Estimation of First Round and Selected Subsequent Income Effects of Water Resources Investment

IWR

INSTITUTE FOR WATER RESOURCES

DEPARTMENT OF THE ARMY
CORPS OF ENGINEERS

FEBRUARY 1970

IWR REPORT 70-1

U.S.-C.E.-C

INSTITUTE FOR WATER RESOURCES

In 1969, the Subcommittees on Public Works, House and Senate Appropriations Committees, authorized formation of the Institute for Water Resources to assist in formulating Corps of Engineers water resources development planning and programming. Improved planning and evaluation concepts and methods are needed if development and utilization of these resources are to meet the public objectives of economic efficiency, protection and enhancement of the environment, regional development, income distribution, and general well-being of people. The Institute and its two Centers--the Center for Advanced Planning and the Center for Economic Studies--carry out both in-house and by contract, research studies to resolve conceptual and methodological problems involved in the development of our Nation's water and related land resources.

The Institute welcomes views and comments on its program and publications.

R. H. GROVES
Brigadier General, USA
Director

70-1

ESTIMATION OF FIRST ROUND AND SELECTED SUBSEQUENT INCOME EFFECTS OF WATER R IENT

ESTIMATION OF FIRST ROUND AND SELECTED SUBSEQUENT INCOME EFFECTS OF WATER RESOURCES INVESTMENT

A Report Submitted to the

U. S. Army Engineer Institute for Water Resources
206 North Washington Street
Alexandria, Virginia 22314

by

University of Chicago
Chicago, Illinois

Under Contract Number DA49-129-CIVENG-65-11

Edited by

George S. Tolley
Department of Economics
University of Chicago

CLEARINGHOUSE
for Federal Scientific and Technical Information
Springfield, Virginia 22151

FEBRUARY 1970

IWR Report 70-1

NOTE TO THE READER

This report is a product of research conducted by staff members associated with the North Carolina State University and subsequently with the University of Chicago under contract to the Corps of Engineers.*

The report develops methods for estimating several effects of water resource projects. These include: benefits of providing water-related recreation opportunities, effects of providing additional water supply on industrial location decisions, and impacts of water resource development on unemployment and income. Additional research considers reasons for variation in regional multiplier values.

The research draws on and contributes to closely related research conducted for the Corps of Engineers by Washington University.** Whereas the primary concern of the effort at Washington University was to estimate local effects, the primary concern of the research considered in the present report is to estimate national income benefits.

Since this study presents results independently arrived at by the researchers, it does not necessarily reflect the official position of the Corps of Engineers. Any comments you may have will be most welcome.

---

*Contract No. DA-49-129-CIVENG-65-11.

**See Development Benefits of Water Resource Developments, published by the Institute for Water Resources, November 1969, IWR Report 69-1.

TABLE OF CONTENTS

Title	Page No.

Title	Page No.

LIST OF FIGURES

LIST OF TABLES

PREFACE

The research effort underlying this report included the participation of many individuals. Authors of the several research papers included in this report are listed below:

Each of these studies contributes to the overall purpose of evaluating water resource projects in depressed areas.

A depressed area may be defined as a locale where indicators of well-being are persistently below those for the economy at large. The indicators are low even after allowing for differences in age composition, education and other characteristics associated with skill differences. Unemployment rates tend to be high, and there is usually underemployment, i.e., people even though working do not appear to be earning as much as they are capable of earning.

Increased expenditures on water resource projects have been among the measures prominently considered in attempting to raise indicators of well-being of people in depressed areas. As a result, there is an intensified need for methods which can be use in planning projects in depressed areas and in deciding whether to carry them out. The specific aim of the present report is to contribute to estimating quantitatively the extent to which projects in depressed areas contribute to national income. As will be discussed below, to increase national income is not the only goal of projects. Yet it is an important goal having been given the force of Congressional legislative mandate, and it is the only goal which has been consistently accepted by the Executive Branch in quantitative estimation of project benefits.

Chapter I of this report ("National Income and Local Income Effects of Water Projects") provides a framework for the remainder o of the chapters, giving an overview of what is and is not attempted in the report. Many of the national income benefits of projects in depressed areas should be estimated in the same way as national income benefits in non-depressed areas. The present report therefore is supplementary rather than providing a substitute for previously used procedures. Chapter I identifies three ways in which benefit estimation for depressed areas may be different from benefit estimation for non-depressed areas. These are the concern of the remaining chapters.

One way in which benefit estimation for depressed areas may be different is in type of benefit. Many depressed areas are located in mountainous regions with special natural resource conditions. One of the most promising uses of these resources is for recreation. It has been recognized that recreation demands are growing rapidly, but methods for valuing recreation benefits have not been fully developed, particularly as regards differing qualities of recreation experience. Chapter II ("Quality of the Recreation Experience: Estimation of Its Benefits") develops a method for estimating effects of crowding of a recreation facility on the benefits obtained from it. The difference in benefits of two facilities that are the same except in degree of crowding is estimated by valuing the extra time and money costs people would be willing to bear to avoid the extra crowding. The estimate is made possible by observing the relations between degree of crowding among a set of facilities and their distance from a population center whose residents dominate use of the facilities. The residents tend to distribute themselves among the facilities so that they are indifferent between bearing a given degree of crowding at a facility close to the population center or travelling further out where there is less crowding. The reason for the tendency toward indifference is that people have incentives to change from one facility to another as long as there is no indifference. The extra distance from one facility to another thus gives an indication of the effort people are just willing to bear to attain the reduced degree of crowding observed at the more distant facility.

Chapter III ("Demands of Multiple-Purpose Travellers and Route Diverters for Reservoir Recreation") is concerned with visitors whose place of residence is a great distance from a recreation facility. Commonly the entire cost of trip from home to facility and back has been used as an estimate of cost the visitor is willing to bear to visit the facility. This procedure may be satisfactory for most visits not requiring a night away from home. However, the greater the distance, the more likely it is that the trip will have purposes other than solely to visit the facility. For these multiple purpose trips, rather than attributing the entire cost of the trip to visiting the facility, it is more accurate to attribute to the visit only the route diversion or extra cost borne to reach the facility. The method more commonly used should be applied to persons on single purpose trips, and a separate estimate of benefits for persons on multiple purpose trips should b made. Chpater III indicates how to identify visitors at a facility who are on multiple purpose trips and how to estimate recreation benefits for them. Applications are presented based on interviews at several Corps of Engineers reservoirs.

Chapter IV ("Effects of Water Development on Location of Water-Oriented Manufacturing") turns to another type of benefit of special importance for some depressed areas. Many depressed areas have pre-conditions for further industrialization, such as good labor supply conditions. An increase in water availability made possible by a water resource project may therefore lead industries which are affected by water to locate in a depressed area.

In Chapter IV regression analyses are carried out explaining employment at the county level in water-oriented industries. The analyses are based on a model of industry location containing product demands, production functions and supplies of productive factors. One of the factors is water, which in some counties is an economically free good and in others has positive value because it is economically scarce. Water availability can be expected to affect industry location only for those counties where water is scarce, and hence where extra water has positive value. The counties where water has positive value are estimated from iterated regressions. The coefficients of water availability, from the regressions for counties where water has positive value, indicate the effect on industry location to be expected from additional water.

A national income benefit results from the effect of additional water availability on industry location, because a less costly location than would otherwise be chosen is provided for obtaining output of the industries. The saving in cost to firms is dealt with in a study by Edgar Hoover under the companion research endeavor directed by Charles Leven at Washington University. National income benefits in addition to saving in cost to firms will result if effects on industry location result in greater net national employment. Methods for estimating net national employment effects of water-oriented industry location decisions and of other changes in the location of employment due to a project are considered in later chapters of the present report.

So far consideration has been given to the first way in which benefit estimation for depressed areas is different from benefit estimation for other areas namely that a different mix of project purposes is encountered. This is clearly a matter of degree. Just as traditional benefits such as flood control and electric power may be obtained from projects in depressed areas, so recreation and water-oriented manufacturing benefits may be obtained in non-depressed areas. Furthermore, recreation and water-oriented manufacturing benefits are not the only benefits likely to be more heavily concentrated in depressed areas than in other areas. Studies of water transportation benefit have been supported under other Corps of Engineers research contracts. The provision of industrial sites in mountainous depressed areas, where such sites may be scarce, can be a source of benefits not considered in this report. Scale effects, particularly as influenced by size of community, may influence benefits from locating activities in different areas. The research of the present report was designed in the belief that benefits of providing industrial sites and possible neglect of community scale effects in benefit estimation are not quantitatively as important as the benefits considered here, but the neglected effects are nonetheless deserving of study.

A second way in which benefit estimation for depressed areas is different in that an increase in economic activity in a depressed area may lead to a net increase in national employment. The need to estimate net changes in national employment has received much attention, and there have long been calls to develop methods to replace the full employment assumptions traditionally used in benefit estimation.

Chapter V ("Structural Unemployment in the Evaluation of Natural Resource Projects") develops a method for estimating net national change in employment that results from locating a project in a depressed area. The labor characteristics of depressed areas referred to earlier, particularly the prevalence of chronic unemployment, gives reason to expect a net increase in national employment if economic activity is increased in a depressed area. Relatively few persons residing in nondepressed areas are faced with the necessity of migrating if they wish to be employed. In contrast, in depressed areas the proportion of persons finding it necessary to migrate to find employment is great. Employment grows slowly or declines, while family structure is such that the number of persons coming of age to enter the labor force is growing rapidly, or has in the past grown rapidly relative to employment. Universally, depressed areas are found now to have high rates of net migration, or to have had them in the past. When the number of persons who would choose to work in an area exceeds the number of jobs, a proportion of the persons in each age group migrates to other areas if necessary to find a job. If employment is made greater than it would otherwise be in a depressed area due to a water resource project, some of the jobs will be taken by persons who would have remained in the area not working. If the relocation of economic activity due to a water resource project is not away from other depressed areas, there will not be a corresponding decline in employment elsewhere. Non-depressed areas will grow a little more slowly because of relocation of some employment to depressed areas, but they will still be attracting people from depressed areas in order to sustain their growth. People residing in nondepressed areas will still have desired employment unaffected by necessity for them to migrate elsewhere to dind a job. The increase in economic activity in the depressed area will thus increase employment in the depressed area without leading to as great an increase elsehhere, i.e. there will be a net increase in national employment.

Chapter V develops a mthod for estimating how much of an increase in employment in a depressed area will be taken up by persons who would have otherwise remained in the area not working. The method utilizes the observed association between migration rates and measures of employment relative to the number of persons surviving in an area in the absence of migration. The analysis distinguishes among persons according to age, sex and skill, and results in estimates of largest net increases in national employment for areas which have high concentrations of middle-aged male workers of lower skill.

There are several types of unemployment. The method developed in this report estimates net national employment increased due to post construction reduction of geographical structual unemployment. This type of unemployment was singled out for consideration in the belief that it provides the major source of net national increases in employment to be realized from projects in depressed areas. In addition to post construction effects, there may be net national increases in employment during the construction of a project. Furthermore, three types of unemplomment are: (1) unemployment due to deficiency of aggregate demand, (2) structural unemployment due to lack of skills, as for example is manifested in high unemployment rates of certain teen age groups found nationally regardless of the area where they reside, and (3) geographical ctructural unemployment. While this report considers only the thir, for some purposes, such as evaluation of federal expenditures in urban areas, there is need to develop methods of estimating effects on the first two kinds of unemployment. A recent contribution complementary to the present report is the development of a method by Haveman and Krutilla of Resources for the Future, Incorporated, for estimating net increases in national employment from reduction of unemployment due to deficiency of aggregate demand during the construction phase of a project.

There may be national income gains from reductions in underemployment not dealt with in this report. Increases in economic activity in a depressed area may result in more productive employment for persons who have remained in the area in relatively low paying jobs. A chief contributor to the unfavorable income comparisons found for many depressed areas is concentrations of farmers on small outmoded units earning less than persons of similar age and education elsewhere in the nation. Estimates are needed of how many of these persons would find more productive employment if economic activity were increased in depressed areas. It is possible that most of these persons experience underemployment analogous to the second farm of unemployment listed above. They have a skill pattern resulting from lack of industrial experience, and many of them might turn out to be hard core underemployed no matter how great the expansion of employment in depressed areas.

Chapter VI ("Income and Education") is concerned with implications of the fact that depressed areas have low per capita levels of local government expenditures. Local government expenditures on education are investments that make contributions to national income. Because of the importance of local tax payer's income as an influence on education expenditures, increases in income due to a water resource project will affect education expenditures. Education expenditures per pupil rise at a diminishing rate as per capita income rises. Therefore, an increase in income in a depressed area, even if due solely to a transfer of activity from areas with higher per capita income, can be expected to increase education expenditures in the depressed area by more than the decreases in education expenditure caused in the higher income areas. If the rate of return on education expenditures is higher than the rate of return on other investments in the economy, an effect of the net increase in education expenditures will be to raise the present value of national income. Chapter VI develops estimates of education expenditure responses and shows how to use them to quantify effects on national income.

A third way in which benefit estimation for depressed areas differs from benefit estimation for other areas is that estimates of effects on local employment and income are required, even if one's concern is limited to national income benefits. The reason for this has already been brought out. One of the steps in the estimation of net national increases in employment as well as in education expenditures, resulting from depressed area conditon, is to estimate changes in local employment and income.

Methods for estimating changes in local employment and income were developed in the companion study directed by Charels Leven. Local employment and income multipliers used depend on input-output coefficients determining sales from one industry to another. Chapter VII ("Refinement of Regional Employment Multiplier Estimates") shows how to find the effect of errors in input-output coefficients on employment and income multipliers, and

and thus contributes to developing reliability ranges for the multipliers.

As implied earlier, while the present report is concerned with developing estimates of national income benefits, water resource projects in depressed areas may have other kinds of benefits. The other benefits are effects of projects on distribution of income. For instance, a greater value might be given to a benefit accruing in a depressed area than in a non-depressed area, simply because the area is depressed. Alternatively, one might attach values to redistributions of income among persons, giving positive values to redistributions that result in greater equality. More projects might then be built in depressed areas, if building projects in depressed areas do in fact lead to greater income equality. In addition to difficulties in estimating distributional effects, an unresolved problem is how to assign values to different amounts of income redistribution. While income distribution problems are beyond the scope of the present report, they are highly deserving of attention.

NATIONAL INCOME AND LOCAL INCOME EFFECTS OF WATER PROJECTS*

Why Be Concerned with Local Income. Estimates of local effects of projects often called secondary benefits have been estimated because of a belief by some that they are important in their own right, based on the idea that the development of a particular region is an end in itself. A question that arises if this purpose is recognized is what type of local income to measure, e.g. for what area, such as a group of counties or a state or a region, and even more importantly what concept of income. Two contenders for the latter would be total sales and income originating, which is smaller than total sales because of netting out of purchased inputs. Income originating perhaps measured as factor payments, or some similar measure fairly closely related to income of residents of an area would appear more meaningful for decision making than simply total sales.

Another reason estimates are needed of local income effects is that certain national income effects resulting from projects may depend on local effects. The usual concept of primary benefits would appear to be "national income benefits from project purposes" as for example increase in national income due to increase in production on flood protected land. Another concept not typically included which needs to be added when considering depressed areas may be called "national income benefits not from project purposes." An example is the employment of workers who would otherwise be unemployed. Usual procedures of costing

*G. S. Tolley, University of Chicago, Chicago Illinois.

labor at market wages implicitly assume that labor would be fully employed in alternative activities in the absence of the project. Another example is possible effects of projects on productive investments undertaken by local governments, notably education, for national income will be increased if the response to increased local income is to raise expenditures on which the rate of return is high. Because they depend on changes in local government revenues, some of these "national income benefits not from project purposes" depend on local income effects. This provides then another reason for estimating local effects in connection with which it is very important to use a method where the national income effects and local income effects estimated are consistent with one another.

Figure 1 shows a situation where there are national income benefits from project purposes of ten at point α in the total economy. In many cases, this primary benefit of ten is the market value of production associated with the project purpose less any cost incurred in the production. In these cases it may generally be expected that these benefits will accrue to the owners of the factors of production affected by the project. For instance, benefits would be expected to be approximately equal to the increase in the value of flood protected lands. Thus, in the case discussed the national income benefits from project purposes would accrue locally at the project site. The increase in production of particular kinds at the project's site, e.g., crop, power, will displace this kind of production presumably at highest marginal cost areas in the rest of the economy, freeing those resources

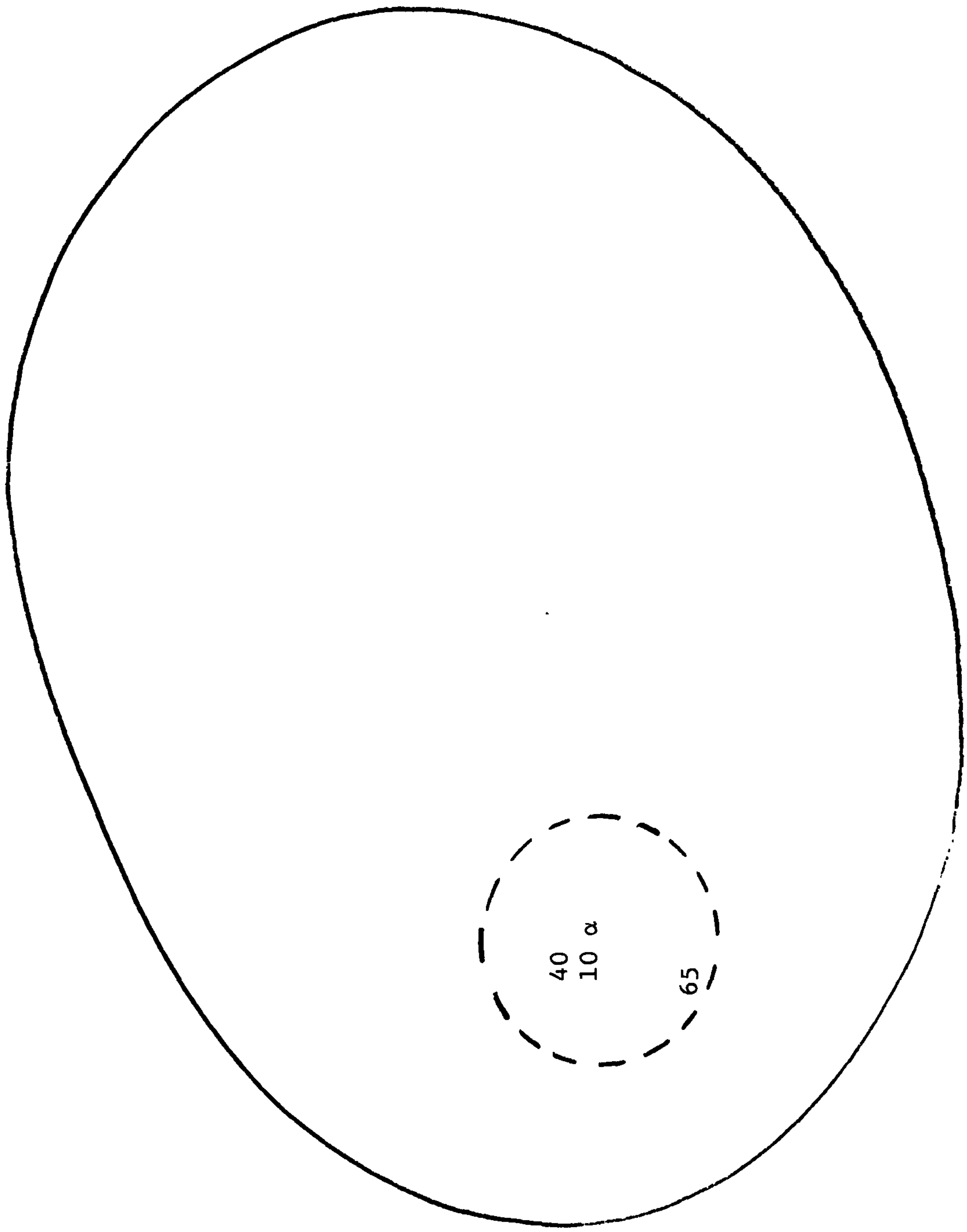


Figure 1 - National and Local Income Effects

to produce the things that are demanded due to the increased income from the project, e.g., houses, automobiles and so forth as determined by income elasticities of demand. If there is definite evidence as to where these displacements from marginal areas are in the rest of the economy, they could be included in estimates of spatial impact. However, if they are in areas that are experiencing general economic growth, it may be expected that the displacements tend mainly to make for a small reduction in rate of growth in those areas as resources are diverted to new kinds of production. Thus in the absence of any better knowledge it may be reasonable to assume that the displacements do not result in unemployment and hence their particular geographical impact on the rest of the economy could be ignored in estimating net national income benefits.

