12-1971

Time lags in the impact of public investment in water resources: The Tennessee Valley Region, 1936-1968

Audley Eugene Hileman

To the Graduate Council:

I am submitting herewith a dissertation written by Audley Eugene Hileman entitled "Time lags in the impact of public investment in water resources: The Tennessee Valley Region, 1936-1968." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Agricultural Economics.

Joe A. Martin, Major Professor

We have read this dissertation and recommend its acceptance:

Charles L. Cleland, Hans E. Jensen, Charles B. Sappington

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

August, 1971

To the Graduate Council:

I am submitting herewith a dissertation written by Audley Eugene Hileman entitled "Time Lags in the Impact of Public Investment in Water Resources: The Tennessee Valley Region, 1936-1968." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Agricultural Economics.

Joe A. Martin
Major Professor

We have read this dissertation
and recommend its acceptance:

Charles Sappington

Hans T Jensen

Charles L. Cleland

Accepted for the Council:

Hilton A. Smith
Vice Chancellor for
Graduate Studies and Research

TIME LAGS IN THE IMPACT OF PUBLIC INVESTMENT IN WATER RESOURCES: THE TENNESSEE VALLEY REGION, 1936-1968

A Dissertation

Presented to

the Graduate Council of

The University of Tennessee

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Audley Eugene Hileman

December 1971

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to the many people who have made completion of this dissertation possible. Dr. Joe A. Martin, Dr. Charles L. Cleland, Dr. Hans E. Jensen and Dr. Charles B. Sappington, the author's graduate committee, have been most gracious with their time. A special word of gratitude is in order for Dr. Martin, the chairman of the committee, for his numerous hours of reading, his undying patience, and his keen insight. The author's peers at the University of Tennessee have also been most helpful. He has benefited a great deal from their comment and discussion. Dr. Thomas J. Whatley and the Department of Agricultural Economics at the University of Tennessee have given financial support to the research.

The author was able to undertake further graduate study as the result of financial support from Maryville College during a study leave from that institution. Dr. Frank D. McClelland, the author's fellow faculty members and, perhaps most of all, his students at Maryville College have been a never ending source of encouragement and motivation.

The author's parents, Mr. and Mrs. William Jaye Hileman, have contributed as only parents can. The greatest source of strength and encouragement have doubtless come from the author's wife Leila and their children Will, Leila Bleecker and Robbie. Without their patience and love the project would never have been completed.

ABSTRACT

The purpose of this study was to investigate the time dimension of the impact of public investment in natural resources upon the economy of urban centered areas. The specific objectives were: (1) to evaluate the effectiveness of estimating distributed lag functions as a technique for studying the time lag associated with the economic impact of public investment in water resources; and (2) to estimate the distribution over time of the impact of investment in water resources upon the economy of urban centered areas by estimating a rational distributed lag function.

The procedure followed was to estimate parameters for a model relating income in urban centered areas to agricultural employment, nonagricultural employment, the national employment rate and annual public investment in water resources. Separate estimations were made using two different measures of public investment in water resources. Based upon the results of estimates using this income model a second model was developed. The second model related manufacturing employment in the urban centered areas to the national employment rate and the cumulative value of all past public investment in water resources in the respective areas. The economic units of observation were the urban

centered areas around Huntsville, Alabama; Paducah, Kentucky; and Asheville, North Carolina.

The two measures of public investment in water resources in the income model were annual public investment in dams and steam plants in the respective areas and annual public investment in dams across the entire TVA region. The estimates failed to show any statistically significant or consistent relationship between income in the areas and either measure of public investment.

A reasonable distributed lag function was obtained for the Paducah area. The distribution had a mean of 5.3 years and a variance of 31.4 years. It was difficult to interpret the lag distribution since the overall impact of annual investments was so weak. The findings of the income model lead to the rejection of the popular notion that public investment in water resources has significant direct impact on area income as a result of the usual income multiplier effects. An alternative hypothesis was developed that the impact of public investment in water resources resulted in increased social overhead capital and consequent increased flow of services available in the area. To examine this alternative hypothesis the manufacturing employment model was used.

Parameters relating manufacturing employment to the cumulative value of public investment in dams and steam plants were quite small. For Huntsville an addition of one million dollars to the cumulative value of public investment in water resources was associated with eight additional jobs in manufacturing; for Paducah, two additional jobs in manufacturing; and for Asheville, one less job in manufacturing. The estimated parameters were statistically significant for Huntsville and Paducah but not for Asheville.

Reasonable distributed lag functions were obtained for the Huntsville and Paducah areas in the manufacturing employment model. The lagged impact for Huntsville was slower and more dispersed than for Paducah. The mean lag for the Huntsville area was 8.6 years and the variance of the lag distribution was 78 years. For the Paducah area the mean lag was only 5.2 years and the variance was 32.9 years.

The analysis provides the basis for several tentative conclusions which are especially helpful in pointing directions in which further research might be most fruitful.

These conclusions are:

(1) The impact of public investment in water resources is apparently the result of increased social overhead capital and its consequent increased flow of services rather than regional income multiplier effects of the public expenditure.

(2) The complex matrix of factors causing economic advance of urban centered areas is so different for different areas that public investment in water resources has different degrees of impact upon the different areas. Also, the time dimensions of the impact for different areas vary rather widely.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Government involvement in natural resource development is not a new phenomenon in the United States. Even the dramatic step of establishing the Tennessee Valley Authority (TVA) is now nearly 40 years old. Many of the public ventures, such as TVA, have involved multi-million dollar investments in the expectation that various types of individual and social payoffs would accrue. Although most of the larger projects have been justified *ex ante* (most generally using the frequently discussed cost-benefit analysis), relatively little evaluation has been attempted *ex post*. As late as 1967 Stefan H. Robock stated that "A thorough evaluation is one of the major unfinished tasks related to the TVA experiment." (22, p. 20) The same could doubtless be said for other Federally supported projects. It is intended that the present study will be a contribution toward meeting this need.

Purpose and Objectives

The purpose of this study was to investigate the time dimension of the impact of public investment in water resources upon the economy of urban centered areas. It is one phase of a broader study evaluating the impact of public investment in water resources. The specific objectives of the study were:

1. To evaluate the effectiveness of estimating distributed lag functions as a technique for studying the time lag associated with the economic impact of public investment in water resources.
2. To estimate the distribution over time of the impact of investment in water resources upon the economy of urban centered areas by estimating a rational distributed lag function.

Justification

The results of the study should be helpful in at least two important areas. First, they should be of assistance in evaluating the overall effectiveness of multipurpose water resource development projects like TVA as a means of influencing rates of economic advance in other regions of the United States as well as in other areas of the world. Knowledge of the time dimension of the impact of alternative types of projects is extremely important information in choosing among alternatives.

The second area where the results of the study should be beneficial is in cost-benefit analysis of proposed projects. One of the chief areas of concern and controversy surrounding cost-benefit analysis is associated with the secondary benefits of projects--specifically, benefits associated with economic growth. A point of focus in cost-benefit is the present value of future benefits; therefore, many projects which may be worth while if the impact is rapid may not be economically sound if the impact is spread over many years or only occurs after a considerable lapse of time. This study should give insight into the very important time dimension of growth in employment and income induced by public investment in water resources.

Previous Attempts at Evaluation

It was pointed out above that very little research aimed at ex post evaluation of public investment in water resources has been done. Furthermore, work which has been done is frequently narrow in scope and limited in content. The primary value of the previous research is the insight it gives into different approaches to the task.

The most common approach taken in previous investigations is to compare the behavior of relevant variables of the impacted area with some control or check area. Examples of this approach include Ratchford (21), the President's Water Resource Policy Commission (20) and Zimmerman (38) and to a somewhat different degree Chappell (5) and Wiebe (37).

In "Government Action or Private Enterprise in River Valley Development: An Economist's View" Ratchford (21) attempted to determine whether or not a differential rate of economic expansion took place under private development or concentrated government action. In his analysis he compared various measures of economic growth in Tennessee (a proxy for the impact area of TVA) with a grouping of nine southern states--Virginia, North Carolina, South Carolina, Oklahoma, Georgia, Florida, Arkansas, Louisiana, and Texas. He compared changes in population, income payments, production workers, wages paid, value added in manufacturing, resources of all reporting banks, and retail sales for the two groups from 1929-30 to 1947-50 (years varied slightly due to data availability). His general conclusion was that "There is nothing in these figures to indicate that TVA has been a major factor in affecting the economic activity of the state in which most of its activities have been centered." (21, p. 306) His data shows that Tennessee grew more rapidly than the United States as a whole, but was in line with the other Southern states.

In a similar comparison in A Water Policy for the American People the President's Water Resources Policy Commission concludes that "[t]he contribution of this product of falling water to the increase in jobs and income is indicated by . . . data relating to manufacturing and employment." (20, p. 35) They then compare manufacturing establishments in the 122 TVA valley counties[1] with the entire seven valley states from 1935 to 1939; also, they compare manufacturing employment in the valley with the Nation as a whole for the same period. Both comparisons show TVA

---

[1]TVA presently designates 125 counties as valley counties and 170 counties as its power service area. Since these are not mutually exclusive groups 201 counties are in either one or the other group. The valley counties are those in which TVA has the widest range of responsibilities. The number of counties in this area has changed from time to time and was 122 in 1950.

in a very favorable light. (One wonders if failure to make parallel comparisons on both series is due to lack of data or failure of the editor to understand its importance.) Comparisons are also made between 1929 and 1947-48 for income payments to individuals and production workers employed in manufacturing. These comparisons showed the per cent gain in the valley counties ahead of the entire seven valley states and both of these were above the Nation as a whole.

The striking fact about these two sources is that, in spite of similar periods being considered, conclusions differ due to differences in areas and data compared. The difficulty for the reader is in determining whether the selection of the data and regions compared was based on some acceptable objective criteria or whether the researchers chose the comparisons which best supported their theses.

Zimmerman (38) compares the Missouri Valley and TVA regions to show the impact of a unified planning approach to river basin development. This work differs from Ratchford and the President's Water Resource Policy Commission in that the areas were selected on the basis of a common resource rather than geographic proximity. Zimmerman cites comparisons of the performance in the Missouri Valley and Tennessee Valley economies between 1930 and 1945. The comparison shows strikingly better performance in the Tennessee Valley in such categories as changes in acre value, value of farm machinery per acre, population change and farm tenancy. Zimmerman points out that the comparisons of the two valleys cannot be made without qualification since the Missouri Valley is much larger and more heterogeneous than the Tennessee Valley. Nonetheless, the comparison is quite favorable to the unified planning approach of the TVA as compared with the less unified (and less complete) approach for developing the Missouri Valley. It should be pointed out that the Zimmerman data was based upon rather limited research. Its main contribution is in pointing to an approach to evaluation of TVA and similar projects which may be very fruitful.

Although the traditional techniques of scientific analysis cause the control-area approach used in all the above studies to be very appealing, it is subject to serious shortcomings in this particular type of use. First, the approach requires (i.e. assumes) that the difference in the dependent variables are due to differences in some independent variable alone. Put another way, in order for the check area to be satisfactory, the observed or controlled variable must be the only differentiating factor between the areas. Clearly this is not the case in any of the above studies; furthermore, the possibility of finding two areas so similar is extremely unlikely.

A second criticism of the check area approach is that determination of the impacted area is difficult. Data availability generally restricts observation to politically determined units which are frequently not very satisfactory from an economic point of view. Even more basic is the fact that the impact of investment in natural resources may be widespread, is not necessarily greatest in areas immediately adjacent to the project site, and may not even occur in geographically contiguous areas.

One critic of the above approach suggests that the impact of investments in natural resources (e.g. TVA) can best be evaluated by looking at specific industries (16). He cites the chemical based industrial complex which developed in the Paducah, Kentucky, area as an example of industrial development resulting from public investment in water resources. This work has special implications for the present study since it suggests--by inference only--that such repercussions are not predictable in advance and may be erratic in both the time and space dimension. Unfortunately the research stops short of any rigorous analysis. Rather, ideas are suggested, along with some preliminary observations, and the topic is dropped.

