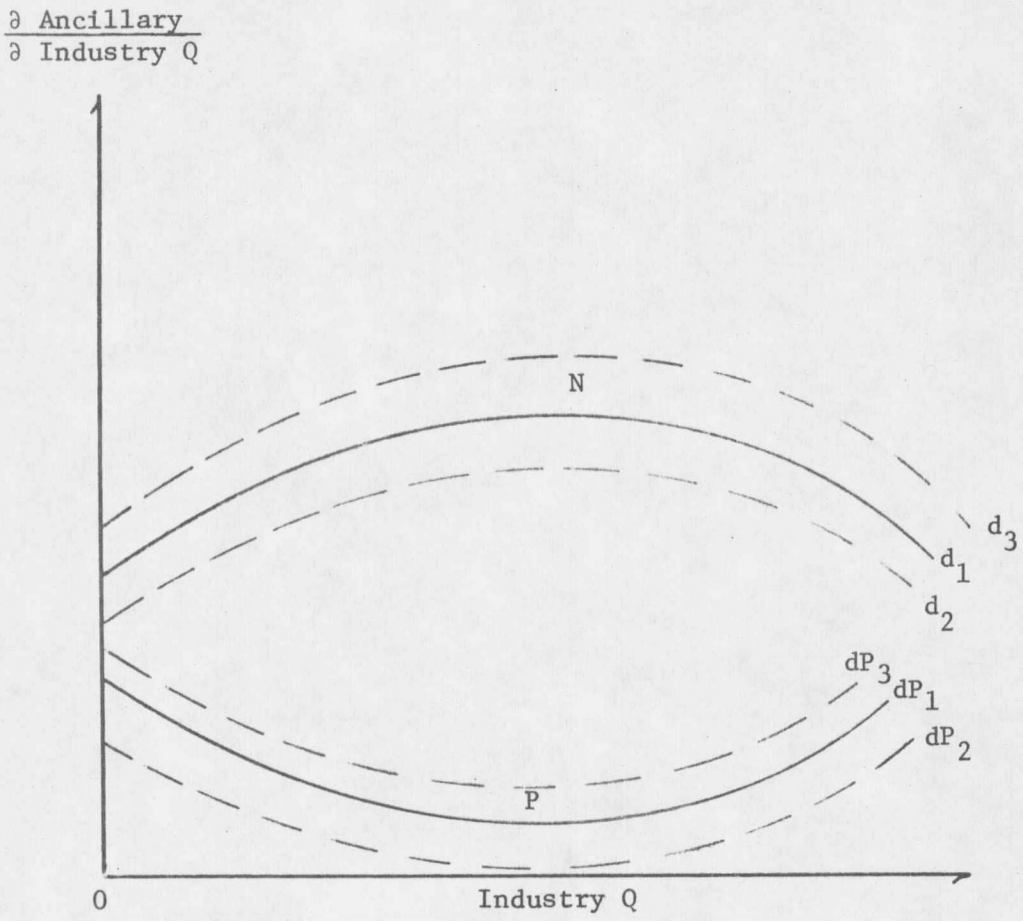


Figure 2. Possible Shapes of the Marginal Multiplier Function Given that Only the Scale of the Activity Varies; Dotted Lines Illustrate the Effect of the Distance Variable.

With distance held constant, the partial derivative in equation (8) can be divided into two parts. One component is those terms which comprise the intercept with the $\frac{\partial \text{ ancillary}}{\partial Q}$ axis, Figure 2. These are the first three terms in equation (8): $\beta_1 + \beta_2 D + \beta_3 D^2$. The second part of the function (terms $2\beta_4 Q + 3\beta_5 Q^2$) defines the curve of the multiplier with respect to scale of activity for a given distance. These two components comprise a simple quadratic which can take on one of two relevant shapes depending on the sign of the $\beta_5 Q^3$ term which after differentiation equals $3\beta_5 Q^2$. 11/ If the $3\beta_5 Q^3$ term is negative, the function is an inverted U, curve N in Figure 2. A positive sign on the $3\beta_5 Q^2$ term dictates a U-shaped curve, curve P in Figure 2. From the signs on the $3\beta_5 Q^2$ terms in Table IX, it can be seen that the agriculture and transportation multipliers plotted against scale of activity will resemble the curve labeled N in Figure 2, while the positive signs on the mining (M^3) and manufacturing (F^3) terms dictate a graph similar to curve P in Figure 2.