The Magnitude of Spatial Impacts. So far then the reasoning suggests primary benefits of ten at point α with no necessity to consider impacts elsewhere in the economy due to these primary benefits. The national income benefits from project purposes may be viewed as resulting from increased factor supplies. In effect the project produced factors of production which are used to increase national income. In the case cited, the supply of land was increased. Other factors of production will be used in the production resulting from the project purpose. For instance, most importantly, labor resources will be used in producing the additional agricultural crops, power or other output at point α. In a full employment economy, without the project, this labor would have been employed in nonproject activity elsewhere in the economy.

Suppose the yearly payment to the labor is 40. Then the increase in income originating at point α is 10 plus 40, or a total of 50.

The increase in production at point α will lead to drawing away of economic activity from other parts of the economy to the vicinity of point α because of inter-industry effects (local purchases of firms) and local household consumption. What can be said about the nature of, in particular the size, of the area in the vicinity of point α to which activity will be drawn, and how can the income of this activity be estimated? As a valid approximation likely to be satisfactory on the average, it is suggested that all local impacts be assumed to occur within a range determined by daily driving times governing commuting habits and firm deliveries. This is the area enclosed by the dashed circle surrounding point α in the figure. It is assumed that the inter-industry effects and local household consumption result in 65 dollars of additional income originating in this area which is drawn away from other parts of the economy.

For estimating the increase in income originating in this area no commonly used technique is available. However, local employment multipliers derived from input-output analysis taking the household sector as endogenous and using coefficients appropriate to the type of area delineated by the dashed circle, could be used to obtain an estimate of well over half of the income originating in the area due to the project. This would be done by treating the increase in employment of forty associated with the project as due to an increase in final demand of the area, estimating the total area effects from the multipliers obtained by inverting the employment input-output matrix. By employment input-output matrix is meant the set of coefficients showing the increase in employment in the j^{th} industry resulting from an increase in employment of one in the i^{th} industry. If employment is expressed in terms of dollar values then the resulting estimates give labor income originating in the area as a result of the project. If the coefficients are in terms of numbers of people or man hours of employment, the answer would have to be converted into an estimate of income perhaps using some estimate of average wage rates.

Ideally there would be available a companion input-output matrix for nonlabor resources and a set of coefficients indicating a fraction of income to these resources paid to residents of the area as opposed to residents of the rest of the economy. Lacking these refinements, it is suggested that the increase in labor income estimated as described in the preceding paragraph be multiplied by what appears to be a reasonable estimate of the ratio of total income in an area to labor income based on examination of census and other statistics.

It has been noted that benefits from saleable outputs due to a project purpose tend to accrue to owners of factors whose value is increased which in most cases will be at the site of the project. Recreation and other nonsold benefits may accrue directly to users rather to owners of factors of production. In so far as these users come from outside the dashed circle, it is incorrect to assume that all the primary benefits accrue in the local area. However, it is believed that the error from assuming all the primary benefits to accrue to the project area is likely to be very small. For Corps facilities, which are ordinarily not national attractions drawing tourists who are several nights stay away from home, almost all the visitors will be from within the local area. The not completely unrealistic numbers given above suggest in any case that the primary benefits are small in relation to the total area income effects, so that the percentage error in any case would be small for the area. The remarks suggest that "national income benefits from project purposes" generally accrue within the dashed circle and are nonwage income. They are thus part of the income considered in the previous paragraph. The national income benefits from project purposes and changes in nonwage income in the area should be considered together to ensure that national income benefits are included but that there is not double counting.

It likewise appears that in the absence of definite knowledge to the contrary, "national income benefits not from project purposes" should be assumed to accrue in the project area. This is where the additional employment from hiring labor which would otherwise be unemployed almost certainly accrues. The returns on local government investments in adding to or improving facilities in the area likewise must be local. The returns on education will accrue to the affected children and depend on where they will live as adults. However, in the procedures of benefit estimates being suggested as a result of our research, the national income benefits due to increased education are not used as a basis for any other estimation and so it is immaterial where they accrue.

Time Streams. This discussion has concerned post-construction production which is the period when national income benefits are realized. We have not been concerned with the construction phase although national employment generating effects might be important for projects constructed in depressed areas. This consideration reducing the social costs of the project could be estimated by adapting the methods we have developed for the national employment effects for the post construction phase, or the method developed for the construction phase by Haveman and Krutilla which is different from our method might be used for the construction phase. Our estimates for income effects in the post construction period assume no additional employment in connection with creating the increased amounts of nonhuman resources in the local area, which might increase employment in the construction industry for several years but which it is hoped in the aggregate would be small especially in a depressed area where many redundent physical facilities exist. Aside from buildings, most nonhuman resources such as machines might be imported.

In our benefit estimation average annual effects would be estimated for a year occurring between a third and a half way through the project life. More ideally the total time stream of effects would be estimated, but we are not convinced of the feasibility of this for the benefits we have been concerned with particularly if the method designed is to be widely used in the field. Choosing a distant year not over half way through the project implies a constant time stream of benefits at the levels estimated for that year. It is desirable that the present value of this time stream be as close as possible to the present value of the actual time stream that will occur where benefits through time may sometimes be rising and sometimes falling. Using a year more than ten years beyond completion of the project gives time for long-run effects to emerge, but choosing a year not over half way through the life of the project may help achieve the approximation in as much as the compounding procedure gives greater weight to near than far years. It appears desirable to base estimates on relatively near years since these effects are more certain and can be estimated more reliably than for later years that would be beyond the half-way point of the project's life.

QUALITY OF THE RECREATION EXPERIENCE-ESTIMATION OF ITS BENEFITS*

Strategies for Benefit Estimation

In our thinking about methods of approach to the problem of benefit estimation, two basically different strategies have suggested themselves to us. The first strategy would be to extend the basic recreation demand analysis, which has been developed by Hotelling and Clawson using travel costs, by Knetsch using land values and by others, into analysis of demand for recreation quality. Given sufficient user data and data on types and qualities of recreation experiences, there could be much fruitful research leading to improved methods for recreation demand analysis. In this paper we report on some of the work we have done along this line of approach.

The second strategy would be research into the possibility of deriving the potential demand for certain basically aesthetic quality characteristics of the recreation experience by considering the demand for characteristics in other fields. One might, for example, observe how much persons are willing to pay for lawns and freshly painted houses. Then, by judgment or interview, one might establish a measure of comparability between aesthetic outputs.

Elements of the Quality Package

Before proceeding to describe more fully the lines of approach and work done, it will be useful to sketch out and classify the elements of the quality package. Four broad classifications can be made:

(1) Quality of the water (water quality);

(2) Quality of the physical surroundings;

(3) Quality of the physical setup; the mix of recreation facilities;

(4) Quality of crowdedness or intensity of use.

---

*V. S. Hastings, Delaware River Basin Commission and G. S. Tolley, University of Chicago.

Now there may be both aesthetic and nonaesthetic considerations in each of these classifications. Thus, for each of the four broad classifications, our first strategy for benefit estimation may simultaneously capture aesthetic and nonaesthetic elements. However, following our second strategy, it may be important to separate off the aesthetic elements, such as scenic, aromatic, and tactile qualities of the water, from the nonaesthetic elements, such as the health and fish survivial qualities of the water.

Two considerations make the second strategy for aesthetic benefits estimation seem important. First, for many recreation experiences the aesthetic elements may be the important elements. Indeed, the aesthetic elements may make all the difference from one park to another. Second, there may not be any easy or immediate way to measure directly the demand for these elements and an indirect means may be needed, at least an interim strategy.

Certain quality elements are actually "unique" elements, such as natural "wonders" like the Grand Canyon. Since these wonders cannot be reproduced, we do not have a direct interest in their value estimation. However, using demand analysis, we would want to separate off the values attributable to such uniqueness.

Research Progress

Our actual research to date has followed the first strategy. Basically, we have extended the Hotelling-Clawson approach. We have so far worked in three fields. The first field deals with the scenic value of a reservoir area. In this the question is asked, "How far will persons go out of their way just to see or pass by a more scenic area?" The second field deals with the quality of the recreation experience with respect to the mix of individual recreation experiences available at a reservoir site. The third is estimating benefits with respect to the quality of crowdedness or intensity of use of reservoir sites.

We will first give a short review of the Hotelling-Clawson approach and then report on some of the conceptualization and research activities in these fields.

Hotelling-Clawson Approach

The original Hotelling-Clawson approach is based on observing the proportion of the population at given distances that bears travel costs plus entrance fee to visit a facility.[1] It assumes that travel cost plus entrance fee is a predictor of proportion of the population that will visit a facility. For instance, suppose one wished to predict the effect on attendance of raising the entrance fee of a facility by a certain amount. The predicted proportion of the population that will visit from a given distance A with the new entrance fee is equal to the proportion that were previously visiting from the further distance which, before raising of fees, had the same travel cost plus entrance fee as now faced by people at distance A.

Among others, Lerner,[2] Boyet,[3] and Merewitz[4] have used this basic approach in their research.

---

[1]Marion Clawson, Methods of Measuring the Demand for and Value of Outdoor Recreation, Resources for the Future, Inc., Washington, 1959.

[2]Lionel J. Lerner, Quantitative indices of recreational values, (paper presented at Reno, Nevada, August, 1962.)

[3]Wayne Elwood Boyet, Area complex and national park recreation demand projection, (unpublished doctoral thesis, North Carolina State University, Raleigh, 1964.)

[4]Leonard Merewitz, Recreational benefits of water resource development, (Harvard Water Program, mimeographed, June, 1964.)

Scenic Value of a Reservoir Area

If on a drive or trip from one point to another a driver diverts to a route that takes him by a reservoir area rather than by the most direct route, it would seem that the costs of this diversion might be a measure of the scenic value of the reservoir area. Indeed, using the Hotelling-Clawson approach, we might in some cases be able to derive demand schedules for the aesthetic enjoyment of the scenery.

Consider a reservoir, R, as shown in Figure 2, enjoyed by travelers traveling over the section of road marked MN. Table 2 shows distances between cities A, B, C, and D by the shortest routes (column 2 and by going by the reservoir (colum 3), the extra distance traveled, and cost (columns 4 and 5). The proportion of travelers bearing the extra costs of diverting to pass by the reservoir are shown in column 8. The regression of the proportion on the extra cost, shown in Figure 3, is the per capita demand schedule for such visits to the reservoir. Data for such analysis might be obtained by interviewing travelers along routes typical of the kind depicted in Figure 2.

Mix of Facilities

Now let us turn to the estimation of the effect of the facilities mix on the quality of the recreation experience. Visitation rate will depend, among other factors, on the character of the facilities mix at various recreation sites. Various facilities, or the activities which go with them, may be complementary, such as boating, fishing, and picnicking. Such complementarity would suggest a high payoff to achieving a balanced mix among such activities at each site. Other activities

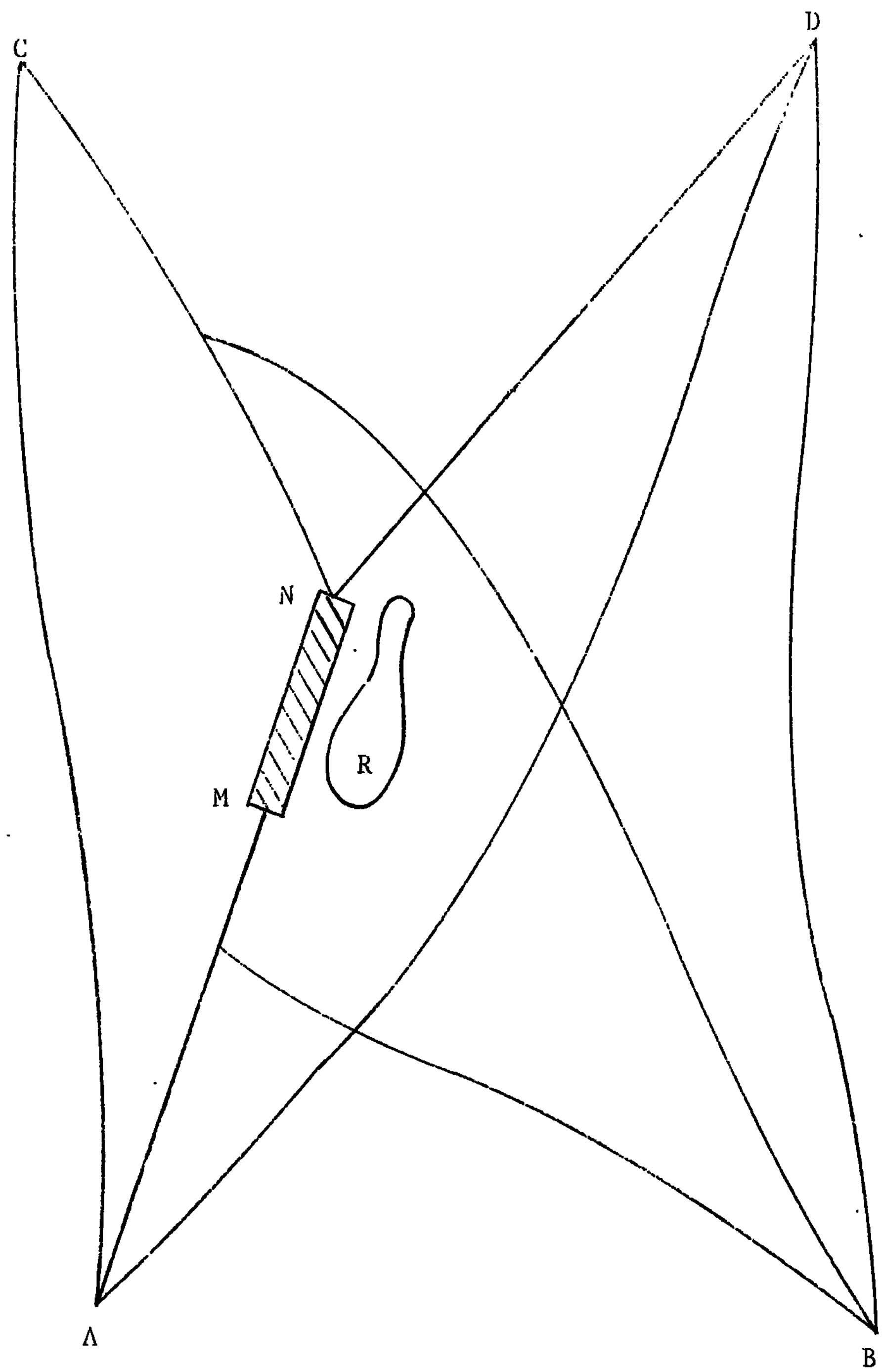


Figure 2 - Reservoir and Route Locations

Table 1. Effects of reservoir on travel

Intercity travel	Shortest route miles	Going by reservoir miles	Distance out of way	Cost	Total intercity travelers per day	Intercity travelers going by reservoir	Proportion going by reservoir
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A - D	20	21	1	$ .10	1,000	300	.30
A - C	15	18	3	.30	2,000	200	.10
B - C	22	27	5	.50	1,500	75	.05
B - D	16	28	12	1.20	3,000	30	.01
Total						605	

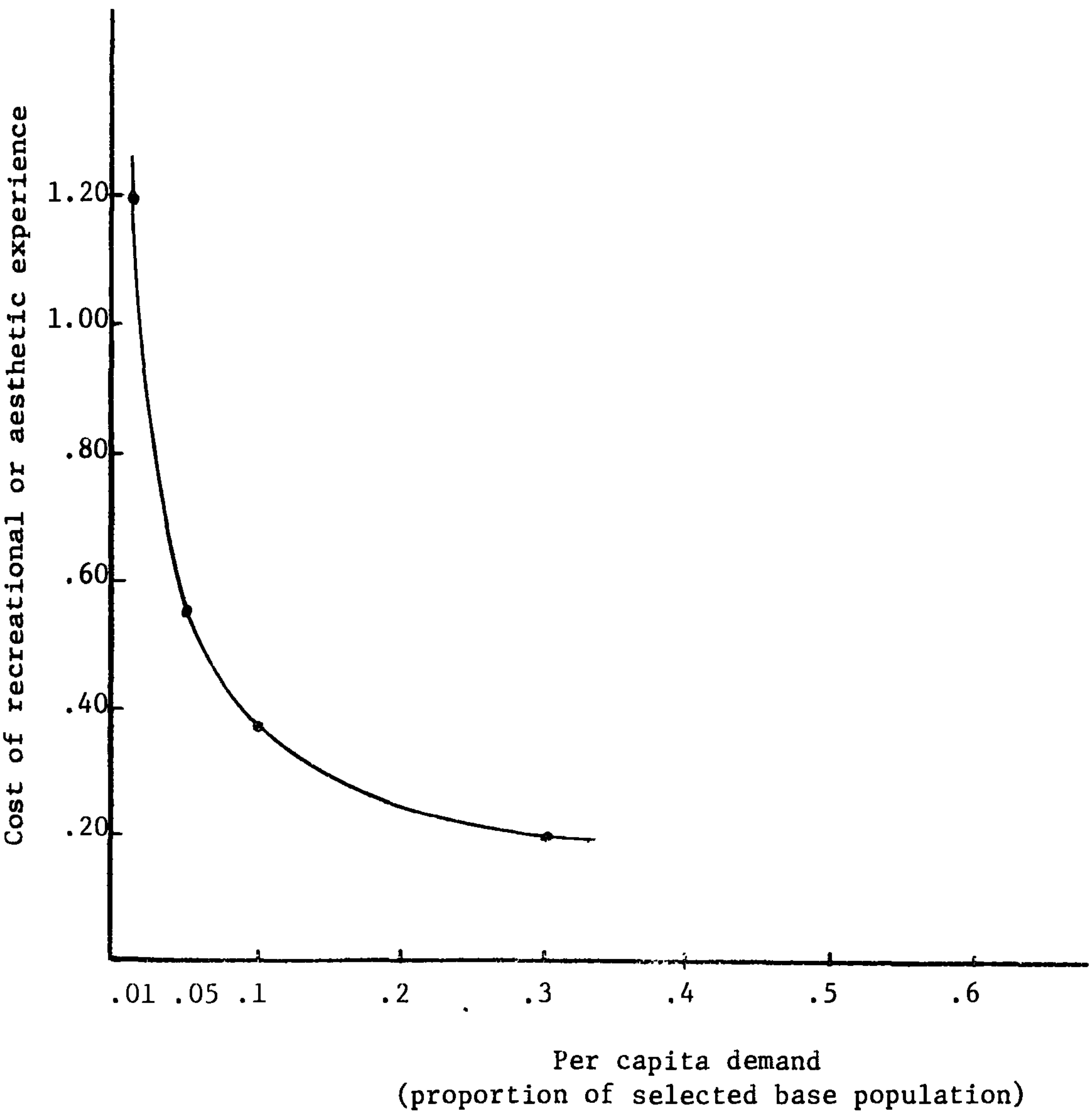


Figure 3 - Per Capita Demand

might be competitive, such as fishing and water skiing. Such competitiveness would suggest a high payoff to specialization of site use, or to separation of a site into zones of use. We visualize obtaining data from various parks and running multiple regressions of the Hotelling-Clawson type. Measures of the amount of facilities for various activities, such as number of picnic tables, length of swimming beach, number of fishing boats, etc., would be used as independent variables. Measures of the effect of interaction among variables would be an important part of this analysis.[1] Finally, because we have a "price" variable in a Hotelling-Clawson type demand schedule, we would implicitly obtain, from the regressions, value or demand schedules for various inputs and mixes. A major problem in this area of research is one of data.