A variation on the method which uses check areas was followed by Wiebe. His objective ". . . was to examine the effects of investments in water resources on regional income and employment." (37, p. 163) He examined the hypothesis that the effects of the investment were in the immediate area of the projects and a second hypothesis that the effects of the investments were more spatially diffused. He divided the counties in the Tennessee Valley into those "directly impacted" and those "not directly impacted" by water resource projects. The directly impacted counties were those geographically contiguous with the projects. He examined income, manufacturing investment and education levels. He concluded that per capita incomes increased in the directly impacted counties but increases in employment were associated with the counties at some distance from the investment projects.

Chappell (5), in a companion study to that of Wiebe, concentrated his effort on 150 counties in the TVA Power area during the decade of the 1950's. He attempted to evaluate the effectiveness of public investment in natural resources as a stimulant to economic growth. Using several multivariate techniques, he constructed an index of economic growth and then isolated socio-economic factors which explained the differences in the counties' growth performance during the decade. He found that counties with high growth indexes were, in general, urban counties located along waterways in which substantial investment in natural resources had taken place. It is noteworthy that investment in natural resources was not an important factor in explaining differences in the growth performance of the counties.

Plan of Work

The decision to focus attention upon the urban centered area as a unit of observation follows directly from the results obtained by Wiebe and Chappell. In the next chapter the concept of the urban centered area is examined and the specific areas used in the study are delineated.

The concept and language of the distributed lag function are put forth in Chapter III. Chapter IV completes the groundwork of the research by presenting the model and discussing the special estimating technique which is required to perform the analysis. Chapter V presents the results of the analysis.

CHAPTER II

SELECTION OF URBAN CENTERED STUDY AREAS

The findings of Wiebe and Chappell are hardly surprising in view of what is known about the process of economic activity. Perloff *et al.* (19) state the hypothesis (generally associated with T. W. Schultz) (24) ". . . that closeness to population and industry concentration directly influences the volume, type, and intensity of economic activities." The ramification of the results of Wiebe's and Chappell's research for the present study is the suggestion that the most meaningful unit of observation for analyzing the impact of investment in water resources is the urban centered area. The precedent for this is quite well established. Smith in *The Wealth of Nations* (26) recognizes the important "functional unity" of the region in the growth process. Von Thunen (29) explicitly focused his analysis on the urban centered area. More recently Fox (8) and his colleagues at Iowa State University have given considerable attention to functional economic areas.

From a strictly economic point of view it is sensible to focus on the urban centered area because the activities within the area are highly complementary. The resource markets, for example, encompass the entire area because resources available in one part of the area are usually available in any part of the area. If this is not absolutely true, the arbitrage effect will cause changes within the market to be quickly reflected over the entire area without depending on significant geographic mobility. Naturally, it would be agreed that this is true to some extent for the entire nation. The importance of the urban centered areas, however, is that they are the building blocks from which the national economy is built. The urban center is also the link between the area and the national economy. It is the area market place for imports into the area as well as the market for products exported from the area. The urban center is the hub of the transportation system, the financial center, and the information center which services the area economy.

Urban centers have impact beyond the strictly economic considerations. In the very important social and cultural area the influence of the urban center upon its hinterland is also well documented. Descriptions of cities in the middle ages point out that they were the places where new ideas formed, where political leadership was located, and where new types of economic activity evolved. Probably the most outstanding aspect of cities in the middle ages is that, even though their population was only a small per cent of the total, they were the chief holders of economic and political power as well as the chief carriers of cultural and intellectual activity. (12)

Hoselitz says that cities are "centers which tend to impress their characteristics on the rest of society as it undergoes economic growth." (12, p. 208) This is true now as earlier and in developed as well as underdeveloped countries. Cities are the centers of learning, the source of many new ideas and the focal point of culture. These social forces at work within cities are very important in transforming society. As Davis and Golden put it, "The requirements of urban living force innovations which those in the countryside, if left to themselves, would never make." (6)

Definition of Urban Centered Area

Descriptions and definitions of urban centered areas or central places are nearly as numerous as writers on the subject. One of the most complete and workable lists of attributes of such areas is given in The Role of Growth Centers in Regional Economic Development. Their "growth centers" fit well the desired attributes of the areas chosen for analysis in this study.

> The linkages of such a growth center within the region include the following:
>
> (1) The growth center is the center of a labor market or commuting area;
>
> (2) The growth center is the center of a major retail trade area--for example, it may contain the only full-time department store in an economic development district;
>
> (3) The growth center typically has one or more newspapers, radio stations and TV stations which cover a territory similar to the retail trade area;
>
> (4) The growth center may contain a junior college or community college and other agencies or institutions for training in vocational skills;
>
> (5) Most residents of the economic development district are familiar with the growth center and avail themselves of its more specialized shopping, recreation, and other facilities;

(6) Retailers throughout the economic development district depend upon wholesalers located in the growth center for many of their supplies (13, pp. 5-6).

A noteworthy attribute of the list is that it draws attention to the linkages between the central place and the outlying area. These linkages are important in studying the impact of public investment in water resources since the investment only rarely takes place at the urban center. Rather, it is made in outlying areas. One of the key assumptions of this study has been that the impact of investment can be measured primarily through observation of the urban centered area as a single entity. The polarity of the urban area is the reason for using it as the unit of observation rather than, for example, counties. Data on individual counties may miss completely the impact of the investment; or more likely, they may give only a partial and biased view of the overall picture.

The urban center is not only linked to the surrounding area, it is linked to the national economy as well. This is important since the influence of investment in water resources upon an area is likely to be in direct proportion to the extent to which investment changes the characteristics of the area vis a vis the national economy. Such changes may come about from reducing the cost of industrial inputs (e.g, electric power) or from reducing transportation costs between the area and the national product and raw materials market.

Criteria for Selection of Areas to Be Studied

Several criteria were established for selecting suitable geographic areas for investigating the time dimension of the impact of public investment in water resources. First, investment in water resources must have taken place over a sufficient period to provide reasonable assurance that the impact had "worked itself out" to a degree consistent with reliable estimates. Second, sufficient data must be available on each of the variables which enter the model. Third, areas should conform to the urban based concept described above.

The first criterion is satisfied quite well by many areas within the Tennessee Valley Region. The TVA has been making investments in water resources since the late 1930's. In all likelihood, any systematic impact of such investment will have shown up over the intervening period. TVA is also a valuable source of data on a county basis--making it possible to build up data for multi-county areas. However, their data are not complete and, in the final analysis, data availability was the limiting factor which determined the choices from among the many urban centered areas within the TVA Region.

The Areas Studied

The three areas selected for analysis were Paducah, Kentucky; Huntsville, Florence, Athens, Alabama; and

Asheville, North Carolina. The areas are shown on Figure II-1 and the specific counties included in each area are listed in Table II-1. None of these conforms exactly to the urban based concept set forth above but it is assumed that the discrepancies will not substantially influence the results.

The North Alabama area satisfies the criteria best. Since it did not have one central urban core, it did not exactly meet the criteria. It was included as a single area because it is comprised of contiguous counties, appears to be homogeneous, and is not dominated by any city outside the area. Over the period studied Huntsville has emerged as the dominant center of the area and throughout the study this area is referred to as the Huntsville area.

The Paducah and Asheville areas share a common shortcoming in that only part of the economic areas associated with these cities is included in the study areas. Indeed, as Figure II-1 shows, Watauga County, North Carolina, is not even contiguous with the remainder of the counties included in the Asheville area.

The reason that the complete market area is not included is lack of comparable employment data for all counties. In both the Paducah and Asheville areas the excluded counties lie outside the TVA power region and, in this sense, would not benefit directly from this aspect of TVA investment. Their exclusion should not materially affect the results of the analysis since the included counties comprise fairly complete cross sections of the areas.

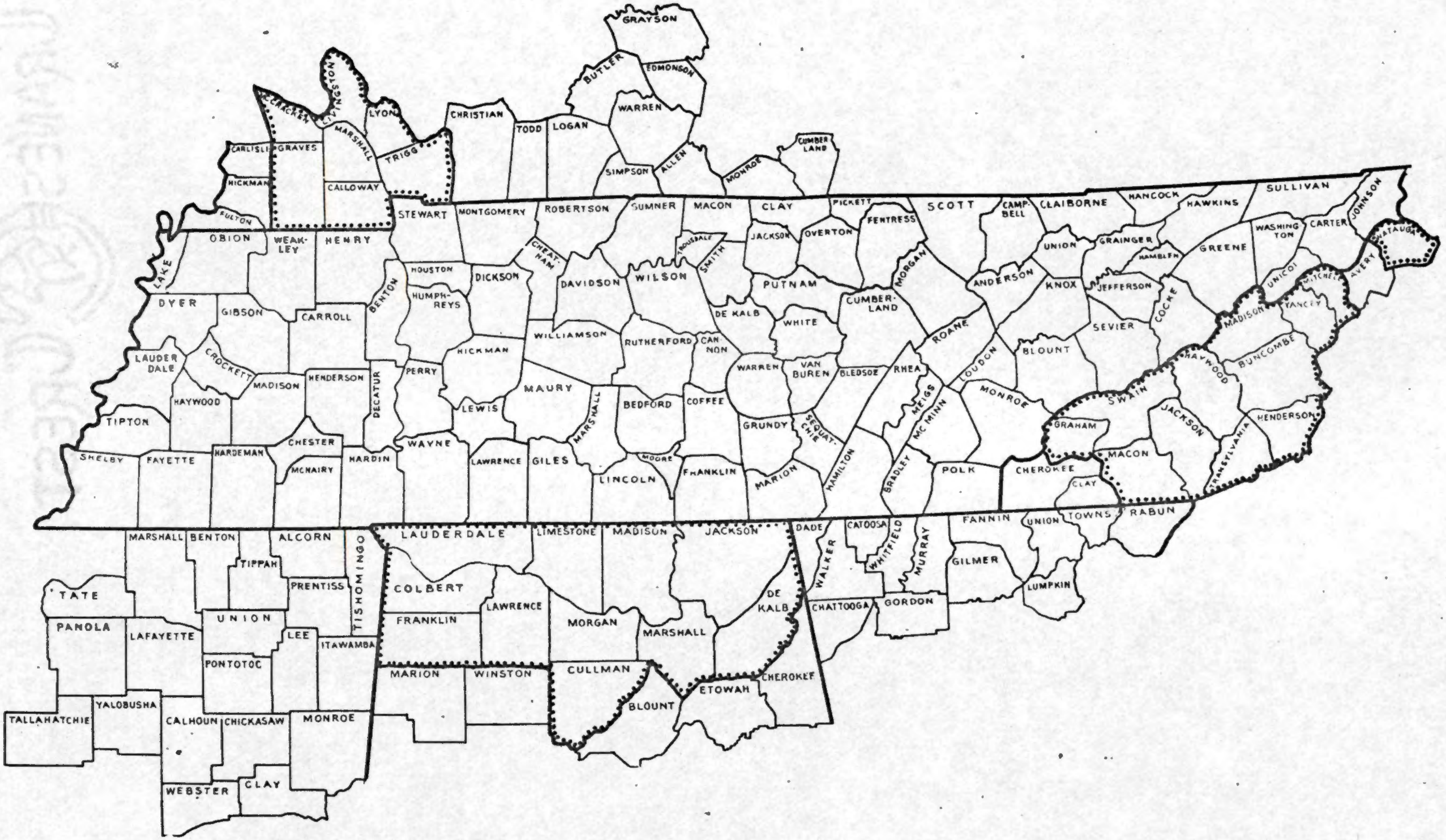

Figure II-1. Counties in Western Kentucky, Northern Alabama, and Western North Carolina comprising the areas studied.