11/ In fact, attempting to give meaning to the possible shapes of this function may be pointless. If there is multicollinearity between various terms which comprise the multiplier for each industry, then effects of the individual terms cannot be separated out from the total effect of all terms taken together. The simple r's presented in Table X do suggest possible multicollinearity, but there is no way to be absolutely sure that this is a problem.

TABLE X. SIMPLE R'S GREATER THAN 0.5 FOR AGRICULTURE, MINING, MANUFACTURING, AND TRANSPORTATION TERMS.

	A	DA	D^2A	A^2	A^3	M	DM	D^2M	M^2	M^3	F	DF	D^2F	F^2	F^3	T	DT	D^2T	T^2	T^3
A																				
DA																				
D^2A		.84																		
A^2	.86																			
A^3	.73			.97																
M																				
DM						.89														
D^2M						.68	.89													
M^2						.94	.87	.65												
M^3						.92	.86	.66	.99											
F																				
DF											.57									
D^2F												.72								
F^2											.93	.58								
F^3											.90	.61		.99						
T																				
DT																				
D^2T																	.82			
T^2																.90				
T^3																.83				.98

Interpreting the negatively signed cubic term (curve N in Figure 2) is straightforward. The upward sloping portion of the curve might represent service sectors which are at capacity. Increases in export employment require larger increases in ancillary employment (diseconomies of scale in the service sectors). The decreasing section of the curve could be indicative of economies of scale in those sectors comprising ancillary employment.

The meaning of curve P in Figure 2 is unclear. The non-increasing portion of the curve could indicate existing excess capacity in the service sectors. A dry cleaning establishment, for instance, might have capacity to handle several new customers without expanding their labor requirements. However, this would not explain a decrease in the multiplier for increasing export employment but rather would indicate a horizontal linear relationship out to the point where the curve turns up. The upward sloping segment of the curve is also difficult to interpret. This section would indicate diseconomies of scale, when in fact one would expect the opposite.

The problem of interpreting the curves in Figure 2 is compounded by the computational method used to fit the data. The algorithm forces the curve to follow the specified functional form. Some portions of the curves representing the cubic terms may be beyond the range of the observations. The fact that three or four cubic terms are not significant could point to this explanation.

The distance terms in equation (8), $\beta^2 D$ and $\beta_3 D^2$, shift the industry scale curves either up or down in parallel fashion. These shifts are illustrated in Figure 2. The entire curves will lie in a different position depending on the distance used in the calculations.

The multipliers calculated for mean distance and mean industry sizes are shown in Table XI. The signs are all correct and all but the transportation figure seem reasonable. The large multiplier for transport employment implies that railroad service centers generate many close local linkages. Railroad repair does not use large volumes of imports and most of the ongoing expenditures are for local labor. This also implies a large multiplier in contrast to a manufacturing activity, for instance, which imports intermediate goods.

As an example of the multiplier, assume the number of manufacturing employees in a county located 100 miles from a major trade center is increased from 200 to 400. The employment multiplier is calculated from the expression remaining after the total regression equation is differentiated with respect to manufacturing. In equation form:

$$\frac{\partial \text{ ancillary}}{\partial \text{ manufacturing}} = \beta_1 + \beta_2 D + \beta_3 D^2 + 2\beta_4 F + 3\beta_5 F^2. \qquad (9)$$

Substituting the assumed values for D and F and using the appropriate regression coefficients from Table IX:

TABLE XI. PLAINS AREA: MULTIPLIERS AT MEAN VALUES, 1970.