Crowding or Intensity of Use

Finally we turn to crowding.

Crowding of recreation areas appears to be a definite phenomenon on summer weekends at recreation sites around metropolitan areas. Probably the best known example of such crowding is Coney Island. Indeed, Coney Island is almost a synonym for crowding.

Crowding and the Demand for Space

It is our hypothesis that crowding has a negative quality value. That is, crowding will decrease the demand for visits to a recreation site. For a given cost for visits, there will be fewer visits than

[1]For an example of this type of investigation see Paul Davidson, F. Gerard Adams, and Joseph Seneca, The social value of water recreational facilities resulting from an improvement in water quality in an estuary: The Delaware, a case study, (presented at Denver, Colorado, July, 1965.)

there would have been had the recreation facilities been built to hold larger capacities. Or, vice versa, for a given number of visits, the price that visitors would be willing to pay will be lower for a more crowded facility.

This is a negative statement. We prefer to turn it around so that we can talk about the demand for something positive, for space or elbow room. Persons will pay more for more elbow room.

We have frequently heard the argument that persons, or at least important fractions of the people, prefer crowding. People get together for a picnic, not apart for a picnic. Teenagers like to find lots of other teenagers. There may be truth in such statements. However, we contend that it is the marginal visitor that counts and, on the basis of limited but persuasive empirical evidence, he moves away from crowds.

The empirical evidence--Our empirical evidence comes from Long Island. Data were obtained from ten water-based state recreation parks. These parks range from lake and stream-based picnic areas to major beaches and range in distance from Manhattan from about 18 miles to about 150 miles (Valley Stream, just outside the city limits, to Montauk Point). The total number of annual visitors ranged upwards to 12 million at Jones Beach. As a measure of space or elbow room, we took number of acres of per annual visitor. We would have preferred to have had some measure of peak use, but this was not available, so we took the next best thing.

With these data, we regressed acres per visitor on distance of the park from Manhattan. Statistically significant relationships were found in the linear and log forms. The highest R^2, 0.88, was found in the quadratic form. The relationship in this form was:

$$D = 17.337 + \underset{(3.01)}{24.725X} - \underset{(1.38)}{1.306X^2}$$

where D = distance and X = acres per annual visitor x 1,000.

We regressed acres per visitor on distance rather than distance on acres per visitor for two reasons. First, to avoid confusion. It is conventional to put the price variable (and the dependent variable) on the y axis. More important, we were concerned with individual decisions and demand schedules. The individual takes crowding at various parks as given and chooses the distance he wants to travel. Thus, looked at in this way, D is the dependent variable despite the fact that distances to parks are not determined by crowding.

Another point about this regression, we have ignored entrance fees and have done this for two reasons. First we think they may be small relative to distance costs. Second, fees tend to be charged at and are about the same at, all ocean and sound parks. Thus, at least among these parks, fees would have little differential effect.

Now what does this relationship between distance from Manhattan suggest? Reflection might suggest a number of hypotheses for explaining the casuality of this significant relationship. However, one plausible hypothesis is that the nearer parks are more crowded because it costs less in gasoline, less in time (which includes lost time recreating), and in irritation on the parkways to get to the nearer parks. But the relationship suggests something further. Recreationists do not all crowd themselves in the park nearest to the population center or in any one park but do spread themselves out. (No data are available at this time on source of visitors to parks by township or borough, but we do believe that a casual observation of traffic on the parkways on summer Saturdays and Sundays would suggest that a large proportion of visitors at almost all parks are from the heavily populated metropolitan center, including the New York boroughs and southeastern Nassau County. One observes traffic moving to and from New York. We would claim that one observes very little traffic moving to and from each park from a pulling radius around each park.) This spreading out suggests that a number of persons, possibly everyone part of the time, is willing to pay extra to get to a less crowded park. We have heard two other explanations for this willingness or desire to travel farther. One is that persons desire variety. This hypothesis is difficult to reject on the basis of the evidence. However, we would point out that a part of the variety which persons obtain in traveling farther is a less crowded park. The second explanation is that, at least part of the time, persons want and enjoy a drive--a longer drive. One of the authors having lived right on the major parkway to Jones Beach, and about eight miles from Jones Beach for five years, we think we can speak with some experience on this point. We do not believe anyone could enjoy a drive on a Long Island parkway on a summer Sunday evening. It sounds good, but that's all. And we do not believe we ever heard of anyone purposely going out on a Long Island parkway inbound to New York on a summer Sunday night <u>just for a drive</u>.

A third hypothesis can be stated concerning the relationship we have found between crowding and distance, that is, that it is an equilibrium relationship. That is to say, everyone has generally satisfied himself with the amount of elbow room which he has "purchased." Persons have spread themselves out in a way which just creates an equilibrium situation between how much room they have and how much more or less room they could get in exchange for more or less travel. (If they felt they could do better, they would go to other parks until equilibrium is established. This is in terms of expectations for any one weekend, of course. Persons could, of course, also stay home. We will take up later, however, the matter of the effect of crowding at all parks on the total visitation rate.)

One can immediately see the elements of a demand relationship--the demand for a quality factor, space at parks, in terms of a price factor, distance. Indeed, the data from the equilibrium situation, if it is in fact brought about by people's demand for space in terms of its cost, then describes an aggregate demand schedule, since in equilibrium everyone must be on his demand schedule.

For a typical demand relationship, however, we want not total value but marginal value: a cost per unit of space, that is, a price. For this purpose, using distance as cost, we take the first derivative of distance with respect to space per person in the original relationship and obtain: $\frac{dD}{dX} = 24.725 - 2.612X$

Figure 4 shows an example of a regression of acrea per visitor on distance and its first derivative, the schedule of demand for space.

This is a typical demand schedule in the linear form, p = a - bq. Here, price is the additional distance to obtain a unit of space, not the total distance traveled. Now what does the equation tell us? The

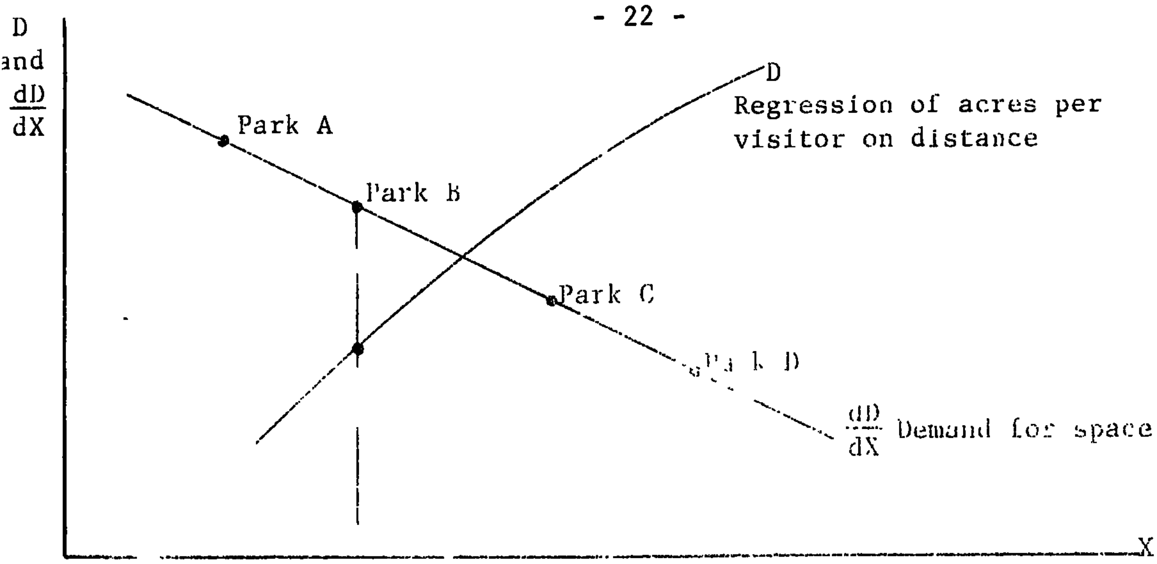


Figure 4 - Relation between distance and space per person

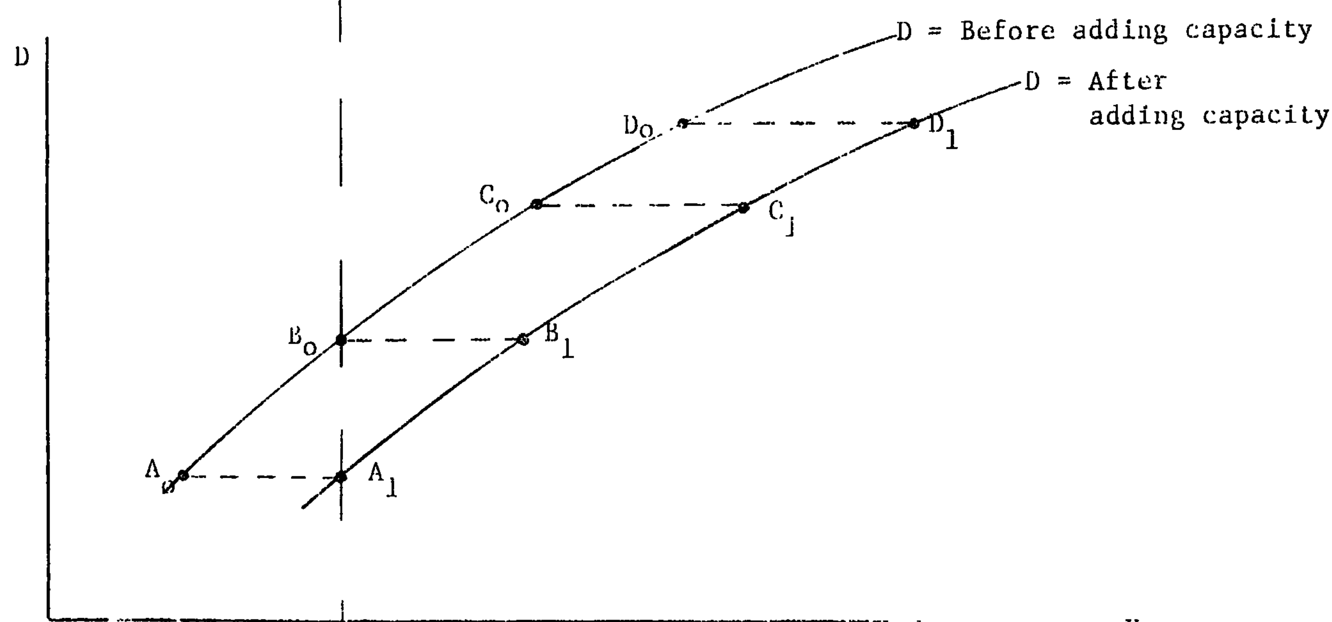


Figure 5 - Effect of adding capacity (redistribution)

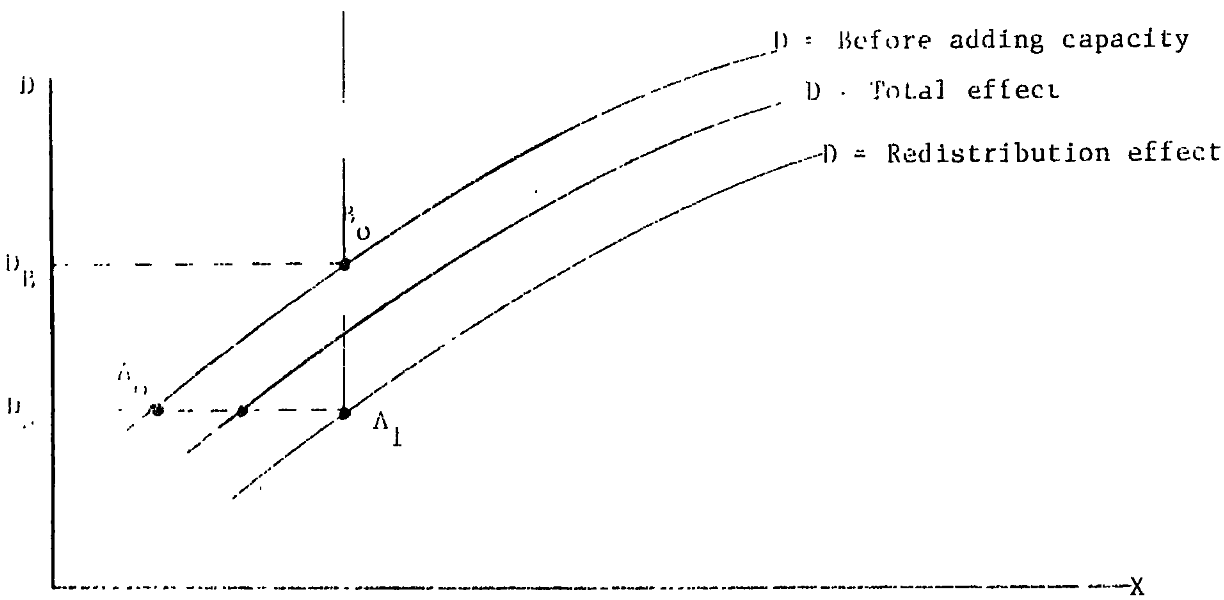


Figure 6 - Redistribution and total effect of adding capacity

equation says that the price, i.e., additional cost one is willing to pay for a unit of space, declines as one obtains more space. A small unit of additional space is more highly valued if one is very crowded but not so highly valued if one is not so very crowded. Indeed, in linear form once space per person reaches a certain level, additional space has no value. We think this is reasonable.

Now this aggregate demand function, which the data describes, is not in every respect the same as the usual quantity-price demand relationship. It is not usual, for example, that each individual finds the price of something facing him moving along the aggregate demand schedule as he decides to take more or less of the thing. That is, it is not usual to find the supply schedule coincident with the demand schedule. But this we have.

Incidentally, if everyone's individual demand for space were the same, we do not think this particular unusual aspect would affect the aggregate demand schedule described. Again, incidentally, if this were the case, specifically where each visitor would be sheer accident. This would be true because, with no differences among individuals, there would be nothing to determine which persons go to near parks and which go to far parks to establish the final equilibrium. A more technical way of looking at this lack of determinacy is this: Since each individual demand schedule is the same as the aggregate demand schedule, which coincides with the supply schedule, there is no intersection of supply and demand.

However, with different individual demands, this unusual aspect may affect the aggregate demand schedule we get. To be more concrete, we have not started out by asking, "If everyone were at the first park with its given crowdedness, how much would the visitors, in terms of aggregate demand, be willing to pay to get to the second park with its given crowdedness." We do not have everyone at the first park, and we do not know for sure what the visitiors at the last park would "demand" if they were at the first park. Differences in tastes have probably dictated partially who is at the first park and who is at the last.

Another situation which we have here, which is probably not usual, is that some persons come from farther distances than others. This may affect individual, and thus aggregate, demand schedules. A person who has already driven an hour may not be as willing to drive another twenty minutes for additional room as a person who just started out.

Despite these differences from the usual aggregate demand schedules, we believe we have what might be appropriately called an aggregate demand schedule, which is useful in itself and useful for further analysis.

Conclusions from empirical evidence.--To recapitulate briefly, we have found that there is a demand for space (for elbow room) at recreation parks. We have measured space in terms of acres per person. We have measured the demand for acres per person in terms of travel distance required to obtain additional space. And, we have found that the additional distance visitors are willing to pay for space is a decreasing function of the space purchased.

Providing Additional Space

If additional space is worth something in the recreation experience, then the above analysis should tell us something about how much any additional space is worth. Then, with knowledge concerning how much space, and its concomitant facilities, and cost, the analysis should begin to tell us how much additional space should be provided.

For the purpose of our analysis which has proceeded from this point, we have assumed that entrance to a park, once the visitor has paid the cost of travel to the park, is free. We have done this not because we think entrance should be free but because a situation of zero or only nominal charges is, in fact, the case in many public parks.[1]

While the elements of the analysis that follows applies in particular to zero entrance fees, the general approach would still be applicable if fees were charged.

To proceed, we know that additional space has a value, and we know that additional space, in the aggregate, can be provided by adding to or building new parks. From this point, we will only sketch in our analysis and where it leads us.

What would happen if an existing park were enlarged, always assuming that concomitant facilities go along with additional space? (The same questions and answers would apply to a new park built close enough to the city.) We can divide the effects into two effects--the redistribution effect on visits and the effect on attracting additional visits.