Table II-1. Counties included in the three study areas

Huntsville, Alabama	Paducah, Kentucky	Asheville, North Carolina
Colbert	Calloway	Buncombe
Cullman	Graves	Graham
DeKalb	Livingston	Haywood
Franklin	Lyon	Henderson
Jackson	McCracken	Jackson
Lauderdale	Marshall	Macon
Lawrence	Trigg	Madison
Limestone		Mitchell
Madison		Swain
Marshall		Transylvania
Morgan		Watauga
		Yancey

Measured changes in the relevant variables for the included counties should reflect with considerable accuracy the performance of the entire area.

In addition to the problem of excluded counties in the Asheville area, some of the counties which are included lie at the margin of the area. This is true of the counties in extreme southwestern North Carolina. These counties are so small--especially if measured by the magnitude of economic and demographic variable rather than geographic area--that they should not have a great impact upon the area. Also, it is no doubt true that the primary direction of economic interchange between these counties and a central place is with Asheville.

An additional point concerning the Asheville area is that the TVA investment has been made at locations relatively distant from the urban center. For this reason, as well as those listed above, special care was taken in the analysis to make certain the conclusions concerning this area were not the result of its special characteristics as compared with the other two areas.

Area Profiles

Table II-2 lists data which give an overview of the three study areas. It is clear that the areas were quite different in their makeup in 1936, the beginning of the period included in the study, as well as in the nature of the change made during the subsequent 30 years. Little is

Table II-2. Statistical profile of the areas included in the study

Item	Year	Unit	Huntsville	Paducah	Asheville
Household income	1936	thousand dollars	140810	83132	165154
Household income	1946	thousand dollars	367684	122010	303675
Household income	1956	thousand dollars	478610	214410	385800
Household income	1966	thousand dollars	1163765	280616	610422
Ag. Employment	1936	thousands	100.1	21.1	40.1
Ag. Employment	1946	thousands	67.6	15.8	39.1
Ag. Employment	1956	thousands	50.6	9.7	19.7
Ag. Employment	1966	thousands	24.8	5.7	11.1
Non-Ag. Employment	1936	thousands	28.7	16.1	28.1
Non-Ag. Employment	1946	thousands	42.2	17.7	52.7
Non-Ag. Employment	1956	thousands	65.2	27.4	62.4
Non-Ag. Employment	1966	thousands	129.8	35.3	85.2
Ratio Ag./Non-Ag.	1936		3.49	1.31	1.42
Ratio Ag./Non-Ag.	1946		1.60	.89	.74
Ratio Ag./Non-Ag.	1956		.78	.35	.32
Ratio Ag./Non-Ag.	1966		.19	.16	.13
Education	1936	median years	6.6	7.7	7.4
Education	1946	median years	7.5	8.2	8.3
Education	1956	median years	8.5	8.8	9.3
Education	1966	median years	9.6	9.4	10.4
Capital intensity index	1936		75.4	67.0	107.3
Capital intensity index	1946		90.5	83.7	107.0
Capital intensity index	1956		108.5	104.5	106.8
Capital intensity index	1966		130.2	131.4	106.5
Daily Newspaper circulation from central urban city	1970	papers	53,363	30,987	70,046

to be gained by extensive summarizing of the table but some observations should be made.

Asheville and Paducah seem to have had about the same ratios of agricultural to non-agricultural employment over the period but the nature of the industrial mix changed more drastically in Paducah as indicated by the capital intensity index. (Computation of this index is explained fully in Appendix A. Essentially it reflects the capital intensity of the industrial mix in the area.) Huntsville seems to have made the most dramatic transformation from an agrarian to industrial economy. It had a much higher ratio of agricultural to non-agricultural employment in 1936 than the other two areas. Even in 1956 this was still true but in the next decade the difference was virtually eliminated.

The fact that Huntsville is not the only urban place in the area would explain the relatively low newspaper circulation figure reported there. Huntsville, like Asheville and Paducah, has a television station and is the communication center of the area.

CHAPTER III

DISTRIBUTED LAGS AND ESTIMATING TECHNIQUES

One of the objectives of this study is to estimate a distributed lag function for the impact of public investment in water resources upon economies of urban centered areas. The two-fold purpose of this chapter is to discuss the terminology of the distributed lag models and to set out the type of distributed lag model used in the study. A complete review of the literature on distributed lags is not included since this has been done so well by Griliches (11). Bauer (2) also gave an excellent summary. Because he is somewhat less inclusive than Griliches, Bauer's summary is frequently easier to follow. The material in this chapter relies heavily upon these two sources.

The Concept of Distributed Lags and a Lag Generating Function

Consider two variables related in a manner such as

(1) $$Y_t = \sum_i B_i X_{t-i} \qquad (i = 0, 1, 2, \ldots ,)$$

where X is an independent or predetermined variable and Y is a dependent variable. The subscript t refers to the time period.[1] Such an equation has been estimated using ordinary least squares regression following the procedure of Alt (1) who suggested that lagged B_i's be estimated until their signs become erratic.

Among the problems with this procedure is that a large number of parameters must be estimated. In order to overcome this problem Koyck (15) estimated a geometric lag distribution requiring only one parameter more than a non-lagged equation. In order to maintain the advantage of a small number of parameters but to gain more generality, use of lag generating functions has come into use (27, 14). Estimation of lag generating functions requires that all B_i's are of the same sign and have a finite sum.

If this is the case we could write the relationship

(2) $$Y_t = B(v_0X_t + v_1X_{t-1} + v_2X_{t-2} + \ldots)$$

If $w_i = \dfrac{v_i}{\sum_{i=0}^{\infty} v_i}$ we could write

(3) $$Y_t = B(w_0X_t + w_1X_{t-1} + w_2X_{t-2} + \ldots +)$$

In this case $0 \leq w_i \leq 1$ for all w_i's. In the form used in equation (3) the w_i's can be identified with a probability

---

[1] In this chapter the constant term and the random term are not written in the equations. This is for convenience only and they are, of course, part of the actual model estimated in the study.

distribution and a probability generating function can be utilized in estimating the w_i's.

Griliches points out the advantage of being able to use the concept of a probability generating function. These are (1) the function

> . . . may reduce itself to a relatively simple algebraic form even though the sequence [w_i] is quite complicated; (2) one can get the parameters (moments) of the probability distribution directly as functions of the derivatives of the generating function; and (3) it facilitates greatly the derivation of distributions of sums of random variables, or for our purposes, the cascading or convolution of several lag distributions (11, p. 19).

The first two of these advantages have proven to be most important in this study.

Lag Operator Notation

Nearly all recent discussions of distributed lags make use of the lag operator notation. Before proceeding with a discussion of the particular lag generating functions used in the study, it is necessary to introduce this notation.

Let

$$LX_t = X_{t-1}$$

$$L^2X_t = X_{t-2}$$

$$L^3X_t = X_{t-3}$$

where L is the lag operator and X is some variable.

If we make use of this notation and rewrite equation (3) we have

(4) $Y_t = B(w_0X_t + w_1LX_t + w_2L^2X_t + . . .)$

or in a more general form

(5) $Y_t = BW(L)X_t$

where $BW(L) = B \Sigma W_iL^i$ $(i = 0, 1, 2, . . . , \infty)$

This is an infinite power series in L and can be treated as such in performing mathematical operations. Many such series can be written in the closed form. A generating function which can be written in this manner is called a rational generating function ("i.e. it can be factored into a ratio of two finite polynomials") (11, p. 22). The rational generating function estimated in this study followed the technique suggested by Jorgenson (14) as an improvement upon the more restrictive techniques of Koyck (15) and Solow (27).

The Jorgenson Family of Distributed Lag Functions

The significant characteristic of the Jorgenson procedure as opposed to earlier techniques (Koyck and Solow) is that he did not constrain the lag polynomial. He proved that the rational form $Y_t = \frac{N(L)}{D(L)} X_t$ can be used to approximate any arbitrary lag function W(L). As a practical matter it is not useful for the polynomial D(L) to be of a higher order than two or three. In higher order polynomials the parameters to be estimated would be so great that one would do just as well to go back to estimating the w_i's directly as in equation (1).

The objective in estimating the distributed lag function is to obtain a lag distribution which best reflects the relationships existing in the data. The procedure followed in the study was similar to that followed by Bauer. In this procedure a "family" of second order lag functions is estimated and the one that ". . . in some sense best fits the data" (2, p. 28) is identified. Bauer identified the best distribution based upon the coefficient of determination of the estimated equation, the signs and magnitudes of the structural parameters estimated, and the reasonableness of the lag distribution.[2] In this study most of the equations yielded distributions which were not reasonable in that they had negative weights or were nonconvergent.

The family of functions P(L) estimated are listed in Table III-1. In line with the suggestion above, these functions are limited to first and second order polynomials for D(L) (the denominator) and zero, first and second order polynomials for N(L) (the numerator). In one of the models

---

[2]Reasonableness in the sense used here and in the remainder of the study refers to the mathematical properties of the distribution and not the reasonableness of the distribution relative to some a priori judgment of its expected or possible forms. A distribution is said to be reasonable if it has no negative weights and converges to zero. This assures that weight (w_i) can be computed so that

$$0 \leq w_i \leq 1 \text{ and } \sum_{i=0}^{\infty} w_i = 1.$$

Table III-1. The family of second order lag functions estimated

Function number	N(L)	D(L)
2	1	$1 - a(L) - b(L)^2$
3	$1 + a(L)$	$1 - b(L) - c(L)^2$
4	$1 + a(L) + b(L)^2$	$1 - c(L)$
5	$1 + a(L) + b(L)^2$	$1 - c(L) - d(L)^2$

used in the study suitable data were not available for estimating the entire family of lag functions and only lag configuration three was used.

One of the problems with the Jorgenson technique, discussed by Griliches (11), and encountered in this study, is that the parameters estimated may not give reasonable lag distributions. Consider the most simple form estimated in this study:

(6) $$Y_t = B\left(\frac{1}{1 - aL - bL^2}\right)X_t$$

(7) $$Y_t(1 - aL - bL^2) = BX_t$$

(8) $$Y_t = BX_t + aY_{t-1} + bY_{t-2}$$

Equation (8) is the form estimated. (Recall that the constant is omitted from the equations in this chapter.) The estimates of this equation will not generally yield weights that are less than zero and sum to unity. For this reason it is merely an intermediate step in the computation of the desired weights. If the polynomial resulting from the estimation is designated as P(L) then the actual lag distribution is $P(L)/\sum_{i=0}^{\infty} P_i$. It has been shown (2, pp. 39-40) that

$$\sum_{i=0}^{\infty} P_i = \frac{1}{1 - a - b}\text{; therefore,}$$

$$W(L) = P(L) \cdot (1 - a - b)$$

for the special case of P(L) being considered.

In order for the resulting lag distributions to be stable the following restrictions must be met:

$$0 < |\lambda_1| < 1$$

$$0 < |\lambda_2| < 1$$

In order for it to have all positive weights

$$0 < \lambda_1 < 1$$

$$0 < \lambda_2 < 1$$

where λ_1 and λ_2 are the roots of the polynomial $1 - aL - bL^2$ as obtained from the quadratic equations.

(9) $\lambda_1 = 1/2[-(-a) + \sqrt{(-a)^2 - 4(-b)}\ ]$

(10) $\lambda_2 = 1/2[-(a) - \sqrt{(-a)^2 - 4(-b)}\ ]$

These can be simplified to

(9a) $\lambda_1 = 1/2[a + \sqrt{a^2 + 4b}\ ]$

(10a) $\lambda_2 = 1/2[a - \sqrt{a^2 + 4b}\ ]$

It can then be shown that

(11) $\lambda_1\lambda_2 = [1/2(a + \sqrt{a^2 + 4b}\ )]\cdot[1/2(a - \sqrt{a^2 + 4b})]$

$= 1/4(a^2 - a^2 - 4b)$

$= -b$

$-\lambda_1\lambda_2 = b$

and

(12) $\lambda_1 + \lambda_2 = [1/2(a + \sqrt{a^2 + 4b})] + [1/2(a - \sqrt{a^2 + 4b})]$

$= a$

In order to meet the restrictions on λ_1 and λ_2 stated above, it follows that $0 < a < 2$ (since $a = \lambda_1 + \lambda_2$). If estimates of b are negative the condition $a^2 \geq 4b$ must be met in order to avoid negative or imaginary roots.