Variable	Mean	Multiplier		
		Intercept	Slope	Total
Agriculture (A)	744	-0.284	1.34	1.06
Mining (M)	62	3.820	-0.9410	2.88
Manufacturing (F)	212	2.811	-0.457	2.35
Transportation (T)	36	6.26	-1.24	5.02

The mean distance from a Rand McNally trade center = 83.7 miles.

$$\frac{\partial \text{ ancillary}}{\partial \text{ manufacturing}} = (6.719) - (.0718)(100) + (.0003)(100)^2 - (.0011)(400) + (.0196 \times 10^{-5})(400)^2$$

$$\frac{\partial \text{ ancillary}}{\partial \text{ manufacturing}} = 2.13. \tag{10}$$

Total employment in the county due to the 400 manufacturing employees would be:

$$400 + (2.13)(400) = 1,324. \tag{11}$$

Estimating the employment with the original 200 manufacturing employees is done in the same way:

$$\frac{\partial \text{ ancillary}}{\partial \text{ manufacturing}_{t-1}} = (6.719) - (.0718)(100) + (.0003)(100)^2 - (.0011)(200) + (.0196 \times 10^{-5})(200)^2$$

$$\frac{\partial \text{ ancillary}}{\partial \text{ manufacturing}_{t-1}} = 2.33. \tag{12}$$

Using this estimation,

$$200 + (2.33)(200) = 670.$$

The total increase in employment in the county due to the added 200 manufacturing employees is 654 (1,324 - 670). Of these, 454 are service and support employees needed to meet the demands generated initially by the additional manufacturing employment.

The marginal multiplier calculated from the partial derivative is a point estimate of the total function defined by the employment model. Using the model to determine multipliers is not an optimum strategy and is done here only for the sake of illustration. The application of the total estimating equation is the best use of the model. The utilization of the entire equation to determine ancillary employment is illustrated and compared to the multiplier approach in Chapter V.

Summary

This chapter tested a disaggregate economic base model in which local multipliers varied by sector, scale of industry, and by community. Ordinary least squares regression techniques are used to estimate the equation. The data matrix consisted of 181 rural counties in the Northern Great Plains region of the United States.

The model is specified in a manner intended to overcome some of the major criticisms of the economic base method as it has been applied in the past. Standard economic base models calculate one multiplier for all industries. To avoid the structural difficulties inherent in this technique, the model in this study calculates a multiplier for each export industry. Prior economic base studies report a single multiplier that is used irrespective of the scale of the activity. The employment multipliers in this model are formulated so that they can vary depending on the size of the industry, thus avoiding this problem.

The statistical results of the model are encouraging. Fourteen of the 26 variables showed significance at the 95 percent level and the coefficient of determination was 0.97. Further, the signs on all variables were as expected when an *a priori* judgment had been made regarding the outcome.

Application of the multipliers yield point estimates of total employment. These estimates can then be used to make inferences about changes in population and the infrastructure that will be necessary to meet the population's needs.

CHAPTER V

ILLUSTRATION, SUMMARY AND CONCLUSIONS

Illustration

Rosebud County in south-central Montana is used to illustrate the use of the income distribution, migration, and employment multiplier models. Colstrip is located 125 miles from Billings, in the center of a rich, medium B.T.U. low sulfur coal field. The advent of the energy shortage and the ensuing search for energy within the national boundaries of the United States is leading to rapid development in the area. The development will be greatly increased mining and some on-site power generation. The direct employment from these activities is expected to be near 660. 1/ This is permanent employment and does not include transient construction employees.

Prior to coal development Rosebud County was similar to the average of the 181 Plains counties which comprise the observation matrix for this study. Its total population in 1970 was 6,032, down from 6,187 in 1960. About 22 percent of total employment in the county in 1970 was directly agricultural. There were 63 miners and 181 manufacturing employees, none of which were involved in power generation. Only 4.6 percent of the labor force was unemployed, reflecting the impact of coal development in the beginning stages.

1/ This projection was made by Montana Power Company.