---

[1]In fact, our analysis suggests guides to pricing policy. A visitor imposes a cost in terms of crowding, which we found to be real and measurable, on others when he enters a park. If optimum park use is to be achieved, he should be charged this cost. He should not be faced with the erroneous signal that he would partake of a nonscarce resour if he entered the park. Our analysis suggests that higher prices should be charged at nearby parks. Units of space are more valuable at nearby parks, as shown in our demand schedule. (In technical terms, individuals faced with a zero price will proceed to use a park to the point where marginal benefits are zero instead of stopping at the point where marginal benefits equal marginal cost. All use beyond this latter point adds more to cost than to benefits and thus decreases net benefits.)

First, consider the redistribution effect. More space per person will be available at the enlarged park and persons will be induced to redistribute themselves toward that park until equilibrium is re-established. If we assume that the aggregate demand schedule for space drived above remains stable, then there is a unique solution for the redistribution effect. Every park will be less crowded. (See Figure 5) If the derivative $\frac{dD}{dX}$ remains constant, then the schedule D moves downward by a constant amount as parks become less crowded. (The distance cost of a unit of space will remain the same at a given crowdedness, but the distance cost will be less at each park since each park is less crowded.)

Now for the second effect. Recreational opportunities of a given crowdedness will be closer to the population center. Thus, in Figure 5, there is the same crowdedness at B_0, park B, before the capacity addition, as at A_1, park A, after the addition. But park A is D_B-D_A closer to the population center. This means park recreation will be more attractive because it now costs less in terms of distance to visit a park with the given crowdedness. Parks, really, are closer despite the fact that no park has been moved.

Now, how do we measure the effect of this shortening of distance on visitation? Will persons visit one more time a year, two more, or what? The conventional Hotelling-Clawson park visitation demand analysis, which gives visitation rate as a function of distance, gives us a unique solution to this problem for a given physical distribution of population.

We could derive this for the Long Island Parks, except that we do not have information on source of visitors. Until we get this, we could use relationships derived from other studies. At present, we can only say this: If the demand for visits were infinitely elastic, the end result would be that parks would be refilled by additional visits to the point where each park was as crowded as it was before. For anything less than infintely elastic, parks would refill only partially.

Measuring Benefits

We have now described the physical effects of adding capacity at a given park. We are now ready to sketch in a procedure for measuring benefits. We can do this by measuring the so-called consumer surplus provided by the park addition. This can be divided into two parts. The first is the part available to original park visits, that is, the visits exclusive of those attracted by the addition. For each of these visits, the travel cost has been reduced by a certain amount by virtue of the fact that a park of a given crowdedness is so much closer. (This distance reduction is the same for all parks.) This part, in distance terms, is equal to the reduction in distance times the number of original visitors. The second part is for the added visits. The added value per added visit will range from the reduction in distance itself for the visits which just needed the added attraction to induce them to zero for those which were just barely attracted by the capacity addition. Since, the linear form, this is just a triangle, the consumer surplus value for the visits attracted is half the reduction in distance times the number attracted. These products would then be translated into dollar units by multiplication by a factor representing cost per unit of distance.

Where and How Much to Build

More acres would be required at a distant park, in order to have the same effect and to provide the same consumer surplus, than at a nearby park. Given the cost of land as a function of distance from the city, one could determine where best to build capacity. And one could compare this cost to the consumer surplus provided by any capacity addition to determine how much capacity to build.

Summary and Conclusions

We have shown how the demand for a particular quality ingredient of the recreation experience can be estimated. We have shown how the benefits of providing more of this quality ingredient can be measured. And we have shown how this can be translated into an action decision. This same kind of analysis could also be used in measuring the demand for and benefits of other quality ingredients and in applying this knowledge to decisions. We are hopeful that the same research methodology can proceed fruitfully into other quality areas, including the areas of improved recreation activity mixes, improved water quality, and improved environmental qualities in general. At the same time, we are hopeful that the indirect approach, utilizing spending for quality fields outside recreation, will provide useful estimates for interim action decisions.

DEMANDS OF MULTIPLE-PURPOSE TRAVELLERS AND ROUTE DIVERTERS FOR RESERVOIR RECREATION*

Summary

A primary assumption underlying travel-distance models for predicting reservoir attendance is that recreators come directly from home to visit recreational facilities. In fact, many recreators "divert" from their travel itineraries to visit a particular facility along the way. Thus, if distance travelled is taken as an indicator of the price paid for a recreational experience (a) new questions must be introduced into visitation questionnaires to permit a more refined analysis of the travel patterns of recreators and (b) modifications of conventional models for estimating recreation benefits are needed.

In field studies of three Corps of Engineers reservoirs, two types of "diverters" were identified. "Route diverters" travel relatively short distances from a tourist artery to make a brief reservoir visit. "Multiple-purpose travellers" make a purposeful effort to visit a reservoir as part of a larger vacation itinerary. The difference between route diverters and multiple-purpose travellers is that the former stop at a reservoir while en route elsewhere, whereas the latter, upon reaching the reservoir, have arrived at one of the destinations of their trip.

Residential population is an inappropriate base against which to measure the per capita recreational demands of diverters. For route diverters, nearby highway tourist traffic is an appropriate base. In principle, the base population for multiple-purpose travellers is the gross displacement of residential populations during the tourist season, in other words, everybody on the move, but no satisfactory measure of this displacement has been devised.

---

*Robert Rugg, University of Ottawa, Ottawa, Canada

The subtraction of diverters from conventional travel-demand schedules decreases the total per capita demand and increases the distance-elasticity of demand. This is because diverters' residences are farther from a reservoir than the residences of most single-purpose visitors. A tentative exploration of route diversion demand indicated that it is both lower and less distance-elastic than single-purpose demand.

There appears to be a need for more detailed attention to factors independent of travel costs which influence reservoir visitation. Demands for different kinds of recreational experiences, and variations in demand according to the social characteristics of recreators, might be analyzed. For example, nearly all recreators appear to be urbanites, suggesting the use of urban rather than county census data in estimating residential base populations.

Virtually all recreators travel in automobiles. Hence, road distance is a better indicator of travel cost than air distance. The use of road mileages increases distance and changes the shapes of distance zones.

Introduction

Concluding a paper on predicting recreation benefits in a way which has become standard among recreation economists, Edward Ullman and Donald Volk noted in 1962,

> To be truly effective ... [travel distance] models call for far better data on recreation than are now generally available. Attendance figures need to be improved, reliable, detailed origin data for small areas need to be collected,

> and some before and after surveys of behavior made in order to have firmer bases for prediction. The increasing importance of recreation calls for some hard facts on which to base hard analysis.[1]

Since 1962, there has been a great effort on the part of the Army Corps of Engineers to conduct systematic visitation surveys at each reservoir being used for recreation. Nevertheless, the conclusions drawn from the new "hard data" have tended to resemble those drawn from presumably softer stuff: Essentially, the Corps is advised to locate reservoirs intended for recreational use near urban centers with large populations which otherwise lack sufficient opportunities for water-based recreation.

While there has been a spate of data collection, the assumptions underlying Hotelling's original travel-distance method for estimating recreational benefits have largely gone unchallenged. Marion Clawson's subsequent suggestion that crowding may have a negative effect on the benefits of a recreational experience is but one of many issues which impinge on the validity and usefulness of Hotelling's model.

From mid-June until mid-July, 1967, the author visited three Corps of Engineers reservoirs: Youghiogheny Reservoir in Pennsylvania and Maryland, Allegany Reservoir in Pennsylvania and New York, and Lake Oahe in the Dakotas. At the reservoirs, some visitation data were collected and recreators were interviewed in an attempt to identify some of the specific behaviors underlying the travel-distance concept.

---

[1]Edward L. Ullman and Donald J. Volk, "An Operational Method for Predicting Reservoir Attendance and Benefits: Implications of a Location Approach to Water Recreation," Papers of the Michigan Academy of Science, Arts, and Letters, XLVII (1962), 484.

The major premise of the studies, suggested by Marion Clawson in 1959,[1] was that significant numbers of multiple-purpose recreational travellers would be found. A method for measuring their travel costs was proposed by George Tolley and V. Steve Hastings in 1965, when they hypothesized that all such travellers had willingly travelled some definite distance out of their way to include the facility in their itineraries.[2] The term adopted to describe such travellers was "diverters" although the phenomenon is analytically distinct from the simple choice among alternatives to which the term was earlier applied by Ullman and Volk.

The studies revealed that diversion phenomena[3] are progressively more important from a reservoir such as Youghiogheny to one such as Allegany to the large Oahe Reservoir. Possible reasons for this variation are discussed below. Meanwhile, the phenomena are pronounced in at least some cases, and they have yet to be accounted for in travel-demand estimates.

A major aim of the studies was to develop a means of interviewing recreators which would be uncomplicated and yet provide an accurate

---

[1] Marion Clawson, "Methods of Measuring the Demand for and Value of Outdoor Recreation." <u>The Resources for the Future Reprint Series</u>, X(February, 1959).

[2] G. S. Tolley and V. S. Hastings, "Framework for Investigating Benefits Lacking Good Market Indicators." Kerr Reservoir Conference, February 26-27, 1965 (unpublished). To paraphrase their essay, this distance is the <u>minimum</u> attributable to a particular site visit.

[3] More than one type of diversion was observed in the course of the studies.

picture of their itineraries, as these included a reservoir and as they would have been without the reservoir. The following set of questions was found to usefully serve this aim:

1. Where is your home? City ________________ State ________

2. Was coming here the only purpose of your trip? Yes ___ No ___

IF NO

a. Where did you stop last before coming here?

b. What will be your next stop after leaving here?

c. What route would you have taken if the reservoir were not here?

The purpose of the present essay is to illustrate how information gathered in such a way would be used to estimate modified travel-demand functions, and to suggest some of the implications of using this approach.

The interviews were conducted in the informal atmosphere of campgrounds and look-out points. Thus, although the studies were aimed at a specific objective, it was possible to use the opportunity for relaxed conversations with visitors to gain some further insights into the nature of various recreation experiences. As a result, the present essay also includes suggestions of several other issues toward which further research might be directed.

Section 1. Visitation at Oahe Reservoir: A Case Study

Part A. The 1965 Visitation Survey. Surveys of visitation at Oahe Reservoir were conducted by the Corps of Engineers in 1963, 1964, and 1965, following the standard procedure recently applied at most

Corps reservoirs across the country. The 1965 survey has been selected for consideration since it was most recent and benefitted from two years of experience by administering personnel.

The survey was in two parts: total traffic was recorded by counting machines at access points throughout the year; and interviews were conducted on representative weekdays and weekend days during each visitation season. The interviews were used as the basis for computing factors by which traffic counts could be converted into particular kinds and volumes of visitation.

The travel-distance component of the interviews sought to determine the proportions of visitors originating in one of four distance zones: 0-25 miles, 26-50 miles, 51-75 miles, and 76-100 miles. The zones were calculated on the basis of straight-line distances from the periphery of the reservoir. Subsequently, estimates were made of the population residing in each zone. The results of these calculations appear in a standard Corps survey summary for 1965.

The summary thus provides the information necessary for plotting rates of visitation from each of the four zones. In addition, the number of visits made by people from beyond 100 miles may be determined by subtraction from the total number of visits made during the year. In this case, since there are Canadian visitors at Oahe, the base population for visitors from over 100 miles includes the remaining population of the U.S. and Canada. Therefore, even though more than half the visitors to Oahe in 1965 came from over 100 miles, their rate of visitation was very small.

Figure 7 shows rates of visitation to Oahe in 1965 as recorded in the survey summary. It may be noted that if the bars for each distance zone are connected, the result is not a very satisfactory curve, since it appears to have a reversed distance-elasticity[1] between 25 and 100 miles.

Step 1. Effect of excluding rural residents. The area around Oahe is not only sparsely populated, but it is also primarily occupied by farmers and ranchers. One of the possibilities suggested by this summer's studies was that every recreator is an urbanite. In interviews with over 50 parties numbering approximately 200 members, not a single farmer was discovered, and the smallest town from which anyone came had at least 500 inhabitants. Therefore, it may be legitimate to consider only the urban population in each distance zone as the base for determining visitation rates. The use of this procedure implies that a smaller absolute number of people is actually competing for recreational opportunities than would otherwise be assumed (see Table 2); hence, per capita demand is greater. More importantly, the relative population of each zone is adjusted according to the degree of urbanization within the zone. Other variables normally used to predict reservoir attendence, such as income and employment status, do not need to be considered insofar as they vary directly with urbanization. On the other hand, no Negro or Indian visitors were observed at any of the three reservoirs, and different access points appeared to attract people with different social backgrounds. Thus, although there may be only one basic socio-economic

---

[1] In this paper, "higher elasticities" will mean higher negative elasticities.

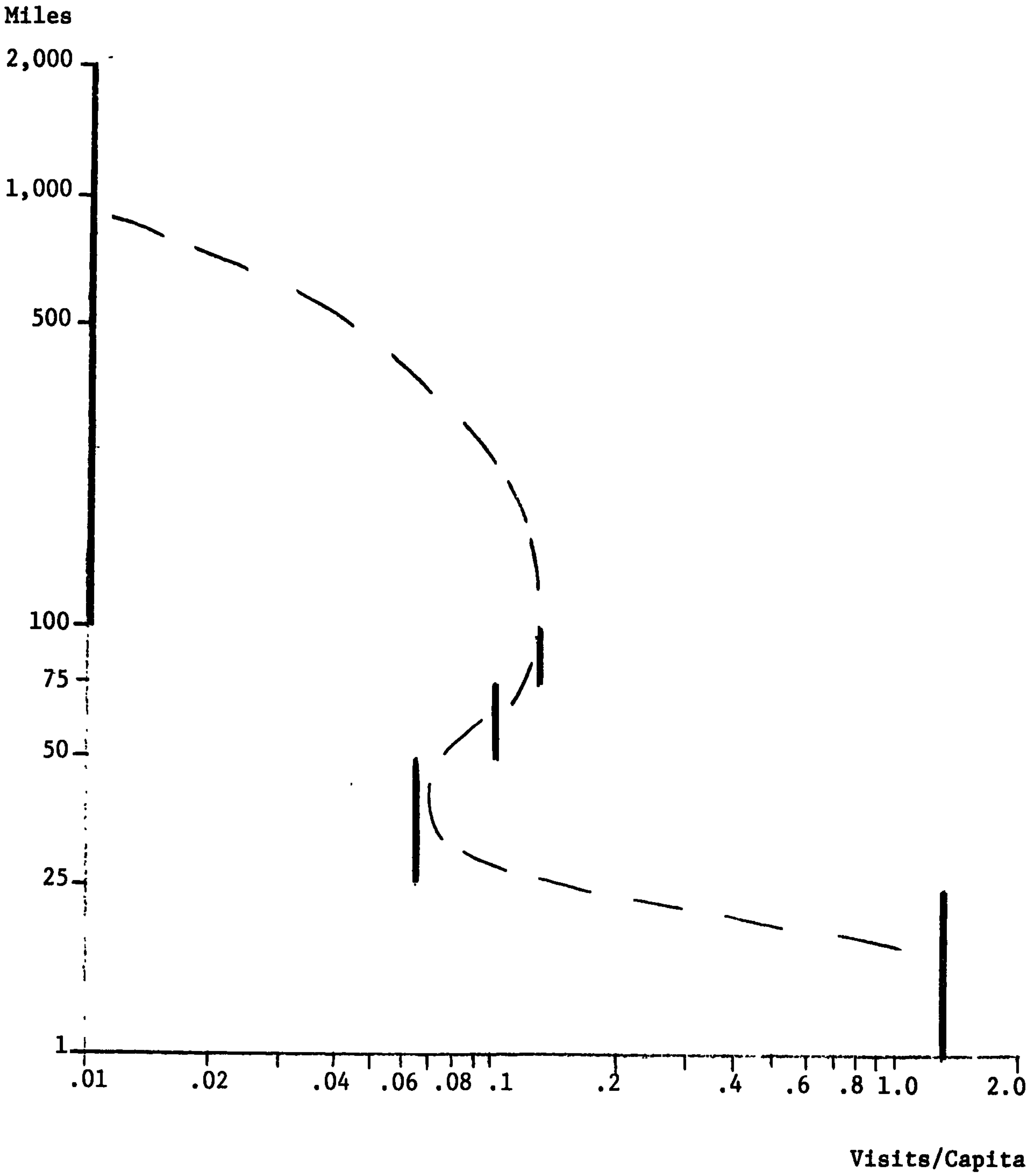


Figure 7 - 1965 Corps survey--man-visits/total population

Table 2. 1965 urban populations near Oahe

Zone (miles)	Total Population[a]	Urban Population[b]
0 - 25	96,000	71,336
26 - 50	47,000	16,270
51 - 75	80,000	16,304
76 - 100	151,000	82,640

[a]Source: Corps of Engineers - Civil Works. Missouri River Division. Omaha District. Oahe Project. "Recreation - Survey Summary," 1965.

[b]Computed by author, using 1965 road atlas with isopleth overlay and estimated 1965 Census of Population figures. An attempt was made to enumerate all towns classed as urban within 100 air miles of Oahe. For more distant zones with much larger populations, classification of cities by size might help simplify the enumeration procedure, or a random sampling method might be used.

variable in gross reservoir attendance, nevertheless, it appears that there are new insights to be gained from more refined field studies of the sociology of recreation.

The use of urban base populations to calculate per capita visitation at Oahe results in a somewhat more desirable scatter of distance bars, as illustrated in Figure 8.[1] In this and subsequent illustrations, a straight-line has been fitted by sight to the bars, which are plotted against logarithmic scales both of distance and of visits per capita.

Step 2. Effect of using road distances. Another possibility which suggested itself during the field studies was the use of road mileages to estimate travel distance. Virtually all reservoir users travel by car, and in any case it is only on traffic counts that visitation surveys are ultimately based. The use of road mileages is more realistic from a number of standpoints: It takes account of physical barriers

[1]Is it possible that the presence of large Indian reservations around Oahe helps account for the exceptionally low rate of urban visitation between 25 and 50 miles?