If the implied division

$$\frac{1}{1 - aL - bL^2}$$

is carried out the weights will be

$$w_0 = \frac{1 - a - b}{1}$$

$$w_1 = \frac{1 - a - b}{1} + bw_0$$

$$w_2 = bw_0 + aw_1$$

.

.

.

$$w_i = bw_{i-2} + aw_{i-1}$$

For all weights to be positive it follows that

$$\frac{1 - a - b}{1}$$

must be positive.

Griliches (11) points out that, even though the above conditions are sufficient to obtain a lag generating function, they will not consistently result in a function significantly different from a Koyck (geometrically declining) distribution. He further points out that, even when the distribution parameters estimated are acceptable, the confidence interval for the estimates is likely to include a substantial range of unacceptable values. Since this is the case other parameters of the lag distribution can be very useful in interpretation of the results. In particular, the mean lag and the variance of the distribution are useful. It is in estimating these that the fact that the lag generating function is a probability distribution proves to be invaluable.

It has been shown (11) that the mean or average lag of a probability distribution $W(L)$ is $W'(L)$ evaluated at $L = 1$. The variance of the distribution is $W''(L) + W'(L) - [W'(L)]^2$ evaluated at $L = 1$. More will be said about these computations in the next chapter, but it is clear that they can be obtained from the estimated lag distribution.

CHAPTER IV

THE MODEL, DATA AND ESTIMATING TECHNIQUE

The purpose of this study was to determine the lags associated with impact of public investment in water resources upon the economies of urban centered areas. Various measures of economic performance suggest themselves as reasonable variables upon which to focus attention in such a study. It could be argued that total income, per capita income, total employment and manufacturing employment could all be reasonable measures. None of these would be satisfactory in every respect and data availability is an ever present problem.

The advantage of income as the dependent variable upon which to focus attention in a model is that it is very aggregative. For this reason it should reflect various types of impacts of public investment in water resources. The measure of income used in this study is comparable to "disposable personal income" as reported by the United States Department of Commerce. The actual data used is "effective buying income" estimated by the publishers of Sales Management (23). It was necessary to use this source since data are not available for the entire study period from government sources.

It must be recognized that total income may not accurately reflect the nature of the impact of all types of public investment. Changes in population and employment could cause changes in per capita income or earnings per employed person that are not reflected in total income data. The possible consequences of population and employment shifts are reduced by using urban centered areas in the analysis. Much of the presumed benefit of public investment in water resources in the Tennessee valley is that it aids economic transformation and expansion of the industrial base. If this benefit is realized, population and employment shifts are frequently within the urban area as defined in the study.

In an effort to measure as close as possible to the point of impact, manufacturing employment might be used as the variable reflecting the impact of investment projects. Krutilla (16) even suggests looking at specific categories of manufacturing activity (e.g. the chemical industry in Paducah). Actually this "leading sector" approach has much to commend it. Unfortunately data are not available on an annual basis in a form that permits estimation of the level of employment (or production) in specific industries in the study areas.

The Income Model

The income model used in the study was

$$I_t = P(L)B_1T_t + B_2A_t + B_3N_t + B_4D_t + B_0$$

where

I = total disposable income in the area measured in millions of dollars per year.

P(L) = rational distributed lag function as described in Table III-1. Four forms of a rational lag function were estimated.

T = annual public investment in water resources measured in millions of dollars per year.

A = private agricultural employment measured in thousands of workers.

N = private non-agricultural employment measured in thousands of workers.

D = employment rate in the national economy during the year. This is a proxy variable for export demand in the area.

B_0 = a constant term

t = time period

The purpose of this model is to examine the popular notion that the flow of public investment in water resources into an area will cause income to increase as a direct result of the publicly stimulated economic activity. A complete description of the data, along with sources, is included in the Appendix. Most of the actual data are not included since they are generally available in widely distributed secondary sources.

The employment variables are included in the equation to reflect internal conditions in the region. The agriculture in these areas is predominantly subsistence and is likely to be a lagging sector of the area economies. In situations like this it is reasonable to expect declines in agricultural employment to be associated with economic growth of the area. It has been pointed out that availability of non-farm employment opportunities normally leads to more rapid migration out of agriculture (4). Private non-agricultural employment should reflect activity in the leading sectors of the economy. The most important component of these data are manufacturing and wholesale and retail trade employment. Both these elements would be expected to be favorably influenced by impact of public investment upon the economy of an area.

The national employment rate is included in the model to reflect the important ties between the area economy and the national economy. If aggregate demand is low in the national economy this should be reflected in the national employment rate. The income generated by the economy of the urban centered area will depend to a considerable extent on primary income which will be influenced by demand in the national economy through export demand for production in the area economy. Also, the growth in the area economies will, to a considerable extent, depend upon outside capital. This capital will not generally be forthcoming during periods of high national unemployment. Throughout much of the study this variable is referred to merely as demand.

Two different measures of public investment were used in the model. These were public investment in dams and steam plants within the respective study areas (following the precedent set by Wiebe and Chappell) and public investment in dams in the entire Tennessee River Valley Region (plus Barkley Dam, a Corps of Engineers project on the Cumberland River near Paducah).

The Manufacturing Employment Model

The manufacturing employment model used in the study was

$$M_t = P(L)\ B_1 G_t + B_2 D_t + B_0$$

where

M_t = manufacturing employment in the area measured in workers.

G_t = total value of all past public investment in water resources in the area measured in millions of dollars.

B's, D, and P(L) are the same as defined in the previous section.

The logic behind the model is that public investment in water resources would influence internal factors in the area which would influence manufacturing employment. The national aggregate demand was a necessary component of the model since most of the products manufactured in the areas studied would be sold outside the area.

This model was developed from the results of the income model. The purpose of the model is to examine the influence of increased social overhead capital and consequent flow of services. Conceptual as well as statistical difficulties arise in choosing an appropriate variable to use as a measure of the stock of social overhead capital for the model. An appropriate variable must reflect the multipurpose nature of public activity in water resources in the study areas. This eliminates such things as miles of navigable water channel or peak load electric generating capacity. Since time lags are a prime consideration in the study, the measure of the stock of social overhead capital should also reflect a potential rather than an actual flow of services. This eliminates actual use data such as river traffic or electric power sales. A third consideration is that the variable should reflect the services available in the respective urban centered study areas.

The variable decided upon was the sum of all the annual investments made by TVA in the respective areas (plus Corps of Engineers Barkley Dam in the Paducah area). This is merely the cumulative amounts of the annual investments included in the income model. This variable can be conceptualized either as a stock of social overhead capital or as proxy for the potential flow of services from past investments.

The chief difficulty with the variable is its property that G_t is not independent of G_{t-1}. That is, G_{t-1}

comprises a major part of the observed value G_t. Such autocorrelated data are frequently a problem in statistical analysis but the distributed lag model may compound the difficulty since G_t, G_{t-1} and G_{t-2} may appear in the same estimating equation. In order to minimize this difficulty, only the lag configuration with a zero order polynomial in the numerator was estimated for this model. This configuration did not include lagged values of G in the estimating equation.

The Estimating Technique

Using the income model and the most general form of the lag function for illustration, the equation to be estimated is

$$(14) \quad I_t = \frac{1 + aL + bL^2}{1 - cL - dL^2} B_1T_t + B_2A_t + B_3N_t + B_4D_t + B_0$$

Multiplying by the denominator to remove the fraction this becomes

$$(15) \quad I_t - cI_{t-1} - dI_{t-2} = B_1T_t + aB_1T_{t-1} + bB_1T_{t-2} + B_2A_t - cB_2A_{t-1} - dB_2A_{t-2} + B_3N_t - cB_3N_{t-1} - dB_3N_{t-2} + B_4D_t - cB_4D_{t-1} - dB_4D_{t-2} + B_0(1 - cL - dL^2)$$

For estimating purposes equation (15) can be written

$$(16) \quad I_t = cI_{t-1} + dI_{t-2} + B_1T_t + aB_1T_{t-1} + bB_1T_{t-2}$$

$$+ B_2A_t - cB_2A_{t-1} - dB_2A_{t-2} + B_3N_t - cB_3N_{t-1}$$

$$- dB_3N_{t-2} + B_4D_t - cB_4D_{t-1} - dB_4D_{t-2}$$

$$+ B_0(1 - c - d)$$

where a, b, c, d and the B_i's are parameters to be estimated and the other variables are as defined above.[1]

It will be noted that the equation is non-linear in the parameters; i.e. the actual equation coefficients are combinations of more than one parameter. This condition necessitates use of special estimating techniques.

Edwards (7) has suggested a method of estimating non-linear equations which Bauer (2) used in estimating lag distributions. The technique, which Edwards refers to as generalized regression, is an iterative process which includes the following steps:

1. Choose an initial set of values for the parameters to be estimated.

2. Take the first partial derivative of the equation

---

[1]The form in which to enter the constant term raises some questions in this equation. An alternative would be to assume that $B_0cL = 0$ and $B_0dL^2 = 0$. There seems to be little reason to favor this over assuming that $B_0L = B_0$ and $B_0L^2 = B_0$ as was done in equations (15) and (16). The advantage of the latter assumption in the present case is that estimates of B_0 cannot be obtained by the method used in the study if the former assumption is made. A second alternative is to assume no constant term. Estimates using this approach were made but yielded lower R^2 coefficients; also, parameter estimates were not as reasonable as those obtained using the approach of equation (16).

with respect to each of the parameters to be estimated.

3. Compute the residual $I - \hat{I}$ for each observation of I and the independent variables. $\hat{I}$ is calculated using the parameters specified in step 1.

4. Using ordinary least squares regression, calculate adjustment factors Δi for each of the parameter estimates using the equation

$$I - \hat{I} = \Sigma \Delta_i \frac{\partial I}{\partial B_i}$$

(i = the number of parameters to be estimated)

where $\frac{\partial I}{\partial B_i}$'s are treated as dependent variables.

5. Calculate the new parameter estimates by adding the Δ_i's to the parameter estimates made in step 1.

6. Using the new parameter estimates repeat steps 3, 4, and 5, adjusting the parameter estimates calculated in step 5 of the previous iteration with the calculated Δ_i's. Repeat the procedure until all Δ_i's converge to zero.[2]

The partial derivatives of equation (16) are

$$\partial I/\partial a = B_1 T_{t-1}$$

---

[2]Fuller and Martin (10) show that the standard error of the parameters can be estimated in the usual manner as the square root of the product of the relevant elements of the c_{ij} matrix obtained in the final iteration of the procedure and the mean square error of the estimated equation.

$\partial I/\partial b = B_1 T_{t-2}$

$\partial I/\partial c = I_{t-1} - B_2 A_{t-1} - B_3 N_{t-1} - B_4 D_{t-1} - B_0$

$\partial I/\partial d = I_{t-2} - B_2 A_{t-2} - B_3 N_{t-2} - B_4 D_{t-2} - B_0$

$\partial I/\partial B_1 = T_t + aT_{t-1} + bT_{t-2}$

$\partial I/\partial B_2 = A_t - cA_{t-1} - dA_{t-2}$

$\partial I/\partial B_3 = N_t - cN_{t-1} - dN_{t-2}$

$\partial I/\partial B_4 = D_t - cD_{t-1} - dD_{t-2}$

$\partial I/\partial B_0 = 1 - c - d$

Naturally, these would be adjusted for different lag functions.