Calculating the addition to total employment in Rosebud County, given the 660 new basic employees from the coal development, is necessarily the first step. This estimate is then the input to compute changes in migration flows and income distribution patterns. In Chapter IV, the total predicting equation for ancillary employment was differentiated with respect to the various basic industries and the resultant expressions were used to estimate marginal employment multipliers. This was done to establish a point of reference and enhance comparability for the discussion of the various multipliers, their respective assumptions, the pros and cons of the models, and the characteristics of the approach used in this study. The partials would not be used in actual application of the model, however, because to do so would be discarding information. The marginal multipliers garnered from the partial derivatives are superior to the average multipliers used in prior studies but they are still point approximations of the non-linear function described by the total equation.

The formulation of the total regression equation makes the calculation of multipliers an unnecessary procedure. Instead a four-step process is followed in which the entire regression equation for ancillary employment is applied:

1) Ancillary employment in time t (before development) is estimated using the actual values for the variables in time t and the estimated regression coefficients.

2) Ancillary employment in time t+1 (after the change in basic employment) is estimated analogous to step (1) except that employment in the affected basic industries is augmented.

3) Ancillary employment in time t is subtracted from ancillary employment in time t+1 giving the total change in ancillary employment resultant from development. The change in ancillary is then added to the change in basic employment to give the change in total employment from period t to t+1.

4) Finally, the change in total employment resulting from the development is added to actual total employment in the economy in period t supplying estimated total employment in time t+1.

Applying this procedure to Rosebud County, Montana, furnishes the following results:

1) Ancillary employment in 1970, period t, is estimated to be 1,339.

2) Ancillary employment in time period t+1, around 1976, after the addition of 500 mining and 160 manufacturing employees is estimated to be 2,508.

3) $\text{Ancillary}_{t+1} - \text{Ancillary}_{t} = 1,169$. The change in total employment due to the addition of 660 basic employees is $(660 + 1,169) = 1,829$.

4) Total employment in Rosebud County in 1970 was 2,346. Predicted total employment after coal development is 4,175,

approximately a 78-percent increase.

Although not the ideal application, the ancillary employment regression equation can be used to derive an aggregate employment multiplier. 2/ This multiplier is usually expressed as

$$\frac{\text{basic employment + ancillary employment}}{\text{basic employment}} .$$

For Rosebud County the aggregate employment multiplier would be 2.9 producing a total employment change of (2.9 x 660) = 1,914, or 85 ancillary employees greater than the regression prediction. The multiplier used by the Montana State Department of Planning and Economic Development is 2.7. This multiplier produces a 1,782 change in employment given 660 new basic jobs, a difference of 47 ancillary employees from the number estimated using the regression equation. In this case, the two methods provide extremely close results. The fallacy of using the 2.7 figure is that it is used in all Montana counties regardless of location, industrial base, or scale of activity. To illustrate this problem, an aggregate employment multiplier is calculated for Carbon County, Montana, from the regression equation. The value of the multiplier is about 1.9--vastly different than 2.7. The point is

---

2/ Basic employment is mining, manufacturing, agriculture, and transportation. Ancillary employment is that service or support employment associated with the above four basic industries and it is estimated by the regression equation.

that there are numerous ways to estimate average and marginal employment multipliers, many of which give different results, but none of which contain all the information of the regression model outlined in Chapter IV.

The increase in employment can be used to estimate the change in population. Using a ratio of 2.73 3/ for employment to population yields a population increase of 4,993 given an estimated 1,829 new jobs in the SEA. Of the 4,993 new residents 1,570 will be accounted for by "natural increase," defined as the gross birth rate minus the gross death rate. 4/ The remaining 3,423 persons will then be either new in-migrants or persons who would have migrated out but did not because of the coal development.

The income (I_i) and unemployment (UE_i) variables in the SEA will possibly be altered leading to a reduction in out-migration. Also, the average non-agricultural wage (AW_i) in the SEA may change because

---

3/ The ratio is estimated by regressing the number of employees against total population. The observation units are 181 non-metro Plains counties. The R^2 is .98 and t-value for the employment variable is 118.