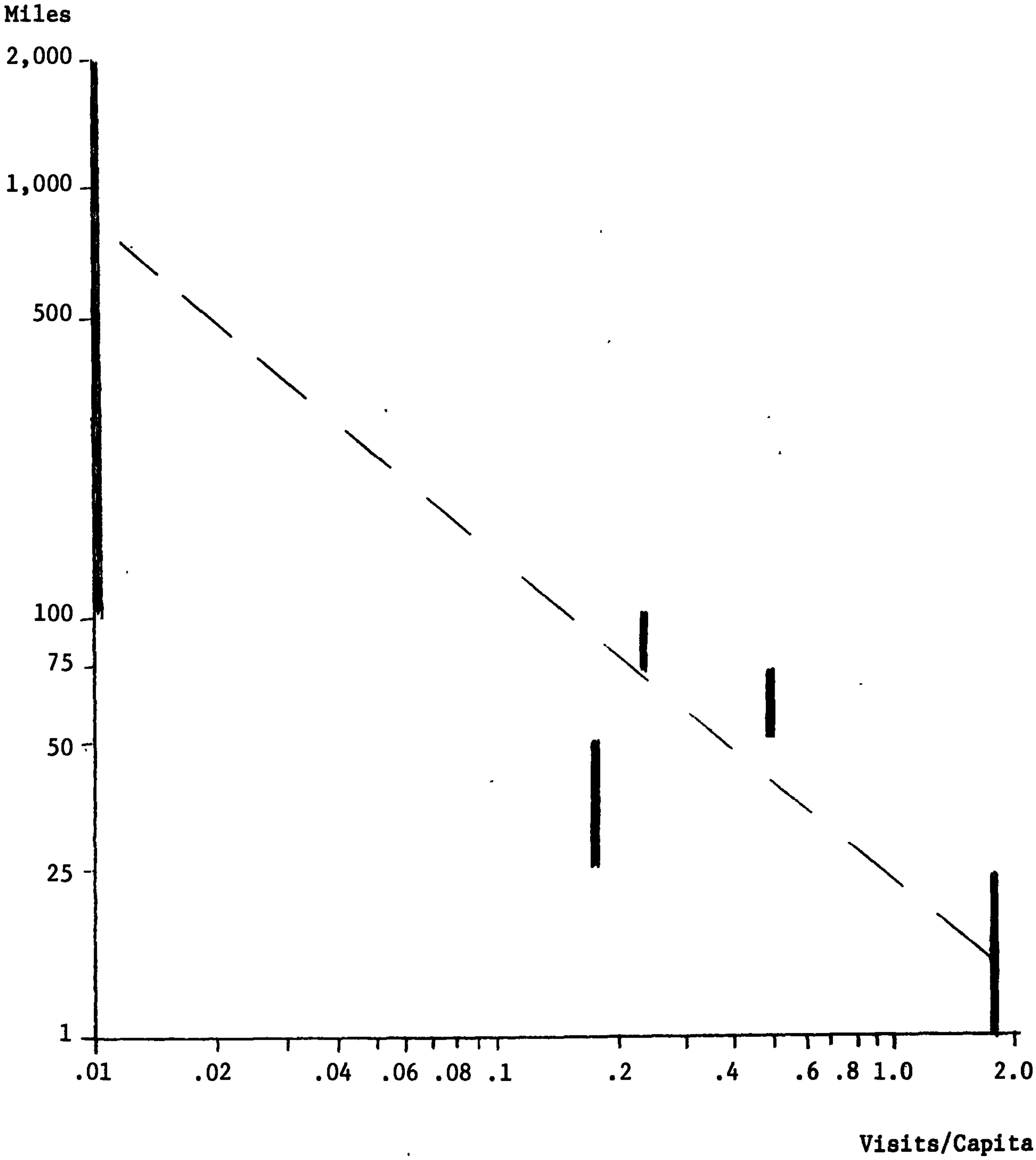


Figure 8 - 1965 Corps survey--man-visits/urban population

such as mountains and large water bodies which seriously affect the accessibility of a given recreational resource; it accounts for reduced visitation rates in areas with underdeveloped road networks; and it permits actual measurement of distances to particular access points along the reservoir, an important consideration at Oahe, where much of its 250-mile length is inaccessible by road. In addition, road distances are simply longer than air distances (see Table 3), and the difference increases as one travels farther in some directions from a reservoir.

Figure 9 shows how the use of road distances as the basis for delineating resident population zones further reduces the spread of distance bars around the hypothetical demand curve. This illustration is only half complete, however, since data were not readily available for classifying the number of visits from each zone according to the same road distance measure. In the remainder of the discussion, therefore, only air distances will be used.

Table 3. Air and road distances of selected communities from Oahe[a]

Community	Air Miles from Margin	Road Miles from Access Point
Aberdeen, South Dakota	85	100
Bismarck, North Dakota	6	11
Grand Forks, North Dakota	189	266
Minot, North Dakota	105	125
Rapid City, South Dakota	104	169
Sioux City, Iowa	238	261

[a]Computed by author.

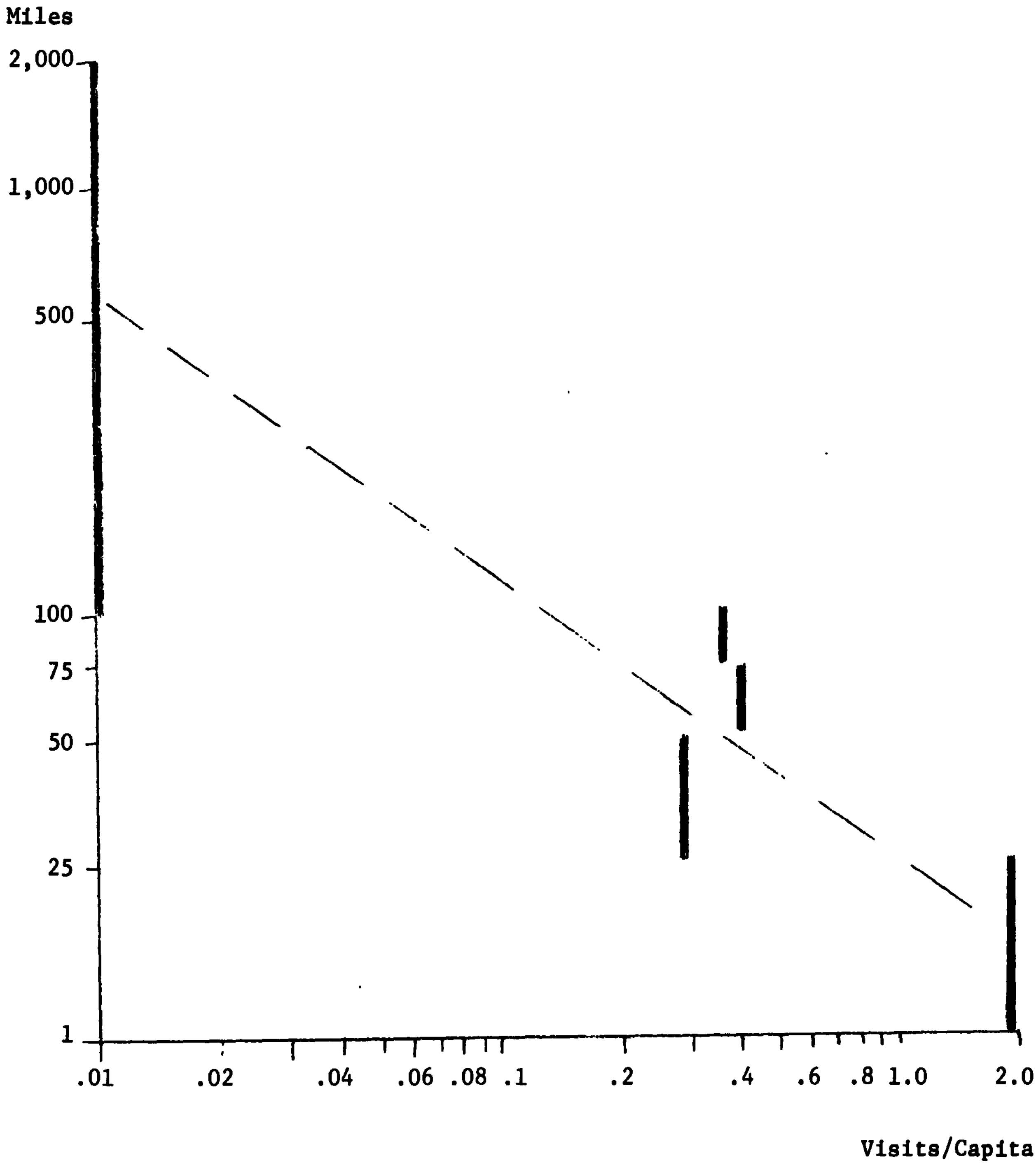
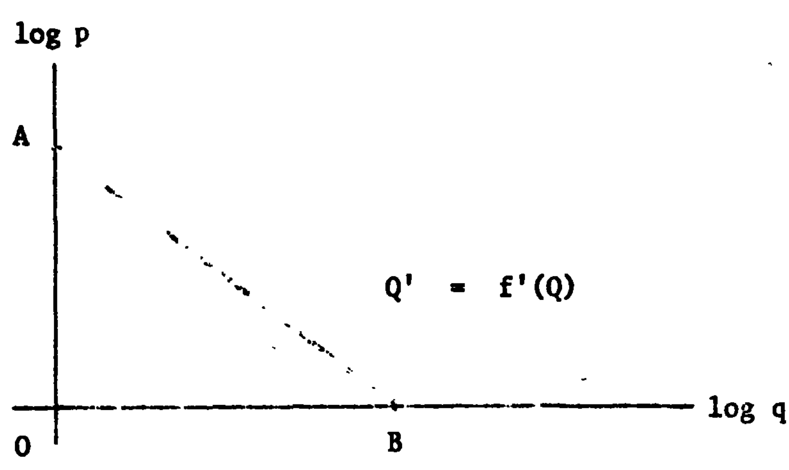


Figure 9 - 1965 Corps survey--man-visits/urban population, by road distance zone

Step 3. The distribution of long-distance visits. It has been mentioned that more than half the visitors to Oahe in 1965 came from more than 100 miles away. Hence, the use of distance zones of up to 100 miles was somewhat inappropriate in this case, and data on visitation rates for zones beyond 100 miles should be revealing. Again, these data exist but were not readily available during the time allowed for this particular study. In lieu of the original data, estimates were made of visitation rates from longer distance zones on the basis of visitation to one developed campground during the five-day period of the Oahe study. During this period, there were fairly high rates of visitation from air distances up to 400 miles from the reservoir, and the maximum distance of origin of a visitor was 1500 miles. The application of these estimates to the 1965 Corps data is illustrated in Figure 10. The elasticity of the demand schedule of Figure 10 may be read from the plotted curve, as follows:



where p = distance in miles and q = visits per capita.

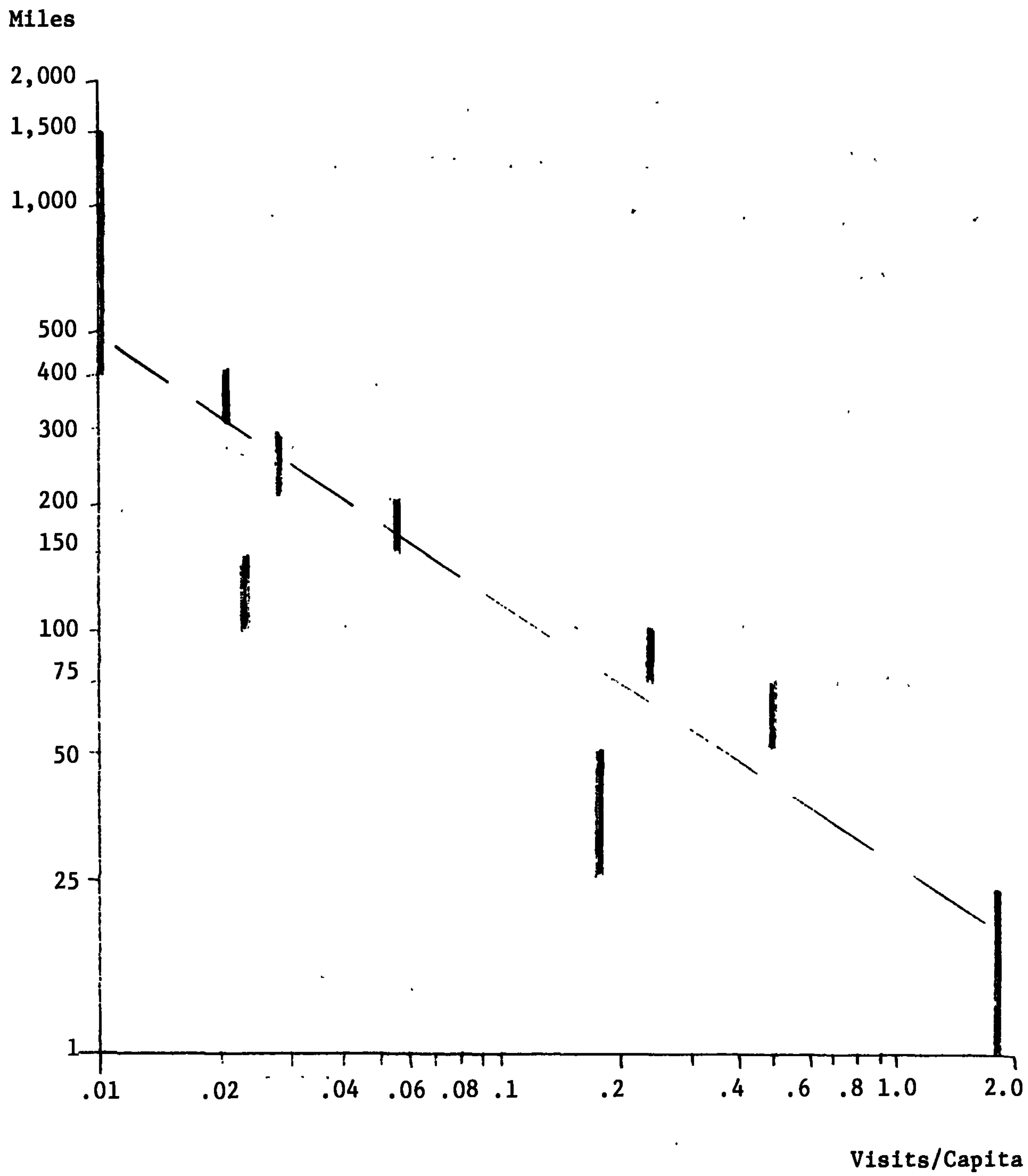


Figure 10. 1965 Corps survey--man-visits/urban population, long-distance visitors apportioned according to campground sign-ins

The distance-elasticity of demand is OB/OA.[1] In Figure 10, it would appear to be about 1.5.

Part B. The 5-day Oahe study. A study was made at the outflow area campground at Oahe Reservoir from July 7 through July 11, 1967. The study, modeled after the conventional Corps survey procedure, consisted of two parts: recording of data on origins and numbers in party obtained at a sign-in booth at the entrance to the campground, and interviews with every third party staying in the campground on a weekend day, Saturday, July 8, and on a weekday, Monday, July 10. The sign-ins were used to estimate rates of visitation by distance zone, as illustrated in Figure 11. The campground, with a capacity of 60 parties, was quite heavily visited during this period, and the distance bars show a fairly consistent pattern of demand, except for visits under 25 miles. This discrepancy is possibly explained by an imperfect market situation: an entrance fee was charged for the campground studied, while a nearby picnic area, familiar to local residents, was available for camping free of charge. The elasticity of the curve in Figure 11 appears to be about 1.4, somewhat less than that for the entire 1965 season. This might reflect either a change in demand or the saturation of substitutes during the height of the tourist season, when the interviews were taken.

Part C. Types of diversion. An interview of the kind presented in the Introduction was used in the campground, as the basis for estimating the proportion of multiple-purpose travellers for each distance

---

[1]The general form of the equation for function Q' (above) is log q = $\eta \cdot \log p + \log b$; where η = tan /OAB = slope OB/OA, and log b = B. The equation of the integral of function Q', or function Q, turns out to be $q = b \cdot p^{\eta}$. Hence, by definition, η = OB/OA is the elasticity of Q.

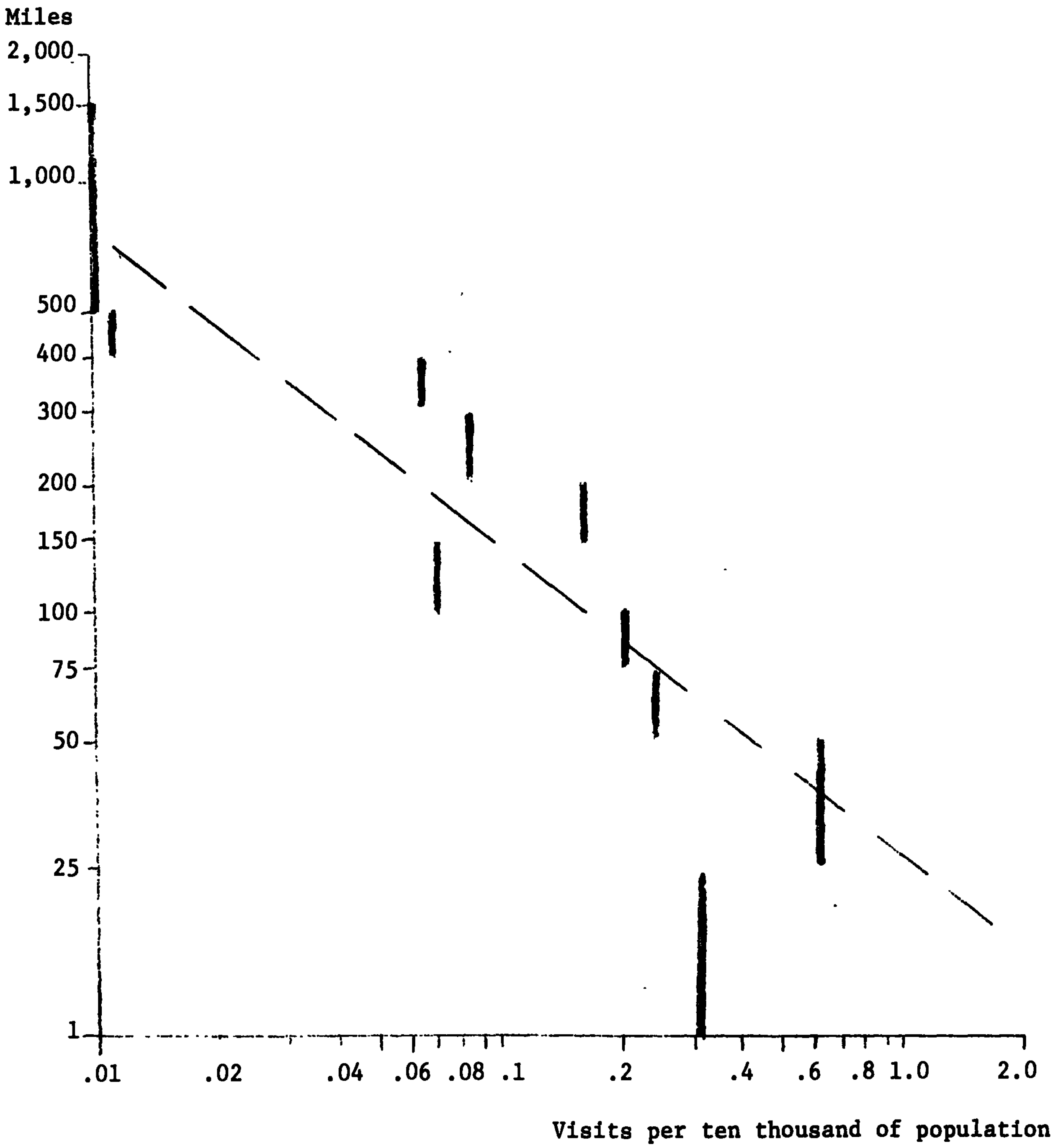


Figure 11. Campground sign-ins--man-visits/urban population

zone. For distances up to 200 miles, there were no diverters. This finding would seem to support the use of simple origin data to calculate travel costs to reservoirs which have no significant visitation from over 200 miles. Some diversion occurred between 200 and 400 miles, and above 400 miles everyone was taking a multiple-purpose trip. Such diverters incurred some proportion of the travel cost of their entire trip to visit the reservoir; this proportion was always lower than the distance of a single-purpose round trip to the reservoir from home, as illustrated in Table 4. The ways in which they might be reclassified are discussed below. The first step is to remove them from the origin-type demand schedule for the campground as illustrated in Figure 12. The curve in Figure 12 has an increased distance-elasticity which would appear to be about 1.7.