Edwards demonstrates that this procedure yields least square estimates but the distributed lag model introduces violations of the least square assumptions which can cause the estimates to be biased and statistical tests to be of limited value. It has been shown that the lag estimating technique used may cause autocorrelation of the error terms. Also, multicolinearity is a rather serious problem since variables in year t enter the equation along with the same variable in the year t-1 and t-2. An additional source of autocorrelation exists in models where cumulative investment is used as the measure of public investment in water resources since one component of any observation is the quantity existing in the previous time period.

It should be pointed out that the 1 in the numerator of the lag functions is an arbitrary constant which ideally should be an estimated parameter. The reason for its use can be seen if a parameter g is substituted in this position in the equation. If this was done the vector $\partial I/\partial B_1$ would be the linear combination $\partial i/\partial a + \partial I/\partial b + \partial I/\partial g$, giving a singular matrix and making the estimation impossible. Bauer assigned a value of one arguing that the value made little difference to the final outcome since it only had direct influence on the first two weights and was rapidly diminished in impact on future weights. Following Bauer's precedent the same procedure was used in this study. Testing other values of g demonstrated that they did indeed have practically no effect upon the weights of the lag distributions.

The Lag Weights

If the division of N(L)/D(L) is carried out for the lag configuration

$$\frac{1 + aL + bL^2}{1 - cL - dL^2}$$

the following weights are computed:

$$p_0 = 1$$

$$p_1 = a + cp_0$$

$$p_2 = b + dp_0 + cp_1$$

.

.

$$p_i = dp_{i-2} + cp_{i-1}$$

It was pointed out above that P(L) is an intermediate step in obtaining the desired rational lag function W(L) since there is no reason to expect

$$\sum_{i=0}^{\infty} p_i = 1.$$

In order to assure that the weights sum to one a new function is formed.

$$P(L) / \sum_{i=0}^{\infty} p_i = W(L)$$

It has been shown (2) that

$$\sum_{i=0}^{\infty} p_i = \frac{1 + a + b}{1 - c - d}.$$

From this it follows that:

$$w_i = p_i / \frac{1 + a + b}{1 - c - d} = p_i \left(\frac{1 - c - d}{1 + a + b}\right)$$

or

$$w_0 = \left(\frac{1 - c - d}{1 + a + b}\right)$$

$$w_1 = a\left(\frac{1 - c - d}{1 + a + b}\right) + cw_0$$

$$w_2 = b\left(\frac{1 - c - d}{1 + a + b}\right) + dw_0 + cw_1$$

$$w_3 = dw_1 + cw_2$$

.

.

$$w_i = dw_{i-2} + cw_{i-1}$$

Mean and Variance of the Lag Distribution[3]

It was pointed out earlier that the mean of the lag distribution is the first derivative of the distribution evaluated at L = 1. In the case of the distribution

$$W(L) = P(L) / \sum_{i=1}^{\infty} p_i$$

$$w_i = \frac{1 + aL + bL^2}{1 - cL - dL^2} \Big/ \frac{1 + a + b}{1 - c - d}$$

$$\left.\frac{\partial W(L)}{\partial L}\right|_{L=1} = \left(\frac{(1 - c - d)(a+2b) - (1 + a + b)(-c-2d)}{(1 - c - d)(1 - c - d)}\right) \cdot$$

$$\left(\frac{1 - c - d}{1 + a + b}\right)$$

$$= \frac{a + 2b - bc + c + 2d + ad}{(1 - c - d)\ (1 + a + b)}$$

The average lag indicates the center of gravity of the estimated distribution.

The variance of the lag distribution can be shown to be:

---

[3] Proofs for the assertions in this session can be found in Griliches (11).

$$\frac{\partial W(L)}{\partial (L)^2} - \frac{\partial W(L)}{\partial (L)} - \left(\frac{\partial W(L)}{\partial (L)}\right)^2$$

evaluated at L = 1. Due to its computational similarity to the average lag the derivation will not be carried out. The variance of the distribution gives a measure of dispersion and is useful in comparing one distribution with others in determining the pattern of the impact.

CHAPTER V

RESULTS OF ESTIMATIONS

The task of this chapter is to examine the results of the regression analysis for the income and manufacturing employment models. Hypotheses are made as to the nature of the influence of public investment in water resources upon income and employment in urban centered areas. Information regarding the distributed lag functions is examined to determine the effectiveness of the tool in isolating lags in the impact of public investment in water resources.

The Income Model

A total of 24 equations were estimated for the income model. In order to facilitate the discussion of these models a three-digit code has been assigned to each equation. The designations for the code are given in Table V-1. The first digit designates the area; the second, the measures of public investment; and the third, the lag configuration used in the equation. For example, equation 4-1-2 is for the Huntsville area, uses annual public investment in dams and steam plants in the area as the measure of public investment and uses the lag configuration $1 + aL / 1 - bL - cL^2$.

Table V-2 summarizes the structural parameters for the 24 equations estimated for the income model. The numbers in the column and row headings refer to the code numbers

Table V-1. Code designations for the equations reported in the study.

Code	Designation
First digit	<u>Study region</u>
4	Huntsville, Alabama
5	Paducah, Kentucky
6	Asheville, North Carolina
Second digit	<u>Measure of public investment</u>
1	Dams and steam plants in the area
2	Dams in the entire valley region
Third digit	<u>Lag function</u>
2	$1 + aL + bL^2 / 1 - cL$
3	$1 / 1 - aL - bL^2$
4	$1 + aL / 1 - bL - cL^2$
5	$1 + aL + bL^2 / 1 - cL - dL^2$

Table V-2. Structural parameters for the regression equations estimated using the income model.[a]

Lag Function	Huntsville (4) Area investment (1)	Huntsville (4) Regional investment (2)	Paducah (5) Area investment (1)	Paducah (5) Regional investment (2)	Asheville (6) Area investment (1)	Asheville (6) Regional investment (2)
			Demand (national employment rate)			
2	5.891	4.867	.583	1.125	2.304	2.203
3	5.953	4.395	.528[b]	.225	1.888	1.278
4	2.164	4.329	.554	.070	.630	1.358
5	6.087	4.858	.903	.623	1.879	1.433
			Agricultural Employment			
2	.692	-.117	-5.478	-4.232	-.083	-.091
3	.746	-.444	-10.252[bc]	-10.328	.897	1.318
4	.821	-.458	-10.419[c]	-10.229	-5.740	1.037
5	.612	-.246	-10.231[c]	-9.991	.841	1.012

Table V-2 (cont'd)

Lag Function	Huntsville (4) Area investment (1)	Huntsville (4) Regional investment (2)	Paducah (5) Area investment (1)	Paducah (5) Regional investment (2)	Asheville (6) Area investment (1)	Asheville (6) Regional investment (2)
			Non-agricultural Employment			
2	1.093	2.048	1.663	1.391	2.658	2.211
3	1.199	2.008	1.667[b]	1.725	2.350	2.154
4	2.429	2.020	1.719	1.841	6.741[c]	2.392
5	1.148	2.089	2.012	1.634	2.672	2.343
			Public Investment			
2	-.460	.279	-.042	-.128	-.039	.029
3	-.533	.105	.015[b]	.017	.407	.153
4	.271	.116	.202	-.019	.015	.103
5	-.610	.290	-.003	-.121	.152	.116

[a]Income is measured in millions of dollars per year, national employment rate in percent, employment in thousands of workers and public investment in millions of dollars per year.

[b]Reasonable lag distributions were obtained in this equation.

[c]These parameters were significant at the .05 level.

described in Table V-1. The data are organized so that all estimates of the parameters are grouped together for easy comparison. Ideally the eight parameters estimated for export demand (i.e., the national employment rate), private agricultural employment and private non-agricultural employment should be the same within each area. Likewise the four parameters for each measure of public investment should be equal for each area. This follows directly from the assumption that independent variables in a regression model have no covariance. Although the estimates varied rather widely in some cases, they are generally close enough that the approximate magnitude of the parameters can be determined.

The table also indicates those parameters which were significantly different from zero at the .05 level. It was pointed out above that caution must be used in interpreting the significance tests in the model since the "least square assumptions" were not fulfilled by the model.

Demand

The parameters associated with external demand for the areas' products, i.e., the national employment rate, were positive in all cases. Substantial differences in the influence of this variable upon area income are suggested. The Huntsville area was most influenced by national demand. In that area a one percent increase in the national employment rate was associated with an increase of between five and six million dollars in annual income in the area.

The estimates for the other two areas were somewhat more mixed but most of the estimates for Asheville were between 1.5 and 2.5.[1] In the Paducah area the estimated parameters for the different equations range from an increase of 70 thousand dollars to 1.125 million dollars in area income for a one percent increase in the national employment rate. None of the equations yielded parameters which were significant at the .05 level.

Since Paducah is somewhat smaller than the other two areas, the absolute influence of a change in the national aggregate demand would be expected to be somewhat lower there than in the Asheville and Huntsville areas. However, size alone cannot explain the difference in the estimates. Two additional factors come to mind as possible causes for area differences in the demand parameters. The first is the nature of the key industry or group of industries which are the basis of the area's growth. The second is the importance of the relationship between the area's economy and the national economy. A plausible explanation for the low

---

[1]It is noteworthy that the parameters for equation 6-1-4 are substantially different than estimates for other equations estimated for the Asheville area. This raises some doubt as to the accuracy of the estimate. Fuller and Martin (10) suggest that the estimating technique may not always give one unique solution. They further suggest that when this case is suspected, an alternative set of original estimates of the parameters should enable a researcher to find the true least square solution. Although several attempts were made to find estimates in this equation which were more consistent with the other results, none were successful.

parameters for the Paducah area is the first; namely, the nature of the industrial base in the area. It is possible that the rapid growth in the chemical industry swamps the cyclical influences in the national economy upon the economy of the Paducah area. Asheville, with a more diversified and somewhat older industrial base, is influenced significantly by the cyclical forces in the national economy. Federal munitions and space projects play an important role in the Huntsville area's economy. Since such expenditure is probably not greatly influenced by national aggregate demand (it could even be countercyclical), it is surprising that the Huntsville area had such high parameters associated with the national aggregate demand variable.

<u>Private Agricultural Employment</u>

In the Paducah area a decrease of one thousand workers in private agricultural employment is associated with an increase in disposable income for the area of from 5 to 10 million dollars (depending upon the equation used). The estimates for both the Asheville and Huntsville areas are very mixed both in sign and magnitude suggesting a weak relationship between agricultural employment and income in these areas. Three of the equations for Paducah yielded estimates which were significant but none were significant for Asheville or Huntsville.

The parameters estimated for the three areas suggest a different type of economic transformation occurred in the three areas during the period studies. Apparently the income growth in the Paducah area was associated with substantial reductions in agricultural workers. They may have been drawn into non-agricultural employment within the area or they may have migrated outside the area in search for a better livelihood. The strong inverse relationship between agricultural employment and total income apparently did not occur in the Asheville and Huntsville areas. This may be the result of newly created employment being filled by workers from outside the area (e.g., physics Ph.D.'s in Huntsville). Or, it could result from continued large numbers of small subsistence farmers which persist in these areas.

Private Non-Agricultural Employment

The estimated coefficients for private non-agricultural employment are practically all positive and, compared to the other parameters, the different areas have relatively similar magnitudes. The estimates for Paducah indicate that one thousand additional workers in private non-agricultural occupations are associated with between 1.5 and 2 million dollars of additional income in the area. The estimates for Asheville show between two and three million dollars of additional income per thousand additional private non-agricultural workers (except in equation 6-1-4). The estimates for Huntsville were, for the most part, very close to two million dollars per additional thousand workers in non-agricultural employment.