4/ The gross birth rate for 1970 in Montana was 18.2 persons per thousand and the gross death rate in 1970 was 9.5 persons per thousand. The rate of natural increase is then the difference of the two which is 8.7 persons per thousand population. The population of Montana SEA#4 in 1970 was 36,182.

of the development, thus affecting in-migration flows. The population in the SEA (P_i) is an independent variable in both models. The regression coefficients on P_i in the two models are nearly equal as expected and the effect of this variable is neutralized.

By using the employment estimate for the coal development it is possible to calculate a portion of the expected in-migration flows. Assuming that all the new jobs are non-agricultural, then the increase in non-agricultural employment will equal the change in total employment which is 1,829. Total change in non-agricultural employment in Montana SEA#4 from 1965 to 1970 was 1,049. The added employment is a 174-percent increase. The regression coefficient for non-agricultural employment from the in-migration model is .2065 percent (remembering that the coefficients are elasticities). This produces a percentage change in the number of in-migrants of:

(.2065%)(174%) = 36% increase.

The total number of in-migrants in the previous period (1965-1970) was 5,649. The expected increase in migration into Montana SEA#4 resulting from 660 new basic jobs in Rosebud County is:

(5,649)(.36) = 2,033 in-migrants.

This leaves 1,390 persons needed to meet the population prediction of 4,993. This amount will, according to the migration models, be

accounted for by a combination of changes in (1) the average non-agricultural wage in Montana SEA#4 (AW_i), (2) the unemployment rate in Montana SEA#4 (UE_i), and (3) median income in Montana SEA#4 (I_i). It is unlikely that the development will significantly affect the median income in the SEA. The new jobs added will be less than 10 percent of the total employment in the agriculturally dominated SEA and many of the incomes generated will be distributed around the 1970 median income in SEA#4 of $6,100. In this case, then, only AW_i and UE_i will be significantly affected. The ranges over which these variables might reasonably vary to induce the population needs are listed in Table XII. The average non-agricultural wage in Montana SEA#4 from 1964 to 1970 was $4,916 and the unemployment rate was 3.9 percent. As Table XII indicates, migrants are considerably more sensitive to job availability than they are to changes in the non-agricultural wage level. 5/ It is assumed that nearly all of the 4,993 population increase will be centered in Rosebud County where the jobs are being created. The severity of this increase is more startling when it is realized that the greatest part of the impact will be felt at Colstrip, the only existing town near the coal mining and proposed generation locations.

5/ The large segment of the total migration flow which is due to a reduction in the number of out-migrants is not unreasonable. It is only a 14 percent decline in total out-migration for the SEA. It is assumed that those persons who do not migrate out of the SEA because of the coal development in Rosebud County will move within the SEA to Rosebud County to avail themselves of the increased employment opportunities.

TABLE XII. LEVELS OF AW_i AND UE_i WHICH INDUCE A 1,390 PERSON POPULATION INCREASE IN MONTANA SEA#4.

Levels of Variables		Migration Effect		Total Change in Population
AW_i	UE_i	AW_i	UE_i	
--Pct.--				
+ 5.0	3.3	67	1,323	1,390
+10.0	3.2	133	1,257	1,390
+20.0	3.1	266	1,124	1,390
+40.0	3.0	532	858	1,390

The number of low-income families can be calculated very easily. Using the appropriate regression coefficients and dividing the increased basic employment into manufacturing and mining employment with values of 160 and 500, respectively, the computations are:

Manufacturing: No. of low-income families = (.2599)(260) = 41

Mining: No. of low-income families = (-.0211)(500) = -10.

A net increase of 31 low income families is due to the increase in basic employment. Note that this estimate does not expressly include the 1,169 service employees. There is no variable in the income distribution equation which accounts directly for this sector of the working force.