Two distinct types of diversion were observed among the campers at Oahe. Many travelers found the campground a convenient night's stopover along the route they would have travelled in any case; sight-seers who spend even less than a night's stay at the reservoir might also fit into this category; the phenomenon which they represent could be termed "route diversion," implying a brief stopover at the reservoir while en route elsewhere. Other recreators fit more closely into the conventional concept of "multiple-purpose travellers." Such travellers had been attracted by the relatively high level of recreational resources in and around Oahe, and generally made more of a purposeful effort to include the reservoir in their travel itineraries. It may be that "purposefulness" in this sense is reflected in the total time spent visiting a reservoir, either in absolute terms or as a proportion of

Table 4. Sample diversion travel distances of Oahe campers

Origin	Distance Zone (miles)	Diversion Route	Route No.[a]	Diversion Distance (miles)
Colorado Springs, Colorado	401-500	Black Hills, Colorado Springs	M.P.	149
Indianola, Iowa	401-500	Black Hills, Minnesota	16	39
Menomee Falls, Wisconsin	601-700	Black Hills, Wisconsin	14	3
Omaha, Nebraska	301-400	Black Hills, Omaha	14	3
Billings, Montana	301-400	Eastern North Dakota-Montana	M.P.	77.5
Minneapolis, Minnesota	301-400	Black Hills, Northern Minnesota	M.P.	6
Milhill, New Jersey	1201-1400	Black Hills, Fort Randall	M.P.	57
Los Angeles, California	1001-1200	Chamberlin, South Dakota-Los Angeles	16	39
Mitchell, Missouri	501-600	Gavins Point, Black Hills	18	58

[a]Route No. shown if "route diverter." "Multiple-purpose travellers" indicated by "M.P." See discussion on previous page.

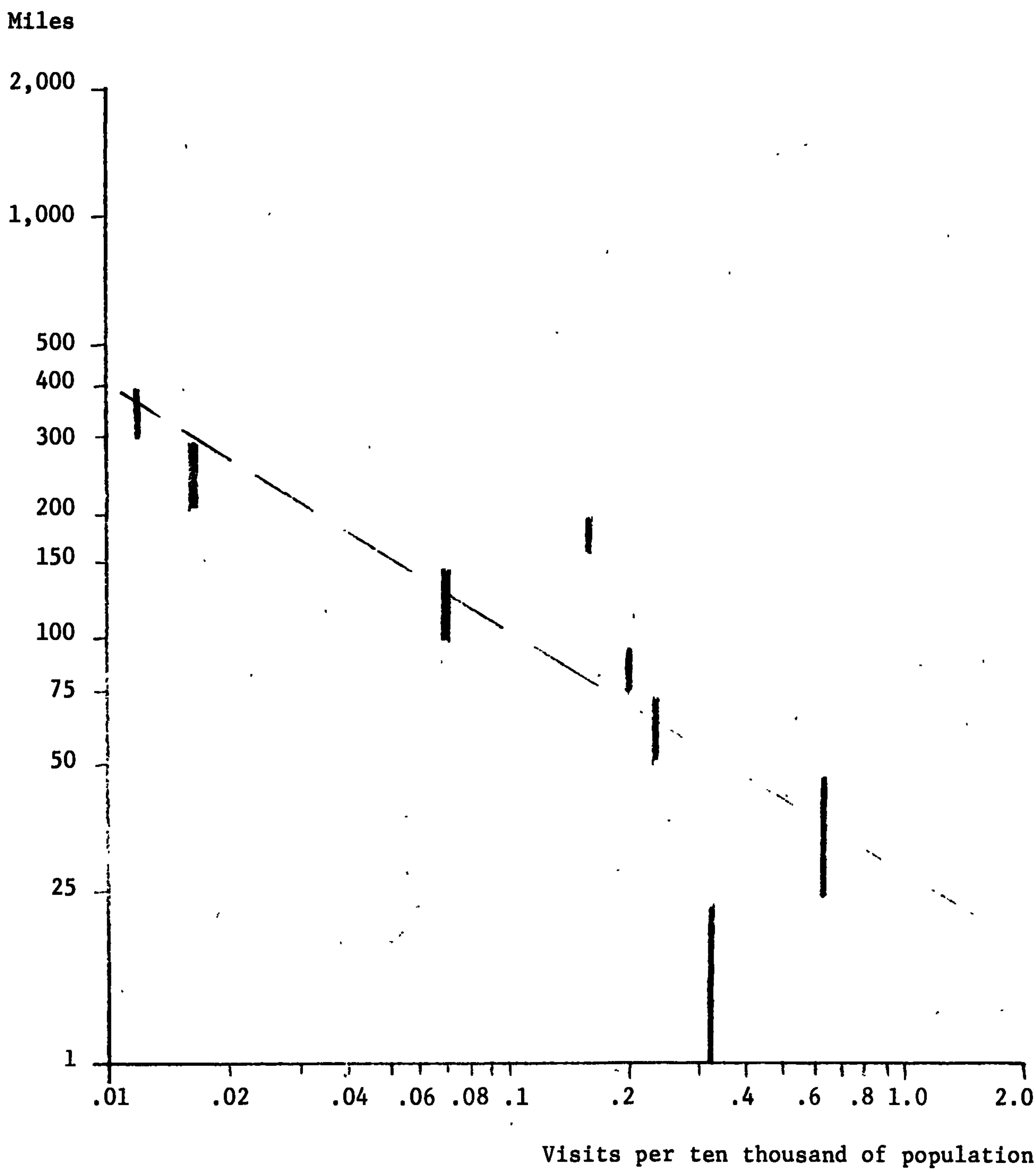
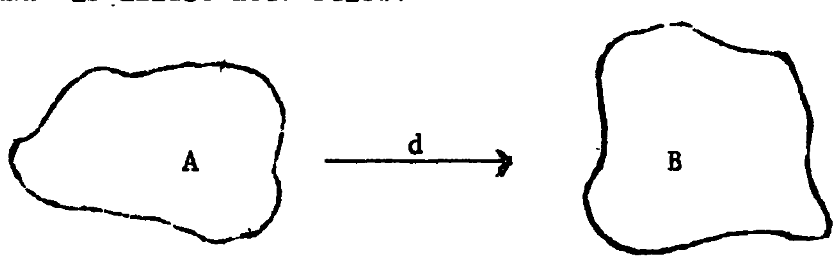


Figure 12. Campground sign-ins--diverters subtracted according to interview sample

the vacation time available to each visitor. However, further subdivisions of multiple-purpose travellers on such a basis were not attempted during this study.

To reclassify each type of diverter requires not only a more precise estimate of travel costs, as determined by the suggested interview method, but also a decision regarding the base populations from which such recreators may appropriately be considered to have been drawn. The reclassification problem becomes quite unwieldy when one attempts to resolve it within the conventional model based on residential population potential. A simple extension of this model, however, may eventually provide a suitable framework for a much broader consideration of issues of recreational demand, including the diversion problem.

A suggested framework for broadening the conventional view of recreation demand is illustrated below:



"A" and "B" are geographic regions, "d" is a highway. It is assumed that region B has a greater recreation resource relative to its population than region A has. This imbalance in resources results in a net seasonal migration from A to B over route d during the summer months, when recreation resources are in greater demand relative to other kinds of resources. Depending on a number of factors of which temperature and sunshine are but two, the rate of migration along route d increases to a climax in mid-summer, during which a significant proportion of the population from region A has relocated temporarily in region B.

Such a framework has direct application to the problem of reclassifying recreational diverters. "Route diverters" are highway travellers, and the base population from which they are drawn is the transient recreational traffic along the route from which they diverted. "Multiple-purpose travellers" belong to the population which has already shifted from region A to region B; they, together with the single-purpose visitors from region B, belong to the adjusted summer base population of the region. Alternatively, they could be considered to constitute a separate base population of people "on the move."

Step 1. Route diversion. The route diverters who stopped at the Oahe campground were travelling east and west along routes U.S. 14, 16, and 18, en route to or from the Black Hills. Comparing various possible means of reaching the Black Hills via the reservoir from any of these routes, the writer estimated "diversion distances" for each one. For comparison with other travel measures which reckon one-way distance from residence, these figures were calculated as one-half the total increment to travel distance necessary to include the reservoir in an itinerary.

Transient recreational populations were estimated from 1965 traffic counts by assuming a base utilitarian traffic represented in the months of May and October and calculating the increment to this base traffic during the month of July. The increments were multiplied by 1/6 to find traffic during a five-day period, and the traffic was multiplied by the load factor used to estimate man-visits from vehicle-visits in the Corps survey (about 3.18 in the 1965 survey).

Actual traffic surveys taken at points where diversion was expected would have been desirable, but were not feasible. Such surveys would ideally include brief interviews, although for rough estimates one could count the number of people travelling in out-of-state cars.

Only those interviews at the campground which clearly fit into the Black Hills-route diversion pattern were used to estimate the route diversion demand schedule among sign-in campers illustrated in Figure 13.[1] Although the estimate is made from very little data, it would seem that the recreational demand for route diverters is relatively distance-inelastic, apparently having an elasticity of about 0.7.

Step 2. Multiple-purpose trips. When asked what route they would have followed if the reservoir were not there, travellers who "would have followed a more northern route," who "were on the way back to Denver but decided to meet their relatives from Iowa here," who "came out by route 10, but wanted to go back a different way," and so on, were classified as multiple-purpose travellers. Their travel costs were estimated as one-half the difference between the miles they travelled and the minimum number of miles they would have had to travel to carry out their alternative itineraries. Their "diversion routes" were thus scattered randomly throughout the vicinity of the reservoir rather than adhering to the direct route diversion pattern.

---

[1] The per capita diversion from each route shown in Figure 7 is the number interviewed who diverted from each route divided by the total number interviewed, multiplied by the number of sign-ins for a five-day period and divided by the estimated tourist traffic for the same five-day period.

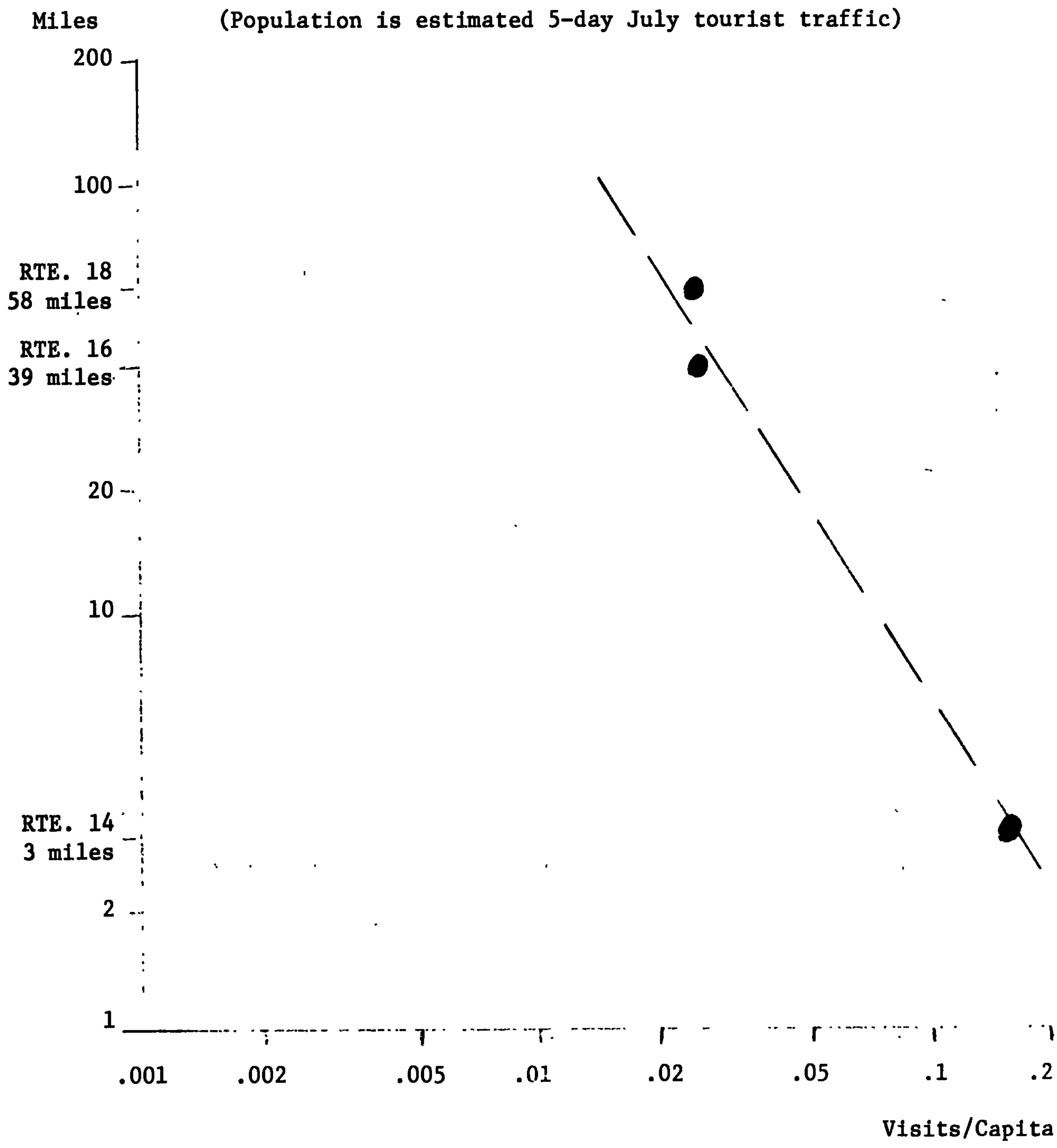


Figure 13. Campground sign-ins--route diverters apportioned according to interview sample

Data were not readily available to permit proper adjustments of base populations around the reservoir within the framework suggested above. Indeed, the identification of such population shifts is the major unsolved problem posed by the application of a diversion-demand scheme.

Part D. Effects of proposed method on the 1965 Corps survey schedule. Although the method for dealing with diversion phenomena outlined above may be applied to any reservoir or other recreational facility, the particular results of the Oahe study apply only to visitation at a single campground during one five-day period in July, 1967. It is nevertheless of interest to consider what effects the method might have had if it were applied during the 1965 Corps survey.

Data from the campground interviews showed that higher percentages of visitors from longer distance zones than of visitors from shorter zones were diverters. These diverters were subtracted from total visitors to arrive at the hypothetical single-purpose demand schedule shown in Figure 8. This compares with Figure 4 in which all visits are included. To the extent that diversion visits vary directly with highway tourist traffic, a route diversion schedule for the entire 1965 season, if calculated on the basis of data collected for this study, would exactly resemble Figure 13.[1]

[1]To estimate route diversion demand in 1965, more information would be required concerning tourist traffic flows near the entire length of the reservoir. Such information would include the total summer tourist traffic along U.S. routes 10, 12, and 212 as well as along 14, 16, and 18 which were obtained for the present study. It would also include knowledge of the ratio of tourist traffic during the interview period to total summer tourist traffic, so that allowance

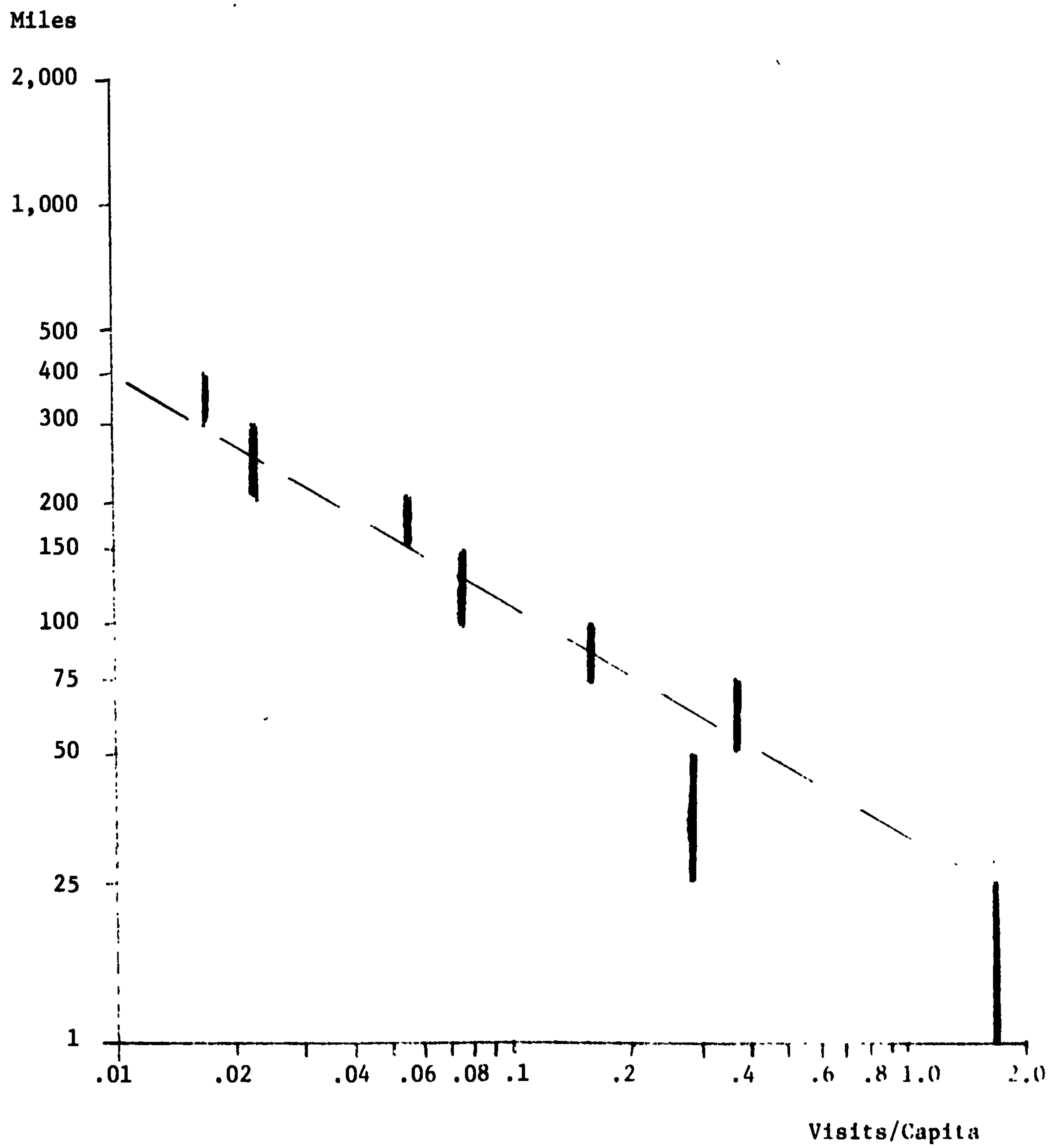


Figure 14. 1965 Corps survey--diverters subtracted according to interview sample and sign-ins

It is interesting to note that both Figures 13 and 14 show a markedly reduced scatter of distance bars and points around the hypothetical demand curves, and appear quite reasonable when compared with other estimates of recreational demand. When compared with Figures 8, 9, and 10, Figure 14 shows a demand function which is more distance-elastic. This follows from the subtraction of "diverters," who are primarily long-distance travellers. On the other hand, the route diversion function in Figure 13 is less distance-elastic than the functions including single-purpose visitors. Relative inelasticity may be a characteristic of route diversion demand for Corps reservoirs, at least within a distance range where travel costs are competitive with alternatives such as motel costs.