The most plausible reason for the higher estimates for the Asheville area is that, for the study period as a whole, both the level of education of the population and the capital intensity of the industrial mix were somewhat higher in the Asheville area. It is important to note, however, that the gap in education has practically disappeared and the capital intensity of the industrial mix had reversed in the latter years of the period.

Public Investment in Water Resources

The coefficients listed for the two measures of public investment indicate the additional dollars of income annually in the respective areas associated with each additional dollar of public investment in water resources (both variables were actually measured in millions of dollars). Very few of the estimates were significant and substantial variation in both the sign and magnitude of the parameters exists. Unlike the other parameters in Table V-2, these estimates would be expected to differ for the different measures of the variable. For the reasons pointed out above, they should be the same for the different lag models estimated for each equation.

The coefficients obtained when annual investments on dams and steam plants within the area were used as the measure of public investment were quite different both among the areas and for the different lag configurations for the respective areas. None of the estimated parameters were significant at the .05 level. For Huntsville all were negative but one. In Paducah two were positive and two were negative. In Asheville three were positive and one was negative. In the equations where investment in dams in the entire TVA region was used as the measure of public investment both Huntsville and Asheville had consistently positive signs on all four coefficients estimated. Huntsville had estimates ranging from nearly .3 to just over .1; Asheville had a range from .03 to .15. Paducah had one positive and three negative coefficients. It is clear that the direct measurable impact of the TVA investment, whether in the immediate area or not, was quite weak.

In trying to interpret these results one must ask what type of impact would be expected from each of the measures of public investment in water resources. If income is increased as a result of the usual Keynesian income multiplier, an increase in public expenditure within the area would result in increased income. Completion of the project and the consequent reduction of public expenditure should result, other things equal, in a return of income to the previous level. If this pattern was shown in the data, the coefficient associated with annual public investment in dams and steam plants would be positive. If, on the other hand, income increased as the result of increased social overhead capital which resulted in a continued growth in income after the project was completed, the sign on the coefficient for annual public investment in dams and steam plants would probably be negative. Annual investment data would not be suitable for investigating the influence of increased social overhead capital or increased availability of services resulting from public investment in water resources. From the point of view of economic policy, the results of the study show that the short-run impact of public expenditure in water resources upon the income and employment or urban centered areas is quite small--indeed not measured by the models estimated. It can be hypothesized that any impact which these investments have would be in the longer run and are the result of increased flow of services.

Distributed Lag Functions

A distributed lag function for the income model shows how the influence of a change in the annual rate of public investment in water resources upon the income of the area is distributed over time. In view of the fact that the analysis failed to reveal any consistent relationship between annual levels of public investment in water resources and income, it is not surprising that only one reasonable lag distribution was estimated by the model. It occurred in the equation for Paducah using investment within the area and lag configuration $1/1\text{-}aL\text{-}bL^2$. The weights of the distribution are shown in Figure V-1 and

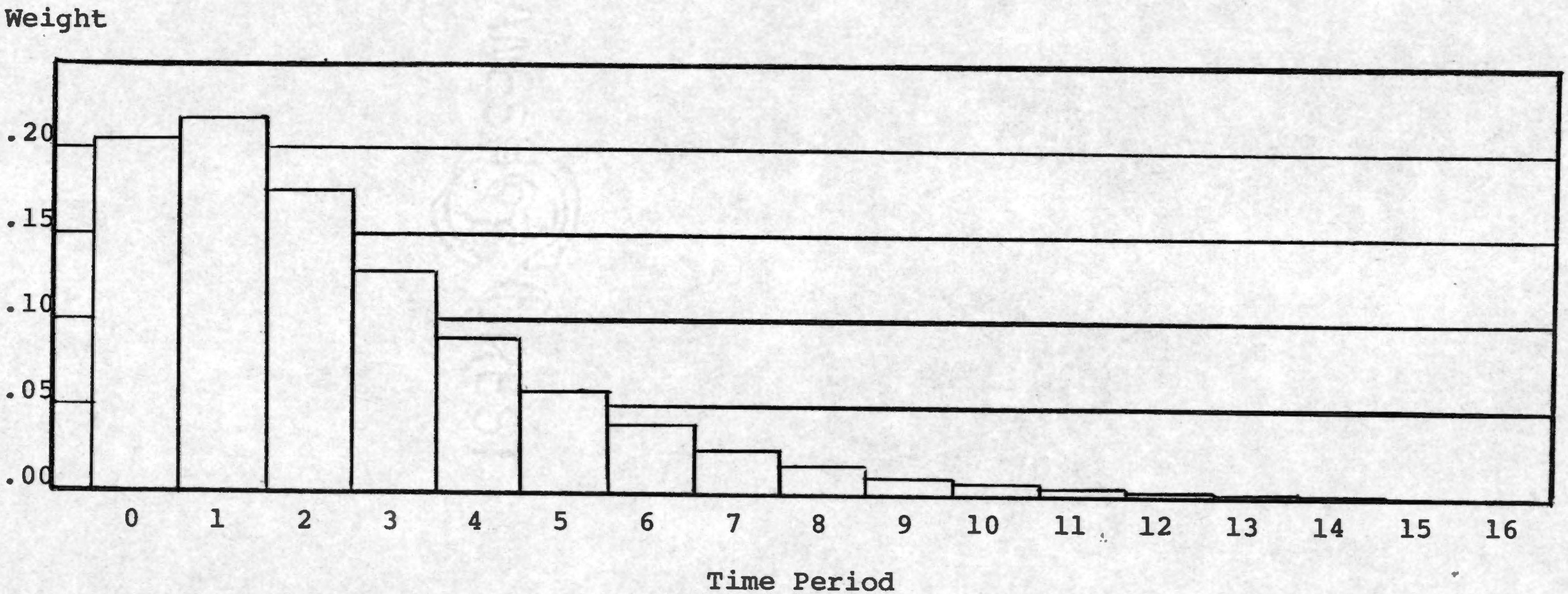

Figure V-1. Histogram of the reasonable lag distribution obtained in the income model for the Paducah area.

in Table V-3. Twenty percent of the impact was felt during the year of the investment and nearly 22 percent the following year. The mean of the distribution is 5.3 years and the variance is 31.4 years. Actually this would probably be a rather long lag if it were the result of lags in the usual multiplier effect of public expenditure. In view of the inconsistent estimates for the coefficients on the public investment parameter the lag is very difficult to interpret. Rather, it would seem to be better to turn to the manufacturing employment model for evidence of the lagged impact of public investment in water resources upon income and employment in the urban centered areas.

The Manufacturing Employment Model

The manufacturing employment model was used to test the hypothesis that public investment in water resources influences income and employment of an area because of permanent increases in social overhead facilities. The flow of services from these facilities results in a change in the competitive position of the area relative to other areas. This, in turn, causes private investment to flow into the area and income and employment to increase. The choice of manufacturing employment as the dependent variable was based on the assumption that economic expansion will generally result from increased manufacturing activity. It must be recognized that other types of industries may

Table V-3. Weights, mean lag and variance of the lag distribution obtained for the Paducah area in the income model.

Time period	Weights
0	.206
1	.218
2	.176
3	.129
4	.090
5	.061
6	.041
7	.027
8	.018
9	.012
10	.008
11	.005
12	.003
13	.002
14	.001
15	.001
16	.001
Mean lag	5.3
Lag variance	31.4

benefit from the improved competitive position of the area, however.

The Structural Parameters

It was pointed out in Chapters III and IV that statistical problems in the distributed lag model made use of all but one of the family of lag functions unwise in this model. The parameters estimated in the equations are listed in Table V-4.

The parameters estimated for the national employment rate have the expected positive sign and reflect the same relative importance of this variable in the three areas as did the income model. Huntsville had the strongest association between national demand and manufacturing employment. A one percent increase in the national employment rate was associated with 586 additional workers in manufacturing in the Huntsville area. In Asheville it was associated with 304 additional manufacturing workers, and in Paducah a one percent increase in national employment was associated with only 141 additional workers in manufacturing. All the parameters for this variable were statistically significant.

The measure of public investment in water resources was the cumulative amount of all past public expenditures on dams and steam plants in the respective areas made by TVA (plus U. S. Corps of Engineers expenditures in Barkley Dam for the Paducah area). Huntsville and Paducah had

Table V-4. Structural parameters estimated for the regression equations estimated using the manufacturing employment model.[a]

	Demand (national employment rate)	Cumulative value of public investments
Huntsville[b]	586[c]	8[c]
Paducah[b]	141[c]	2[c]
Asheville	304[c]	-1

[a]Manufacturing employment is measured in workers, national employment rate in percent and cumulative value of public investment in millions of dollars.

[b]Reasonable lag distributions were obtained for these areas.

[c]These parameters were significant at the .05 level.

estimates which were positive and statistically significant at the .05 level. The estimate for Asheville was not statistically significant at the .05 level and was negative. In the Huntsville area the addition of one million dollars to water resource facilities was associated with eight additional jobs in manufacturing. In Paducah the same investment was associated with two additional workers in manufacturing. The estimate for Asheville showed that an additional one million dollars in cumulative investments in water resources was associated with a reduction of one worker in manufacturing.

One could hypothesize that at least some of the difference in these parameters was due to location. It is logical, for example, to argue that the reason for the higher coefficients in Huntsville area is its location in the center of the Tennessee Valley. Because of its location, it is in a position to have gained from the increased barge transportation while Asheville cannot be reached by river traffic. Paducah would have gained little by improvement of river transport since it is located on the Ohio River at the mouth of the Tennessee. For the Asheville area, an even more compelling explanation for the weak relationship is that the public investment in water resources is small in that area and the sites of such investment are at rather remote locations in the area.

The Distributed Lag Functions

The distributed lag function for the manufacturing employment model shows how the influence of an increase in the cumulative value of public investments in water resources upon manufacturing employment is spread over time. The weights of the distributed lag functions estimated for the Huntsville and Paducah areas are shown in Figure V-2. Table V-5 contains the weights, the mean lag and the variance of the lag distribution.

The discussion of the structural parameters showed that the total impact of public investment upon manufacturing employment was greater for Huntsville than for Paducah. The distributed lag function on Figure V-2 shows that the impact was felt more quickly in the Paducah area. Sixteen percent of the impact was felt in the year that the investment was made in the Paducah area and only eight percent of the impact was registered in that year for the Huntsville area. The mean lag for Huntsville was 8.6 years and for Paducah was 5.2 years. The variance of the distribution was 78 years for the Huntsville area and only 32.9 years for the Paducah area. The mean and variance figures indicate that not only was the impact slower for the Huntsville area but it was also more widely dispersed over time.