To summarize, the addition of 660 new employees, 500 in mining and 160 in manufacturing, will increase total employment in Rosebud County from 2,346 to 4,175. The population will increase from 6,032 to approximately 11,025 with 70 percent of the increase accounted for by migration movements. The implications of such an increase are far reaching. If the ratio of school children to total population is .32 6/ then about 1,600 more school children can be expected. As mentioned,

---

6/ The ratio is estimated by regressing the number of school children against total population. The observation units are 181 non-metro Plains counties. The R^2 is .98 and the t-value for the school children variable is 115.

most of these will be in Colstrip where school enrollment in 1970 was only about 400. Also, one out of every eight new residents will be aged, 7/ assuming the age distribution of the population remains approximately stable. Finally, from the direct employment, 31 low-income families are predicted to enter Rosebud County. This represents a 10 percent increase in the number of low-income families.

In constructing an actual impact statement for Rosebud County the planner could use the estimates of the employment, migration, and income distribution models to produce detailed implications from the coal development. Besides educational costs there are welfare costs from an increased aged population, medical needs, sewer and road requirements, and other important elements of a county or community budget which must be financed through taxation.

Summary and Conclusions

Single equation least squares regression models for employment, in- and out-migration, and income distribution are specified and tested in this study. The observation units are the 181 non-metro Plains counties and only secondary data are used.

7/ The ratio is estimated by regressing the number of aged persons against total population. The observation units are 181 non-metro Plains counties. The R^2 is .92 and the t-value on the aged variable is 44.

The results of the equations are parameter estimators which can be used by planners to project impacts with statistically determined confidence intervals. Using the estimates the planners can determine the change in total employment from a change in the economic base. This prediction can then be used to determine the population change due to migration and the change in the number of low-income families. These estimates have education, welfare, and other tax related infrastructure implications which greatly affect the local community.

The employment model illustrated that the multiplier effect does vary according to industry, scale of operation in the various industries, and location of the activity in economic space. This suggests that the disaggregate least squares model used in this study is a possible alternative to the aggregate economic base and input-output approaches.

The statistical results of the regression equations show each model to be directly related to the economic base of the Plains area. This is an important finding in that it supports the economic theory approach in explaining migration flows and income distribution patterns. Further it allows the results of one model to be directly used as an input into the other models.

The models developed in this study are all static by design. The time which will lapse before the total impact is realized cannot be determined systematically. This is a major question which has not been attacked here. However, the ease with which the models can be updated,

due to the use of inexpensive secondary data, will enable the recalculation of the impacts at various intervals of development, thus to some degree, eradicating the time problem.

Further research is needed in several aspects of the models. Locating better measures of variables which more accurately pinpoint the desired relationship would enhance the predictive power of all three models. Also, defining the models with the same variables where possible would be helpful in making a systematic impact analysis.

The distance of the activity from a regional trade center is an important determinant of ancillary employment. Because of this, further research should be done in determining the location and span of regional trade centers rather than using the Rand McNally designations. To augment the employment predictions the employment model could be converted to an income model thus providing the planner with both the multiplier effects normally associated with impact studies.

LITERATURE CITED

1 Petto, Anthony C. "An Economic Analysis of Migration for the Ozarks Region, 1950-1970." Unpublished Ph. D. dissertation, University of Missouri, 1972.

2 Raimon, Robert L. "Interstate Migration and Wage Theory." Review of Economics and Statistics, XLIV (November, 1962), pp. 428-438.

3 Fabricant, Ruth A. "An Expectational Model of Migration." Journal of Regional Science, X (April, 1970), pp. 13-24.

4 Greenwood, Michael J. "An Analysis of the Determinants of Geographic Labor Mobility in the United States." Review of Economics and Statistics, LI (May, 1969), pp. 189-193.

5 Rogers, A. B. "A Regression Analysis of Interregional Migration in California." Review of Economics and Statistics, LIII (June, 1971), pp. 262-267.

6 Nelson, Philip. "Migration, Real Income, and Information." Journal of Regional Science, I (Spring, 1959), pp. 43-79.

7 Sjaastad, Larry A. "The Costs and Returns of Human Migration." Journal of Political Economy, Supplement Issue, LXX (October, 1962), pp. 80-93.

8 Tolley, George S. "Population Adjustment and Economic Activity: Three Studies." Papers and Proceedings of the Regional Science Association, XI (Summer, 1963), pp. 85-89.