---

might be made for variations in the base population in estimating per capita visitation.

Finally, the problem of the variability of route diversion rates would have to be met. To illustrate this problem, one might imagine a reservoir with only one nearby tourist artery. On a given interview day, ther were ten route diverters, and another ninety visitors were single-purpose travellers. On that day, there were only fifty tourists travelling on the highway, but during the whole season there would be 100,000. Now, suppose during the whole season there are also 100,000 visitors to the reservoir. How many of these are route diverters? If one used the ratio of 10 diverters for each 100 visitors during the interview day, one would conclude that there were 10,000 diversion visits during the whole season. If, on the other hand, one used the survey to estimate that ten out of every fifty tourists on the highway would divert to visit the reservoir, one would conclude that there were 20,000 diversion visits during the whole season.

How can one decide whether there were 10,000 or 20,000 diversion visits in this hypothetical year? It would be important to select a number of interview days at various times during the summer, to check whether the number of diversion visits varies more closely with highway tourist traffic flows on the one hand, or with total visits to the reservoir on the other. The conclusion noted in the text assumes that diversion visits are strictly a function of highway tourist tràffic and, of course, of distance. Thus, in this hypothetical year, one would conclude that there were 20,000 diversion visits and only the remaining 80,000 visits were single-purpose ones.

Section 2. Visitation at Youghiogheny and Allegany Reservoirs

Roughly the same procedure as that described for Oahe Reservoir was applied during studies of visitation at Youghiogheny and Allegany reservoirs. The results were strikingly different, however.

At both of these reservoirs, there were no sign-in booths, because there was free public access to all recreational areas. Therefore, estimates of the origin of visitors from various distances had to be made on the basis of license plate counts. For Youghiogheny, visitors from Pennsylvania, Maryland, or West Virginia were considered in-state and all others out-of-state; for Allegany, non-New York or Pennsylvania license plates were taken to be out-of-state. In each case there would presumably be a trade-off between long-distance visitors from in-state (for example, Philadelphia visitors would count as short-distance travellers even though they would travel over 300 miles to reach either of the reservoirs), and relatively short-distance ones from out-of-state (for example, Cleveland visitors had to travel only 160 miles to reach Allegany). This assumption was validated at Youghiogheny, although at Allegany no interviews were taken among in-state visitors to ascertain their true travel distances.

The Youghiogheny study covered a period of one full week in late June, 1967. During this time, 817 license plates were counted and 30 parties were interviewed. The vast majority of visitors were found to be from within 100 miles, with Pittsburgh being the fartherest point from which such local traffic originated. A scant 2.8 percent of the cars at Youghiogheny bore out-of-state license plates, in other words a total of about 20 cars. Four out-of-state parties were interviewed.

Of these, two were visiting relatives or friends nearby who normally came to visit the reservoir; the other two were diverters travelling along from one campsite to the next on extended trips. Probably, the parties visiting relatives would give the relatives' address as a place of origin in a standard interview, since they were all travelling together and the heads of parties were the ones familiar with the reservoir. Therefore, it might be concluded that little harm would result from using standard Corps survey information in assessing the travel demand for Youghiogheny Reservoir.

In such an area, there would seem to be a great deal more gained from the technical refinements of measurement recommended earlier, however. There, on the edge of Appalachia, rough terrain and poor roads would double road distance compared with airline distance from many surrounding points, and indeed most visitors were found to proceed from towns along the main routes to Pittsburgh straight down to the reservoir, while very little visitation originated in areas off to the side of these routes. Also, the use of urban base populations would be particularly appropriate: Virtually all of those interviewed held urban jobs, such as truck diving and steel working, even though the area is characteristically rural.

The Allegany study included only one day of car counts and interviews at a point overlooking the new Kinzua Dam. Little more could be accomplished since the recreational facilities of the reservoir are still under construction. It is interesting, however, that visitation at "Big Bend Overlook" was found to be intermediate in terms of diversion between Youghiogheny and Oahe. One should bear in mind that diversion phenomena would be expected to increase during the period from late June until late July in which Youghiogheny, Allegany, and Oahe were studied in that order, as extended summer vacations were increasingly being taken, permitting more long-distance recreational travel. It is possible, therefore, that the differences noted in diversion rates at the three reservoirs are not as radical for the summer as a whole as they appeared within the constraints of the study.

The interviews at Allegany were taken on Saturday, June 24, during four one-hour periods of the day when a total of 279 cars passed through the overlook station. On the basis of both static and frequency counts, out-of-state vehicles were found to represent just about 20 percent of the total number. Of these, 15 parties were interviewed, that is, about 30 percent of the out-of-state traffic. Every one of those interviewed were diverters, except one party which still had Oklahoma license plates but had recently moved to New York State. None were "route diverters," probably because of the absence of large, efficient through highways in the vicinity of the southern end of the reservoir. All had been attracted to the general area, and were on multiple-purpose trips.

The Allegany Reservoir is located on the edge of Allegheny National Forest, a large wilderness area of Pennsylvania and New York. The fact that the forest is purposely kept "wild" helps account for the maze of travel routes found among visitors, while its existence is the most powerful force attracting them to within visiting range of the reservoir. Here, a somewhat different type of diversion was observed.

Most long-distance visitors had either previously lived in the area and were returning to visit the old homestead, or they had come to live and work in the area. A couple from Texas were visiting relatives in Pittsburgh; a party of sisters and their husbands from California and Florida were rejoining their stay-at-home sister from nearby Bradford, Pennsylvania; two parties of students from across Canada were spending the summer as interns at the Warren County Hospital just 10 miles down the road; another party from Florida had been visiting the family homestead in Jamestown, New York; one from Kalamazoo, Michigan, was back home in Warren; another from Michigan was returning to Marienville, Pennsylvania; and a group from Binghamton, New York were on their way to see friends in Coutersport, Pennsylvania. Although this sort of phenomenon will not be susceptible to full-scale investigation at Allegany until the reservoir is completed and better data are available, it appears to indicate a definite link between recreational travel and broader patterns of geographic mobility. This link might be of great importance in anticipating recreation demand in areas characterized by high migration rates of a nonrecreational variety.

Conclusions

In sum, the brief field studies conducted at three Corps reservoirs permitted the identification of two distinct patterns of recreational travel in addition to the direct residence-resource pattern normally assumed in survey questionnaires. They are "route diversion" and "multiple-purpose travel."

A general framework of migrations resulting from resource imbalances may be a way of adjusting bases for analyzing visitation rates. If such a general framework were used, an estimate of the probable benefits of a new reservoir would first have to take account of the impact of the reservoir on permanent and seasonal migrations from and into the vicinity of the project.

Distinctions of different types of travel demand lead one to the conclusion that it is very important to know the rates of route diversion to be anticipated at a new reservoir. The schedule presented in Figure 14 for single-purpose visitation shows a greater distance-elasticity of demand than the conventional schedule illustrated in Figure 10. On the other hand, the sample route diversion schedule of Figure 13 shows a reduced elasticity of demand.

Of perhaps similar importance are the rates of visitation by multiple-purpose travellers to be anticipated. No conclusions may presently be drawn regarding the elasticity of such demand.

The possible attributes of types of diversion demand suggest the modification of criteria for the optimal location of reservoirs intended for recreational use. There would seem to be increased benefits from location close to urban centers compared with those discovered by conventional demand estimates. There is an undertermined value in location in regions of high tourism, as well as a need for estimating the impact of a new reservoir on rates of tourism in a region. Location near a major tourist artery would also appear to have a value. However, an inquiry should be made to identify alternative means of meeting such a demand. Campsites, for example, may be equally pleasant along a gurgling brook as along a reservoir margin.

Ideally, all recreational activities anticipated in a proposed reservoir project would be investigated separately with respect to the special characteristics of the demand for the activity as well as to the social characteristics of the base population which is to participate therein. For each such activity, and therefore for the project as a whole, the benefits of the project would be limited to the "alternative costs" of providing the same benefits without the reservoir.[1]

A few of the social factors affecting reservoir attendance were noted in the course of the studies. These relate not only to visitation rates themselves, as in the case of low visitation from Indian reservations, but also to types of activities engaged in, as in the case of emphasis on combined camping and water-skiing by the working people visiting Youghiogheny. Attention to such factors may provide further refinements of demand estimates as well as greater precision in the identification of optimum facilities required to best satisfy the demand.[2]

In general, the method outlined above is designed to permit more precise estimates of the travel demand for reservoir recreation, with important implications noted for the optimal location of recreational facilities. With this increased precision can come more realistic, flexible, and efficient proposals for future recreation developments.

---

[1]See U.S. Senate. Document No. S97. 87th Congress, 2nd Session, 1962. Policies, Standards and Procedures in the Formulation, Evaluation, and Review of Plans for Use and Development of Water and Related Land Resources. Washington: Government Printing Office, 1962, p. 8, VDS2.

[2]See, for example, Robert C. Lucas, "The Recreational Capacity of the Quetico-Superior Area." U. S. Department of Agriculture. Forest Service. Lake Forest Experiment Station: St. Paul, 1963. (unpublished).

EFFECTS OF WATER DEVELOPMENT ON LOCATION OF WATER-ORIENTED MANUFACTURING*

Introduction

Regional development and industrial location are determined by the availability of raw materials, proximity to markets and their relative importance for each industry. Water is considered an important factor of production for many industries. However, the cost of using water is relatively small in many industries and hence is given only little weight in the location decision process of these industries. This study makes a distinction between industries according to the importance of water as a production and location factor, and deals primarily with industries in which water is an important factor. The purpose of this distinction is to provide a better basis for analyzing the conditions in which water is restricting industrial development and to estimate the effects on employment in water-oriented industries.

Conceptual Discussion

Water-oriented manufacturing as defined for the purposes of this study includes those industries in which water has a significant part in the production cost structure and therefore plays an important role in the location decision. Important factors in the location of water-oriented manufacturing are water availability and water quality. Water is used as a raw material in the production process, as a means for cooling and as conveyer of raw materials and wastes. In studying the effects of water projects on location of industry and on regional growth, it is important to determine first the behavioral relationships between water, employment and production in water-oriented manufacturing.

---

* Shaul Ben-David, University of New Mexico, Albuquerque, New Mexico.

Supply and demand for water can be derived by methods common to other factors of production. But we have to note that water as a raw material has special characteristics. Water is usually not bought (unless received from municipal or private utilities sources). Most large plants pump water directly from a nearby stream or lake with no direct payment for the water. Optimum use of water is therefore determined by the overall cost of using water rather than the nominal price of water itself. The actual price (cost of using) water to the plant includes: (1) pumping and hauling from the source to the plant, (2) treatment of intake water to improve quality of water, (3) treatment of wastes to comply with downstream regulations, (4) deterioration of water-using equipment (corrosion, etc.) caused by low-quality water (brackish). The actual price a plant is paying for each unit of water will depend on factors affecting these costs.

It seems reasonable to assume constant or increasing marginal costs for water as there is movement along the output supply schedule. For water-oriented manufacturing the marginal cost for water (alone) would rise. Additional output will require higher costs of using water to produce a unit of product. The additional costs may result from additional costs of pumping, treatment, etc. Cost might also increase by changes in production practices in order to use less water per unit of product. These practices include recirculation and air cooling, possibly induced to avoid bearing even higher costs, for example, of more distant water sources or of waste treatment. The adoption of waste treatment practices, either as required by law or by public considerations is a factor causing costs to rise with increasing production.

The position of the supply function for water is obtained by total water availability in the area and by its quality. Quantity of water available and its quality, in a given area, for a new plant, is related to the length of streams and their flow in the area and to present uses by others. Most industrial and municipal water users have a low consumptive use, a large percentage of their intakes is returned to the streams (for example, water used for cooling, washing, etc.). Water can be reused by another user as soon as it is brought to the quality level required by him. The distance between the intake of a new plant and the outflow of an existing one will be affected by the following: (a) the waste load of the existing plant and the degree of waste treatment, (b) the new plants quality requirements and its planned intake treatment facilities, and (c) the flow in the stream. The effects of wastes disposed by one user on the quality of water intake for the downstream user is inversely related to the flow in the stream. A givenwaste load discharged into two streams which differ only in flow will affect quality more significantly in the stream with the lower flow.

Lower stream flows will require a longer distance between two users. Water availability in an area is therefore a function of length of streams and their flows. Water quality is highly affected by other users in the area and the types and quantities of wastes discharged by them. The cost to a water user depends on water quality, which is determined by upstream intensity of use. Higher flows reduce the effect of an upstream waste load by dilution and hence decrease costs to the downstream user.

The demand for water is a derived demand which is determined by the production function and the demand for the final product. The price an entrepreneur is willing to pay for the use of one unit of water will be as high as the value of marginal product of this water unit. We would expect a downward sloping demand curve. An industry will be willing to buy larger amounts of water when water prices are lower.

The quantity of water used and the cost per unit of water to the industry are determined by the supply and demand for water. We will consider four major cases:

(A) <u>Economic use</u> where the firm is using WU^a units of water at a cost P_w^a per unit. (Figure 15)

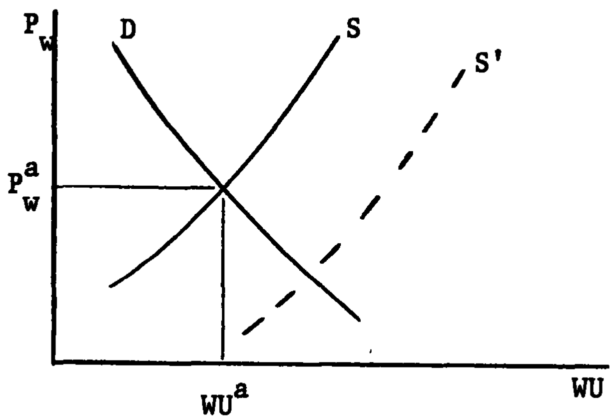


Figure 15 - Economic Use of Water

(B) <u>Free use</u> where the firm water demand (WU^b) can be met at no cost. (Figure 16)

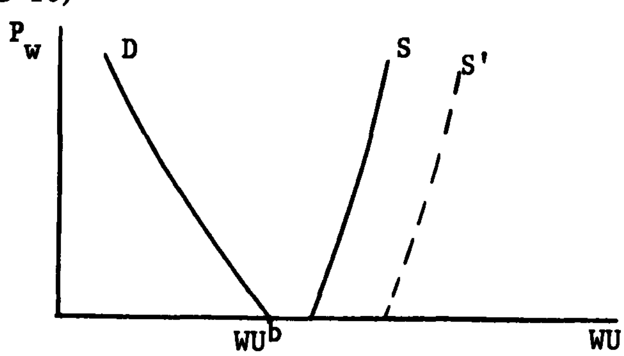


Figure 16 - Free Use of Water

(C) No use-water limiting when water is very limited and even a very small quantity of water will have a unit price that exceeds the value of its marginal product. (Figure 17)

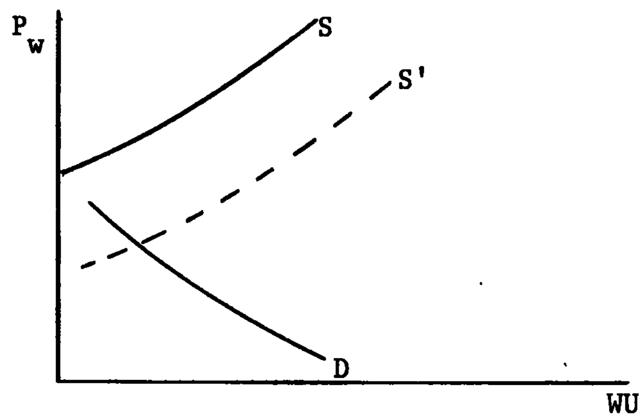


Figure 17 - No use-water limiting

(D) No use-other factors limiting when other production factors are very limited and the firm will not produce at any price of water. (Figure 18)

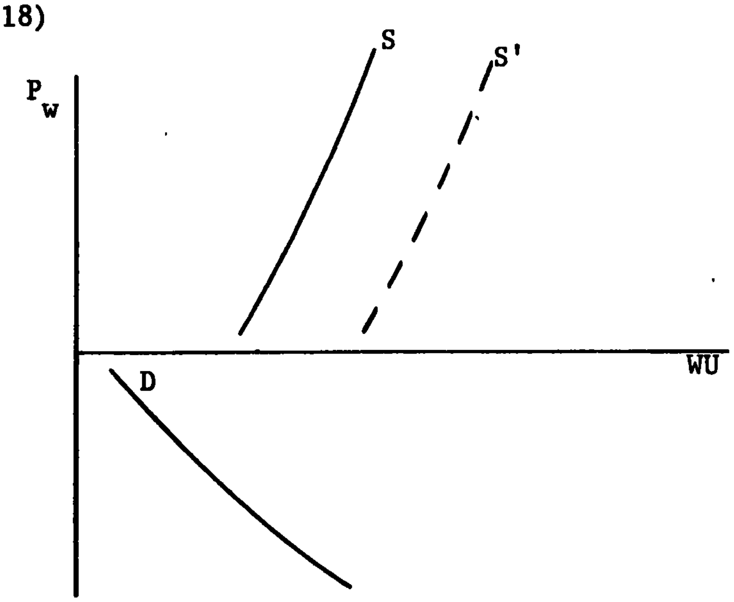


Figure 18 - No use-other factors limiting

The significance of these four cases is realized when we consider the possible effects of water projects on the location of production. A water project which augments water at low flow periods affects the position of the supply function. Larger flows increase water availability and therefore, cause a decrease in intensity of use if use by other manufacturing remains unchanged. Water quality is also affected by low flow augmentation through dilution of the given amount of wastes disposed in the stream. Both the decrease in intensity of using water which appears as a change in the supply function (S' in Figures 15-18) which is positioned now to the right of S. In the four cases discussed we can see that a water project will have no effect on production in Cases B and D, because water was not the limiting factor even before an addition caused by a water project.