The results of this part of the study add to the level of knowledge of the impact of public investment in

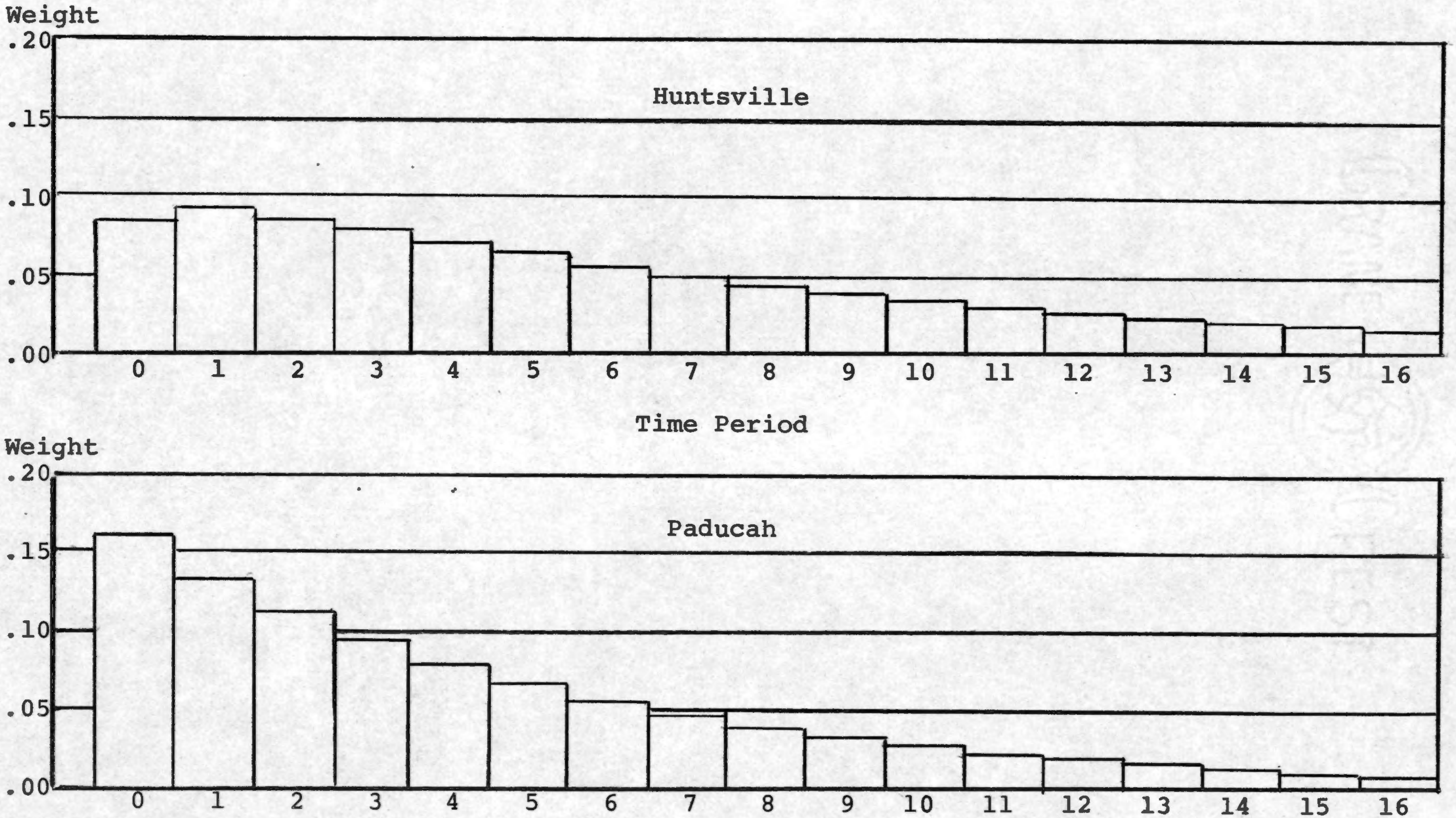

Figure V-2. Histogram of reasonable lag distributions obtained in the manufacturing employment model for the Huntsville and Paducah areas.

Table V-5. Weights, mean lag and variance of the lag distributions obtained for the Huntsville and Paducah areas using the manufacturing employment model.

Time period	Huntsville	Paducah
0	.0842	.1616
1	.0932	.1340
2	.0871	.1126
3	.0786	.0946
4	.0704	.0795
5	.0629	.0668
6	.0561	.0561
7	.0501	.0471
8	.0447	.0396
9	.0399	.0333
10	.0357	.0280
11	.0318	.0235
12	.0284	.0197
13	.0254	.0166
14	.0227	.0139
15	.0202	.0117
16	.0181	.0098
Mean lag	8.6 years	5.2 years
Lag variance	78.0 years	32.9 years

water resources upon urban centered areas. The reasonable lag distributions obtained for the Paducah and Huntsville areas should serve as a basis for further research and would support the use of distributed lag analysis in similar studies.

CHAPTER VI

CONCLUSIONS, EVALUATION, AND SUGGESTIONS FOR FURTHER RESEARCH

The purpose of this study was to investigate the time dimension of the impact of public investment in natural resources upon the economy of urban centered areas. The specific objectives were: (1) to evaluate the effectiveness of estimating distributed lag functions as a technique for studying the time lag associated with the economic impact of public investment in water resources; and (2) to estimate the distribution over time of the impact of investment in water resources upon the economy of urban centered areas by estimating a rational distributed lag function. It is obvious that the success in estimating rational distributed lag functions to achieve the second objective is a most important consideration in arriving at conclusions regarding the first objective.

The procedure followed was to estimate parameters for a model relating income in urban centered areas to agricultural employment, non-agricultural employment, the national employment rate and annual public investment in water resources. Separate estimations were made using two different measures of public investment in water resources. Based upon the results of estimates using this income model a second model was developed. This model related manufacturing employment in the urban centered areas to the national employment rate and the cumulative value of all past public investment in water resources in the respective areas. The economic units of observation where the urban centered areas around Huntsville, Alabama; Paducah, Kentucky; and Asheville, North Carolina.

Results of Estimations

The Income Model

The first structural model examined was referred to as the income model since it used total disposable income in the area as the dependent variable. The least squares estimation procedure yielded estimates of the structural parameters and, in one case, a reasonable distributed lag function relating annual public investment in water resources to income in the area.

The impact of national aggregate demand as measured by the national employment rate was estimated to be somewhat stronger in the Huntsville area than the Asheville area and both were influenced more by this variable than the Paducah area. A one percent increase in the national employment rate was associated with between four and six million dollars increase in income in the Huntsville area, about one to two million dollars in the Asheville area and less than one million dollars in the Paducah area.

Reductions in private agricultural employment were associated with increases in income in the Paducah area but no clear relationship between these two variables was shown for the Asheville or Huntsville areas.

Private non-agricultural employment was directly related to income in all three areas. For Asheville an increase of one thousand people in private non-agricultural employment was associated with an increase of about 2.5 million dollars of annual income in the area. For Huntsville and Paducah the same increase in non-agricultural employment was associated with from one to two million dollars increase in total income per year.

Examination of parameters for annual public investment in dams and steam plants indicates that the multiplier effect of public investment upon the economies of the urban centered study area is very small.

Separate estimates were made of the parameters relating public investment in water resources to income for two measures of public investment. The parameters for annual investment in dams across the entire TVA region had mixed signs for the different equations and none was statistically significant. It is not surprising that only one reasonable distributed lag function was estimated for this model.

The findings of the model lead to the rejection of the popular notion that public investment in water resources leads to significant direct increases in area income due to the inflow of public funds. An alternative hypothesis was developed that the impact of public investment in water resources resulted in increased social overhead capital and consequent increased flow of services available in the area.

The Manufacturing Employment Model

To examine the hypothesis that increased social overhead capital was the primary source of increased employment and income from public investment in water resources, a second model was formulated. In this model manufacturing employment was the dependent variable and the national employment rate and the cumulative value of past investment in dams and steam plants in the respective areas were the independent variables. Because of statistical considerations only one lag configuration was estimated for this model.

The impact of the national employment rate upon manufacturing employment was stronger for the Huntsville area than the Asheville area and both were influenced more by this variable than the Paducah area. A one percent increase in the national employment rate was associated with 586 additional workers in manufacturing for the Huntsville area, 304 additional workers in manufacturing for the Asheville area and 141 additional workers in manufacturing for the Paducah area. All of these parameters were statistically significant at the .05 level.

Parameters relating manufacturing employment to the cumulative value of public investment in dams and steam plants were quite small. For Huntsville an addition of one million dollars to the cumulative value of public investment in water resources was associated with eight additional jobs in manufacturing; for Paducah, two additional jobs in manufacturing; and for Asheville, one less job in manufacturing. The estimated parameters were statistically significant for Huntsville and Paducah but not for Asheville.

Reasonable distributed lag functions were obtained for the Huntsville and Paducah areas. The lagged impact for Huntsville was slower and more dispersed than for Paducah. The mean lag for the Huntsville area was 8.6 years and the variance was 78 years. For the Paducah area the mean lag was only 5.2 years and the lag variance was 32.9 years.

Statistical Considerations

Distributed lag analysis gives rise to special statistical problems. These have been mentioned earlier at several points and must be considered here before a final evaluation of the distributed lag technique can be made.

The most widely discussed problem with the type of distributed lag model used in this study is serial correlation. This may result from transforming the model for estimating purposes or from specification error. The result is that coefficients are likely to be biased. Griliches (11) suggests two possible approaches to the problem. If the form of the serial correlation is known the estimating equation could be adjusted to correct bias. His second suggestion is to use instrumental variables or a two stage procedure. Neither of these was practical in this study, however. The problem with the first technique is that the form of the dependence is not known and operational restrictions on the forms which are computationally practical leave the procedure open to question. Furthermore, in this study the degrees of freedom are somewhat limited giving rise to an additional limitation since estimation of additional parameters is required. The use of instrumental variables or two stage procedures is extremely complex especially in the non-linear equation estimated in distributed lag analysis. Furthermore, results of studies using this technique have not been encouraging (11).

Multicolinearity is a second problem encountered in distributed lag analysis. This is the inevitable result of introducing a variable (X_t) and the same variable lagged one or two years (X_{t-1} or X_{t-2}) into the same equation. The correlation matrix of the variables included in the model confirms the _a priori_ notion that this problem exists. Eliminating this problem is virtually impossible since the lagged variables are an integral part of the estimation procedure.

The result of the statistical shortcomings of the model are that the parameter estimates may be biased and statistical testing may not be reliable. The seriousness of these problems can be quite great and conclusions based upon the model must be treated cautiously.

Evaluation of the Rational Distributed Lag Technique

The results of the estimates for the Huntsville and Paducah areas using the manufacturing employment model give support to the use of the rational distributed lag technique in future research of this nature. The estimated lag function should give useful information to economic planners, project analysts and policy makers.

The technique does have shortcomings which must be recognized. First of all, it complicates statistical estimating procedure and gives rise to problems discussed in the previous section. A second shortcoming of the technique is its inconclusiveness. The failure to obtain reasonable lag distributions is inconclusive evidence that no systematic time relationship exists between total income or manufacturing employment and public investment in water resources.

A different type of difficulty arises from the restrictions placed upon the form that the lagged impact may take. An advantage of the Jorgenson method used in this study is that it is less restrictive than alternative methods, e.g. Koyck. Even the Jorgenson method requires that all weights be positive. This eliminates a lagged impact that, for example, was negative in time periods immediately following investment and then became positive in subsequent time periods. Such might well be the case if a dam removed substantial amounts of farm land from production and, after some lapse of time, caused the lost income to be offset by increased industrial production.

On the basis of the results of this study it must be concluded that the rational distributed lag technique has value in aggregative models for evaluating the time lags associated with the impact of public investment in water resources. Limitations of the technique must be recognized, however.

Conclusions

The analysis provides the basis for several tentative conclusions. These conclusions contribute to the very limited knowledge of the impact of public investment in water resources upon the economy. They are especially helpful in pointing directions in which further research might be most fruitful.

These conclusions are:

1. The impact of public investment in water resources is apparently the result of increased social overhead capital and its consequent increased flow of services rather than regional income multiplier effects of the public expenditure.

2. The complex matrix of factors causing economic advance of urban centered areas is so different for different areas that public investment in water resources has different degrees of impact upon the different areas. Also,

the time dimensions of the impact for different areas vary rather widely.

Suggestions for Further Research

Additional research is needed to explain why the lags reported for the Paducah and Huntsville areas had the dimensions which they did. This is especially true in view of the fact that the lagged impact was apparently different for the different areas. Two directions can be suggested for such research. First, research should be carried out to discover the factors complementary to the public investment which leads to economic expansion. Also, along these lines, the existence of threshold levels for public investment in water resources, as well as for complementary factors, should be explored and identified.

The second suggestion for further research is to focus upon the flow of services resulting from the investment rather than upon the investment itself. If, as was concluded from this study, the important impact of public investment in water resources is from social overhead capital rather than income multipliers, the dollar amounts invested is a poor index of the potential benefit. The importance of the various types of service flows should be explored using price and quantity data on the service itself rather than such aggregative measures as dollars invested in physical facilities.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Alt, Franz L. "Distributed Lags," Econometrica, 10: 113-28, April, 1942.