9 Miller, Edward. "Is Out-Migration Affected by Economic Considerations?." Southern Economic Journal, XXXIX (January, 1973), pp. 396-405.

10 Hathaway, Dale E. "Migration from Agriculture: The Historical Record and Its Meaning." American Economic Review, L (May, 1960), pp. 379-391.

11 Parr, John B. "Outmigration and the Depressed Area Problem." Land Economics, XXVI (May, 1966), pp. 149-159.

12 Sjaastad, Larry A. "The Relationship Between Migration and Income in the United States." Papers and Proceedings of the Regional Science Association, LXVIII (Spring, 1960), pp. 37-64.

13 Tarver, James D. "Metropolitan Area Intercounty Migration Rates: A Test of Labor Market Theory." Industrial and Labor Relations Review, XVIII (January, 1963), pp. 213-223.

14 Wingo, Lowdon. "Issues In A National Urban Development Strategy for the United States." Urban Studies, IX (February, 1972), pp. 3-27.

15 Morgan, James. "The Anatomy of Income Distribution." The Review of Economics and Statistics, XLIV (August, 1962), pp. 270-283.

16 Clossom, C. C. "Some Social Applications of the Doctrine of Probability." Journal of the Political Economy, VIII (March, 1899), pp. 204-237.

17 Champernoure, D. G. "A Model of Income Distribution." Economic Journal, LXIII (June, 1953), pp. 318-351.

18 Rutherford, R. S. G. "Income Distribution: A New Model." Econometrica, XVIII (July, 1955), pp. 425-440.

19 Goldsmith, S. "Size Distribution of Income Since the Mid-Thirties." Review of Economics and Statistics, IIIVI (February, 1954), pp. 7-32.

20 Kuznets, Simon. "Distribution of Income by Size." Economic Development and Cultural Change, XI (January, 1963), pp. 1-80.

21 Soltow, Lee. "The Distribution of Education, Age, and Occupation." Review of Economics and Statistics, XLII (November, 1960), pp. 450-453.

22 Lebergott, Stanley. "The Shape of the Income Distribution." American Economic Review, XLIX (June, 1959), pp. 328-347.

23 Mincer, Jacob. "Investment in Human Capital and Personal Income Distribution." The Journal of Political Economy, LXVI (August, 1958).

24 Kuznets, Simon. "Economic Growth and Income Inequity." The American Economic Review, XLV (March, 1955), pp. 1-28.

25 Friedman, Milton. "Choice, Chance, and the Personal Distribution of Income." The Journal of Political Economy, LXI (August, 1953), pp. 277-290.

26 Budd, Edward C. Inequality and Poverty. New York: Norton, 1967.

27 Booth, E. J. R. "Interregional Income Differences." Southern Economic Journal, XXXI (July, 1964), pp. 44-51.

28 Schultz, R. R. "On the Measurement of Income Inequity." American Economic Review, XLI (March, 1951), pp. 107-122.

29 Adams, F. G. "The Size of Individual Incomes: Socio-Economic Variables and Chance Variations." Review of Economics and Statistics, XL (November, 1958), pp. 390-398.

30 Mayer, Thomas. "The Distribution of Ability and Earnings." The Review of Economics and Statistics, XLII (May, 1960), pp. 189-195.

31 Verway, David I. "A Ranking of States by Inequity Using Census and Tax Data." Review of Economics and Statistics, XLVIII (August, 1966), pp. 314-321.

32 Budd, Edward C., and Radner, Daniel B. "The OBE Size Distribution Series: Methods and Tentative Results for 1964." American Economic Review, Papers and Proceedings, (May, 1969), pp. 435-449.

33 Thurow, Lester C. "Analyzing the American Income Distribution." American Economics Review, LX (May, 1970), pp. 261-269.

34 Olsen, Erling. "Regional Income Differences: A Simulation Approach." Regional Science Association, Papers and Proceedings, XX (1968), pp. 7-17.

35 Wonnacott, R. J., and Wonnacott, T. H. Econometrics. New York: John Wiley & Sons., Inc., 1970.

36 Purcell, Edwin J. Calculus With Analytic Geometry. New York: Appleton-Century-Crofts, 1965.

37 Draper, N. R., and Smith H. Applied Regression Analysis. New York: John Wiley & Sons, Inc., 1966.

38 Tiebout, Charles M. The Community Economic Base Study. New York: Committee for Economic Development, 1962.

39 Park, Se-Hark. "Least Squares Estimates of the Regional Employment Multiplier: An Appraisal." Journal of Regional Science, X, No. 3 (1970), p. 365.

40 Bohm, R. A., and Lord, J. H. "Regional Economic Simulation Modeling, The TVA Experience." Paper presented at Northeast Regional Science Association Meeting. University Park, Penn., April 14-15, 1972.

41 Hamilton, H. R. et al. A Dynamic Model of the Economy of the Susquehanna River Basin. Columbus, Ohio: The Battelle Memorial Institute, 1966.

42 Thompson, G. E. "An Investigation of the Local Employment Multiplier." Review of Economics and Statistics, XLIV (1959), p. 61.

43 Bender, Lloyd D., Green, B. L., and Campbell, R. R. "Ghettos of Poverty in the Ozarks." Planning, the ASPO Magazine (August, 1973), p. 13.

44 Bender, Lloyd D. "Problems Associated With Using Location Quotients in Economic Base Analysis." Unpublished memorandum of the Rural Development Service, USDA (October 12, 1973).

45 Braschler, C. "A Comparison of Least-Squares Estimates of Regional Employment Multipliers with Other Methods." Journal of Regional Science. XII, No. 3 (1972), p. 427.

46 Hildebrand, C., and More, A. "The Employment Multiplier in an Expanding Industrial Market: Los Angeles County, 1940-47." Review of Economics and Statistics, XXXII (1950), p. 241.

47 Weiss, S. J., and Gooding, E. C. "Estimation of Differential Employment Multipliers in a Small Regional Economy." Land Economics, LXIV (1968), p. 235.

48 Bender, Lloyd D. "Relationships Between Ancillary Service Employment and Agriculture, Mining, and Manufacturing in 52 Western Coal Producing Counties." Unpublished mimeo exploratory study conducted by the Rural Development Service, USDA, (October, 1973).

49 Lansing, J., and Mueller, E. The Geographic Mobility of Labor. Ann Arbor, Michigan: Survey Research Center, Institute for Social Science, 1967.

50 Billings, Bruce R. "The Mathematical Identity of the Multipliers Derived from the Economic Base Model and the Input-Output Model." Journal of Regional Science, IX, No. 5 (1969).

51 Bender, Lloyd D. "Ancillary Employment Multipliers--Four Provinces of the Western U. S." Memorandum of Research Progress, ERS/USDA, Montana State University, Bozeman (September, 1974).

52 U. S. Department of Commerce. Bureau of the Census. Social and Economic Statistics Administration. "Characteristics of the Population, Wyoming." I, Pt. 52. Washington, D. C.: Government Printing Office, 1970.

53 _____. "Characteristics of the Population, Montana." I, Pt. 28. Washington, D. C.: Government Printing Office, 1970.

54 _____. "Characteristics of the Population, North Dakota." I, Pt. 36. Washington, D. C.: Government Printing Office, 1970.

55 _____. "Characteristics of the Population, South Dakota." I, Pt. 43. Washington, D. C.: Government Printing Office, 1970.

56 _____. "Characteristics of the Population, Colorado." I, Pt. 7. Washington, D. C.: Government Printing Office, 1970.

57 _____. "State Economic Areas." PC(2)-10B. Washington, D. C.: Government Printing Office, 1970.

58 _____. "Migration Between State Economic Areas." PC(2)-2E. Washington, D. C.: Government Printing Office, 1970.

MONTANA STATE UNIVERSITY LIBRARIES
3 1762 10014699 0

N378
L583
cop.2

Lewis, Eugene P
Empirical least squares regression models for employment ...

DATE	ISSUED TO
APR 17 '75	INTERLIBRARY LOAN

N378
L583
cop2