2. Bauer, Larry Lee. "The Effect of Technology on the Farm Labor Market." Unpublished Ph.D. Dissertation, North Carolina State University, 1968.

3. ________. "The Effect of Technology on the Farm Labor Market," American Journal of Agricultural Economics, 51: 605-618, August, 1959.

4. Bishop, Charles E. "Economic Development and Adjustments in Southeastern Low Income Agriculture," Journal of Farm Economics, 36: 1146-58, December, 1958.

5. Chappell, Vernon Glenn, Jr. "The Identification and Evaluation of Factors Affecting Economic Growth in the Tennessee Valley Region, 1950-1960." Unpublished Ph.D. Dissertation, The University of Tennessee, 1970.

6. Davis, Kingsley and Hilda Herty Golden, "Urbanization and the Development of Pre-industrial Areas," Economic Development and Cultural Change, 3: 6-26, October, 1954.

7. Edwards, Clark. "Non-linear Programming and Non-linear Regression Procedures," Journal of Farm Economics, 44: 100-14, February, 1962.

8. Fox, Karl A. and T. Krishna Kumar, "Delineating Functional Economic Areas," in Research and Education for Regional and Area Development, Ames, Iowa State University, 1966.

9. Friedmann, John R. P. The Spatial Structure of Economic Development in the Tennessee Valley. Department of Geography Research, Paper No. 39 University of Chicago (Chicago: by the author, 1955).

10. Fuller, Wayne A. and J. E. Martin, "The Effects of Autocorrelated Errors on the Statistical Estimation of Distributed Lag Models," Journal of Farm Economics, 43: 71-82, February, 1961.

11. Griliches, Zvi, "Distributed Lags: A Survey," Econometrica, 35: 16-49, January, 1967.

12. Hoselitz, Bert F., "The Role of Cities in Economic Growth of Underdeveloped Countries," Journal of Political Economy, 61: 195-208, June, 1953.

13. Iowa State University of Science and Technology, Department of Economics. The Role of Growth Centers in Regional Economic Development. Washington: The U. S. Department of Commerce, 1966.

14. Jorgenson, Dale W. "Rational Distributed Lag Functions," Econometrica, 34: 135-49, January, 1966.

15. Koyck, L. M. Distributed Lags and Investment Analysis. Amsterdam: North Holland Publishing Company, 1954.

16. Krutilla, John V. "Criteria for Evaluating Regional Development Programs," American Economic Review, 45: 120-32, May, 1955.

17. Moore, John R. (ed.). The Economic Impact of TVA. Knoxville: The University of Tennessee Press, 1967.

18. National Industrial Conference Board. Economic Almanac 1967-1968. New York: The Macmillan Company, 1967.

19. Perloff, Harvey S., Edgar S. Dunn, Jr., Eric E. Lampard and Richard F. Muth. Regions, Resources, and Economic Growth. Baltimore: The Johns Hopkins Press, 1960.

20. The President's Water Resource Policy Commission. A Water Policy for the American People. Volume I. Washington: U. S. Government Printing Office, 1950.

21. Ratchford, B. U. "Government Action or Private Enterprise in River Valley Development: An Economist's View," American Economic Review, 41: 299-306, May, 1951.

22. Robock, Stefan H. "An Unfinished Task: A Socio-Economic Evaluation of the TVA Experiment," in The Economic Impact of TVA, ed. John R. Moore. Knoxville: The University of Tennessee Press, 1967.

23. Sales Management--Survey of Buying Power. Vols. 72-103. 1935-1970.

24. Schultz, Theodore W. "A Framework for Land Economics: The Long View," Journal of Farm Economics, 33: 204-15, May, 1951.

25. Schultz, Theodore W. Transforming Traditional Agriculture. New Haven: Yale University Press, 1964.

26. Smith, Adam. An Inquiry into the Nature and Causes of the Wealth of Nations (The Modern Library Edition). New York: Random House Inc., 1937.

27. Solow, Robert M. "On a Family of Lag Distributions," Econometrica, 28: 393-406, April, 1960.

28. Tennessee Valley Authority. "Notes on Persons Engaged in Agriculture and Wage and Salary Employees in Nonagricultural Establishments, 1929-1952," unpublished monograph, Tennessee Valley Authority, 1963.

29. Thunen, Johann Heinrich Von. Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalokonomies. Berlin: Schumacher-Zarchlin, 1875.

30. U. S. Bureau of the Census. Historical Statistics of the United States Colonial Times to 1957. Washington: U. S. Government Printing Office, 1960.

31. U. S. Bureau of the Census. Statistical Abstract of the United States 1970. Washington, U. S. Government Printing Office, 1970.

32. U. S. Bureau of the Census. United States Census of Population: 1940. General Social and Economic Characteristics, Alabama, Kentucky, North Carolina. Washington: United States Government Printing Office, 1940.

33. U. S. Bureau of the Census. United States Census of Population: 1950. General Social and Economic Characteristics, Alabama, Kentucky, North Carolina. Washington: U. S. Government Printing Office, 1950.

34. U. S. Bureau of the Census. United States Census of Population: 1960. General Social and Economic Characteristics, Alabama, Kentucky, North Carolina. Washington: U. S. Government Printing Office, 1960.

35. U. S. Department of the Army, Corps of Engineers, Personal Interview with Albert Dykes, 1970.

36. U. S. Department of Commerce, Office of Business Economics. Growth Patterns in Employment by County 1940-50 and 1950-60. Volume 5. Washington: U. S. Government Printing Office, 1965.

37. Wiebe, Jacob Edwin. "Effects of Investments in Water Resources on Regional Income and Employment." Unpublished Ph.D. Dissertation, The University of Tennessee, 1970.

38. Zimmerman, Erich W. World Resources and Industries. Rev. ed. New York: Harper and Brothers, 1951.

APPENDIX

APPENDIX

CONSTRUCTION OF VARIABLES

All variables except Public Investment in Water Resources were constructed for each of the three study areas by combining county data for the counties in each of the areas. Public investment in water resources data was reported on a project basis rather than a county basis.

Total Income

Annual estimates of household income are published in Sales Management--Survey of Buying Power (23). Since 1945 net effective buying income has been reported. This series is described as:

> . . . personal income--wages, salaries, interest, dividends, profits and property income, minus federal, state and local taxes. It includes (1) net cash income plus (2) income in kind--payments in noncash goods and services, such as food and housing, and (3) imputed income--food consumed on the farm that produced it and imputed rent of owner-occupied housing. . . . Effective Buying Income is generally equivalent to the Government's "disposable personal income." (23, 1970, p. xiii)

Prior to 1945 gross effective buying income was reported. The difference in the two series is primarily attributable to taxes (23, 1946, p. 19). Both series were reported for 1945, 1946 and 1947. These data, along with estimates of Personal and Disposable income for the U. S. as a whole (30) were used to adjust the data for years prior to 1945 to make the two series comparable. Adjustment of the gross effective buying income series was as follows:

$$n_{ij} = G_{ij} - \left(\frac{P_i - D_i}{P_i}\right)\left(\frac{1}{f_j}\right)(G_{ij})$$

$$f_j = \frac{\sum_{i=45}^{47} \frac{\frac{P_i - D_i}{P_x}}{\frac{G_{ij} - N_{ij}}{G_{ij}}}}{3}$$

where n_{ij} = computed net buying income in year i for area j

G_{ij} = reported gross effective buying income in year i for area j

P_i = National Personal income in year i

D_i = National Disposable income in year i

f_j = adjustment reflecting differences in the national series on Disposable and Personal income and the <u>Survey of Buying Power</u> series on net and gross effective buying income

Income data were adjusted to constant 1957-59 prices using the consumer price index. (30, 31)

Employment

Employment data were based on series published by the Employment Security Commission and the Social Security Commission. Data, furnished by the TVA, had all employment in states allocated to specific counties. The data construction is described in <u>Notes on Persons Engaged in Agriculture and Wage and Salary Employees in Nonagricultural Establishments: 1929-1952</u>. (28)

National Employment Rate

The United States Department of Commerce series on the unemployment rate (30, 31) was used to calculate the national employment rate. The employment rate for any year equals 100 percent minus the unemployment rate in that year.

Public Investment in Water Resources

Three measures of public investment in water resources were used in the study. Data on TVA investment was furnished by TVA. Investment on Barkley Dam, a Corps of Engineers Project on the Cumberland River near Paducah, Kentucky, was obtained from the United States Corps of Engineers (35). Data were converted to constant 1957-59 dollars using the United States Department of Commerce composite construction index (30, 31).

Annual Investment in Dams in the Entire Tennessee Valley Region

Annual investment by TVA in all dams plus investment in Barkley Dam by the Corps of Engineers.

Annual Investment in Dams and Steam Plants in the Study Areas

Annual investment in dams and steam plants on projects within each study area comprise this measure of public investment. The Huntsville and Asheville data include only

TVA expenditures while the Paducah data includes TVA projects plus Barkley Dam, built by the Corps of Engineers.

<u>Cumulative Investment in Dams and Steam Plants in the Study Areas</u>

The projects included in the respective areas for this measure of public investment were the same as the annual investment in public investment in dams and steam plants in the areas. The cumulative investment in year t was the value of all investment prior to year t in projects in the respective areas plus the new investment in year t.

Educational Level

A weighted average of census data by counties on median school years completed by persons 25 and over was computed for census years for each study area. The source of this data was the United States Census of Population (32, 33, 34). These weighted averages were used to establish a trend line (using a semi log scale) from which intercensural years were determined. These data were not included in the regression equations.

Capital Intensity Index

The index of capital intensity was constructed to reflect change in an area's industry mix which has an impact upon capital invested per production worker. Industry categories adopted as well as employment reported in each category was from Growth Patterns in Employment by County 1940-1950 and 1950-1960 (36). Investment data for the various industry categories was Capital Invested per Production Worker (in Book values) in all Enterprises in Specific Manufacturing Industries (18). Where the industry categories did not coincide in the two sources a weighted average figure was used. The capital invested per production worker in each industry used in calculating the index was an average of 1963 (the latest year reported), 1959, 1949, 1939.

The computation of the index was as follows:

$$I_{ij} = \sum_{k} \frac{\frac{E_{ijk}}{E_{ij}}}{C_k} \qquad (k = \text{all } k\text{'s})$$

$$C_k = \frac{\sum_{i} \frac{C_{ik}}{C_0}}{4} \qquad (i = 1963, 1959, 1949, 1939)$$

where

i = year, j = area, and k = industry category

E_{ij} = total manufacturing employment in year i for area j

E_{ijk} = manufacturing employment in industry category k, in year i, for area j.

C_k = Average capital invested per production worker in industry k

C_0 = Average capital invested per production worker in all manufacturing

The index was computed for census years and interpolation was used to assign values for the remaining years. These data were not used in the regression equations estimated in the study.

VITA

Audley Eugene Hileman was born on December 4, 1936, in Armstrong County, Pennsylvania. His parents are William Jaye and Alice Piper Hileman. He attended the public schools of Manor Township and Ford City, Pennsylvania, from which he graduated in 1954. He then attended Warren Wilson College, receiving an A.A. degree in 1956; The Pennsylvania State University, receiving his B.S. in Agricultural Economics in 1959; North Carolina State University, receiving his M.S. in Agricultural Economics in 1962; Duke University; and the University of Tennessee, receiving his Ph.D. in Agricultural Economics in 1971.

Mr. Hileman has been employed by North Carolina State University as a research assistant and as an extension specialist in farm management. From 1963 until 1971 he was assistant professor and chairman, Department of Economics at Maryville College. He has accepted a position as Professor and Chairman, Department of Business and Economics at Central Methodist College beginning September 1971.

He is married to the former Leila Bleecker Harbison. They have three children--William Audley, Leila Bleecker and Robert Eugene.