A water project could affect production in Cases A and C. In Case A any shift of the supply function to the right will increase production as long as other factors are not limiting production. In Case C, where water is very limiting and no water-oriented manufacturing exists in the area, a water project could be effective in attracting industry only if the supply function with the project will be shifted to a position such as to have a positive intersection point with the demand function.

Employment and water inputs in a water-oriented manufacturing industry, may be viewed as determined simultaneously with maximum profits realized when values of marginal products equal to input prices. Let us consider the following production function:

(1) $$Q = f(W, E, R)$$

where the production (Q) is determined by inputs of labor (E), water (W) and other inputs (R). The demand for water as determined by its marginal product will be:

$$P_W^D = P_Q \frac{\partial Q}{\partial W} = g(P_Q, W, E, R) \tag{2}$$

where the cost of water (P_W^D) that a company is willing to pay is related to the price of the product and the level of all inputs.

The demand for labor (employment) will be:

$$P_E^D = P_Q \frac{\partial Q}{\partial E} = h(P_Q, W, E, R) \tag{3}$$

where the price which the company is willing to pay for labor (P_E^D) is related to the price of the product and the level of other inputs.

We would expect the following supply function for water:

$$P_W^S = f(W, WA, WI). \tag{4}$$

The cost of a unit of water (P_W^S) is a function of the quantity of water used by the plant (W), the total water availability in the area (WA) and the intensity of use by other water-oriented plants in the area (WI).

A supply function for labor is expected to be of the form:

$$P_E^S = f(TE, MW) \tag{5}$$

where the price of labor to the plant is related to the total labor availability in the area (TE) and manufacturing wages (MW). A larger labor force in a given area increases the availability of skilled workers and also indicates a higher level of economic activity.

Profits are maximized by equating marginal cost to marginal revenue for each factor of production (equations (6) and (7)).

$$P_W^D = P_W^S \tag{6}$$

$$P_E^D = P_E^S. \tag{7}$$

Substituting (2), (3), (4), and (5) into (6) and (7) gives a system of two simultaneous equations with two endogenous variables (W and E) and six exogenous variables

(8) $$g(P_Q, W, E, R) = f(W, WA, WI)$$

(9) $$h(P_Q, W, E, R) = f(TE, MW).$$

The reduced form equation (where the endogenous variables are expressed in terms of the exogenous variables) can be written

(10) $$W = F(P_Q, R, WA, WI, TE, MW)$$

(11) $$E = G(P_Q, R, WA, WI, TE, MW).$$

The major objective of this study was to determine the net effect of water availability on location of employment in water-oriented manufacturing. This goal can be achieved by estimating the reduced form equation (11) where employment is related to water availability and other location factors.

Identifying Water-Oriented Manufacturing

A search of previous studies does not reveal sufficient information to identify the role of water in decisions as to where to locate new output. In this study, manufacturing industries were included according to their national total annual water intake. It was assumed that the large water uses are the industries to be affected most by water project development. The two-digit industries (SIC classification) included in the analysis were Primary Metal Industry, Chemicals and Allied Products, Paper and Allied Products, Petroleum and Coal Products, and Food and Kindred Products. Table 5 shows that water intake in these five industries accounted for about 90 percent of the total water withdrawals by

Table 5. Water use data for 1964 for plants using more than 20 million gallons

SIC* code	Industry	Total water intake (billions of gallons)	Cumulative percentage of total manufacturing water intake (percent)
33	Primary metals	4,578	32.6
28	Chemicals and allied products	3,888	60.3
26	Paper and allied products	2,071	75.0
29	Petroleum and coal products	1,398	85.0
20	Food and kindred products	760	90.4
	Total manufacturing	14,045	100.0

*Standard Industrial Classification.

Source: Water use in manufacturing, 1963 Census of Manufactures' MC63(1)-10. U. S. Department of Commerce Bureau of the Census.

all manufacturing in 1964. In the first stage these five industries were aggregated into one group for the analysis and the location factors were related to all five water-oriented manufacturing industries together.

Study Area

Estimating the effects of water projects on the location of industries requires a study of the behavioral relationships between employment and location factors. For the estimation of behavioral relationships in the water-oriented manufacturing industries it was important to include areas with a wide range of variations in water availability distance from markets, and degrees of industrial development. Eastern

United States includes highly developed areas with a long industrial tradition (mainly in the North) and areas in which industrial development started only in recent years (South). In some areas industrial development is very limited (Appalachian region).

A contiguous area including fourteen states was chosen to reflect all these variations and to give a basis for application of the results in a wide range of developing areas. The following states were included: Pennsylvania, Ohio, Maryland, Virginia, West Virginia, Kentucky, North Carolina, South Carolina, Tennessee, Georgia, Indiana, Illinois, Michigan and Wisconsin.

The Variables

Employment (E) is the total employment in the five water-oriented manufacturing industries in 1964 by county.[1]

Other employment (OE) is total employment in all other industries (not including the five water-oriented industries in E) in 1964 by county.[1] A positive relationship between employment in water-oriented manufacturing and employment in other industries is expected, since concentration of industry is correlated with the availability of experienced labor in an area.

Manufacturing wages (MW) is average wages and salaries for employees in manufacturing industries in 1964 by county.[2] A negative relationship with employment in water-oriented manufacturing is expected since lower wages are expected to attract more manufacturing.

---

[1] County Business Statistics, Bureau of the Census, 1964.

[2] Ibid.

Market potential (MP) is a population sum weighted by distances to the county. A gravitation model was used to calculated the market potential of each county. Population in each SMSA in continental U. S. was divided by the square of its distance to the county and the results were summed. The market potential is a measure of demand for the product. Product price variations are relatively small in different markets, therefore the market potential is used to represent the price of the product in the location function estimated. The market potential is expected to be positively related to employment in water-oriented manufacturing. Larger market potential will affect mainly the market-oriented industries.

Water availability (WA) is measured in low flow miles. In each county streams were segmented using gauging stations and confluences as segment boundries. The ten-year minimum monthly flow for each segment was then determined and used to calculate low flow miles of the segment (miles of segment times its low flow). Total low flow miles in a county is the sum of low flow miles in all stream segments within the county. When stream segments coincide with county boundries, low flow miles were divided evenly between the two counties. Water availability measured in low flow miles represents the quantitative and qualitative effects of water on industrial location. We would expect a degree of substitution between miles of stream and minimum flows. At a given low flow there will be a minimum distance required between two

[1]Minimum monthly flows by gauging station for a ten-year period are given in "A Compilation of Records, 1950-1960," The Geological Survey.

given plants in order to allow downstream withdrawal by the downstream plant. This minimum distance can be decreased if the minimum low flow in the stream is increased. We would expect a positive relationship between water availability and employment in water-oriented manufacturing.

Form of the Function

Assuming exponential relationships (Cobb-Douglas form) in the structural equations of the model (equations (1)-(5) we obtain exponential reduced form equations (equations (10)-(11)). A Cobb-Douglas type production function allows for diminishing marginal products and interaction between inputs. Exponential relationships in the supply function for water and employment assume increasing costs at an increasing rate when the quantity used increases. By replacing the price of the product by a market potential variable (MP) and including the effect of all other raw materials R in the constant term (A), the explicit form of equation (11) becomes:

$$(12) \qquad E = A(OE)^{\alpha_1} (MW)^{\alpha_2} (MP)^{\alpha_3} (WA)^{\alpha_4} e^u$$

where A is the constant term,u is the disturbance term and α_i are the elasticities.

Empirical Analysis for all Water-Oriented Manufacturing Together

Data for all counties in the study area was collected. The first stage of the study dealt with the determination and elimination of counties in which water was abundant. In the conceptual discussion above we defined four possible cases of water-use relationships. The first two, Economic use (A) and Free use (B), describe cases in which water is presently used by manufacturing and they differ in the degree of water scarcity. Water is abundant in case B while it is scarce in case A. The other two describe cases where water is not used. In case C water availability is the restricting factor while in case D other factors are more restricting for industrial development.

The counties in the study area were first divided into two groups. The first included all cases in which water was used (employment in water-oriented industries was reported) and cases in which water was not used.[1] Only counties in the first group were included for further analysis. Counties in case C in which water was the restricting factor are affected by water were not included mainly because of the difficulty of separating these cases from cases in which there was no employment as a result of the level of other location factors.

The second stage of analysis was directed to the identification of counties in which case A (economic use) exists and their separation from case B counties (free use). We would not expect a significant behavioral relationship between water availability and employment in counties where water is abundant (case B). In counties with abundant water any increase in water availability will have no effect on employment in water-oriented manufacturing. All counties included in this stage with employment in water-oriented manufacturing were ranked according to water intensity (intensity of water use). Water intensity

---

[1]Counties in which employment in water-oriented industries was not reported because of disclosure reasons were also omitted from the analysis. Data was not available for counties in which the number of plants was small and individual plants could be identified easily if data was not disclosed.

intensity was defined as the ratio of total water use by the five water-oriented manufacturing industries (million gallons/day) and the total water availability (cubic feet per second X miles). The county observations were ranked according to their water intensity ratio (the lowest ratio first). The observations ranked by increasing water intensity represented also an increasing scale of water scarcity, starting with the lowest degree of scarcity (or highest degree of abundance). The group of observations to be included in the final location function estimation was determined through a stepwise process. Ten different levels of water intensity were chosen. In each case the observations were split into two parts. Part I in each case included all observations with water intensity below the spliting level and Part II included the observations with water intensity equal or more than the spliting level.

The location function (equation (12) above) was estimated for each of the two parts in all ten groups. Partial results based on 1964 data for 218 counties in the study area are given in Table 6. Water elasticity estimates (b_w) and its t-level for all the different combinations of water intensity (WI) ranges are presented. These elasticity coefficients (b_w) are estimates of net water availability effect in each subgroup estimated through multiple regressions allowing for effects of other location factors. The coefficients of determination (R^2) for each equation are also presented in Table 6.

The choice of the spliting level that separates case A from case B was based on the assumption that there is a significant water effect in case A and an insignificant water effect in case B. Therefore, the

Table 6. Comparison of regression results for different water intensity ranges

Group	Part I				Part II			
	WI less than	b_w	t	R^2	WI equal or more than	b_w	t	R^2
(1)	0.005	0.196	2.1	0.218	0.005	0.096	2.5	0.743
(2)	0.01	0.132	2.2	0.194	0.01	0.132	3.5	0.759
(3)	0.03	0.047	0.8	0.188	0.03	0.169	4.0	0.769
(4)	0.05	0.104	2.3	0.270	0.05	0.163	3.7	0.779
(5)	0.1	0.111	2.5	0.305	0.1	0.241	5.0	0.806
(6)	0.2	0.112	2.6	0.362	0.2	0.288	5.8	0.827
(7)	0.5	0.140	3.2	0.466	0.5	0.376	6.7	0.848
(8)	1.0	0.128	3.1	0.533	1.0	0.457	5.6	0.854
(9)	2.0	0.124	3.4	0.605	2.0	0.476	6.8	0.908
(10)	5.0	0.078	2.2	0.643	5.0	0.581	5.4	0.923

WI = water intensity.

b_w = elasticity of water availability.

t = significance level of b_w.

R^2 = coefficient of determination.

criterion used in determining this level was to choose the group of equations in which water availability is not significant in Part I and significant in Part II and R^2 is low in I and relatively high in II.

Comparing the t-values in Table 6 (significance levels of the water availability coefficients) we observe that values in Part II are higher than the corresponding values in Part I. Water elasticities (b_w) and coefficients of determination (R^2) are increasing when the cut-off point of water intensity level is increasing in Part II. These results suggest that when water becomes more scarce changes in water availability will have larger effect on employment in water-oriented industries.

Increasing the cut-off point of water intensity adds more water scarce observations to Part I equations. No specific trend in the elasticities is observed but they are expected to increase with the increase in number of water scarce observations. One would expect a low degree of significance of these coefficients (low t-values) when water is abundant and increasing significance when the number of observations with water scarcity is increasing.

The lowest R^2 and lowest t-value in Part I appear in equation (3) where the cut-off point was at a water intensity of .03.[1] This WI level was, therefore, chosen to separate the observations between water-abundant and water-scarce counties.

---

[1]Water intensity is measured in units of:

$$\frac{\text{million gallons per day of water use}}{\text{cubic feet per second of low flow X miles of stream}}$$

The first two equations in Part I show a low R^2 but a significant water elasticity (b_w). This result could be interpreted as the result of other water effects like navigation. However, water availability together with the other variables explain only a small part of the variation in employment.

150 counties within the fourteen state area showed in 1964 a water intensity ratio larger than 0.03. The estimated location equation (12) for these counties, written in its logarithmic form is:

$$(13)\quad \text{Log } E = -\ 3.118 + \underset{(12.7)}{.775 \log OE} + \underset{(2.3)}{.796 \log MW} + \underset{(.8)}{.112 \log MP} + \underset{(4.0)}{.169 \log WA}$$

$$R^2 = .729$$

(t-values are presented in parentheses.)

Equation (13) shows a very significant water effect on employment in water-oriented manufacturing in counties where water is not available in abundance. The coefficient of water availability (WA) is positive and shows that an increase of 1 percent in water availability (WA) is accompanied on the average by an increase of 0.169 percent in employment (E). The effects of coefficients of total employment in other industries (OE) and manufacturing wages are also highly significant and about 73 percent of the total variation in employment in water-oriented manufacturing can be explained by the variation in the variables included in the equation ($R^2 = .729$). The effect of total employment in other industries (OE), is positive and high as expected. A 1 percent increase in total employment in other industries will increase employment in water-oriented manufacturing by 0.775 percent. This result is consistant with the hypothesis that industries tend to concentrate. A larger labor force in given area increases the availability of skilled workers and indicates a higher level of economic activity (see the conceptual discussion above). The effect of manufacturing wages (MW) is positive and its coefficient is significant. The result can be explained by the correlation between wages and labor skills. If wages are in fact a proxi variable for skills we should expect a positive relationship with employment instead of the negative effect expected if wages represent only the price of labor.

The effect of market potential (MP) on employment is positive but not significantly different from zero (low t-value). The low significance of the coefficient shows that the production in the five water-oriented manufacturing industries is not consumer-oriented and its location is not significantly affected by distance from consumers. This result could be justified in the heavy industries which produce mainly raw materials for other industries and directly for the consumer.

Equation (13) shows the relationship between employment in water-oriented manufacturing and major factors affecting its geographical location. The major significance of the estimated equation is the positive net effect of water availability. The differences in employment between counties which have the same level of total other employment and manufacturing wages can be explained by differences in water availability. These results can be used in projecting the effects of water development. Water projects that will add to water availability of an area in which water is not abundant (by increasing the minimum low flows) will make the area more attractive to water-oriented industry and we would expect an increase of 0.169 percent in employment for a 1 percent change in water availability.

Individual Industries--Logarithmic Functions

Individual location equations were estimated by applying equation (12) to each of the five water-oriented manufacturing industries. The number of counties included in the analysis of each industry was determined by the number of nonzero observations of industry employment within the counties in which employment in water-oriented manufacturing was reported and water was not in abundance (total of 150 counties).[1]

Regression coefficients for the individual industries and for four of the industries together (excluding the metal industry) are presented in Table 7. A significant water effect was found only in the metal industry (33) and the aggregated group of four industries (26-29). In both cases the coefficients are positive and higher than for the aggregate of all five industries (equation (13)). R^2's are low in the paper industry (26), chemicals (28) and petroleum (29) and relatively high in the food (20) and metals (33). Most of the variation in employment in the food industry is explained by employment in other industries.

A new variable was introduced in the industry equations intending to capture the effect of water use in other water-oriented manufacturing (OWU). The values of this variable represent the total water used by the other four water-oriented industries in the county. This variable had no significant effect in the food, paper and chemical industries separately but significant and negative when four industries (the

---

[1]The logarithmic function is limited to nonzero observations.

Table 7. Industry logarithmic equations[a]

DV	Regression coefficients						
	C	LOE	LMW	LMP	LWA	LOWU	R^2
LE 20	.466	.905	-.455	-.147	.039	-.036	.640
N = 50		(7.5)	(-.9)	(-.8)	(.4)	(-.4)	
LE 26	-6.252	-.003	2.622	-.260	.301	.003	.375
N = 37		(-.01)	(2.7)	(-.70)	(1.6)	(.02)	
LE 28	3.660	.432	-1.030	-.056	.049	.044	.147
N = 33		(1.4)	(-.9)	(-.2)	(.2)	(.2)	
LE 29	4.448	-.348	2.825		-2.034		.270
N - 5		(-.2)	(.5)		(.4)		
LE 33	-9.326	.653	2.295	.422	.549	-.194	.575
		(3.2)	(2.6)	(1.2)	(3.33)	(1.4)	
LE 20-29	-2.3925	.910	.341	.134	.301	-.237	.656
N = 33		(6.5)	(.4)	(.4)	(2.2)	(-2.0)	

[a]DV is dependent variable, C is constant term, LOE is total employment in other industries in logarithms, LMW is manufacturing wages in logarithms, LMP is market potential, LOWU is water use by other industries in logarithms, LE 20 is employment in the food industry, LE 26 is employment in the paper industry in logarithms, LE 28 is employment in the chemical industry in logarithms, LE 29 is employment in the petroleum industry in logarithms, LE 33 is employment in the primary metal industry in logarithms and N is number of observations.

above three plus petroleum) were aggregated. Some significance was also found in the metal industry. In both cases the coefficients are negative as expected.

Individual Industries--Linear Functions

A different functional form was tested for the individual industries attempting to improve the degree of explained variation and to avoid the discarding of all zero employment observations required by the logarithmic form. In addition, the counties were divided into water intensity groups according to total water use by all five industries together. The effect of water availability and other factors (included in the logarithmic form) were estimated for each industry in each of the three water intensity groups. Regression coefficients are presented in Table 8. A significant water effect can be found in most cases and its magnitude is increasing in higher water intensity groups.

A water effect in the petroleum industry (29) can be found only in the high water intensity group. This can be explained by the fact that most petroleum manufacturing employment in the study area is located in counties with high intensity of water use.

The effect of water use by other water-oriented manufacturing (OWU) on employment in a specific industry is significant and negative in most equations. In the metal industry (33) this effect becomes negative and significant only at the high intensity group where the heavy water use in this industry is limited by the high use of other manufactures competing for the same restricted water resources.

Table 8. Individual industries--linear regressions[a]

WI group		DV	Regression coefficients C	OE	MW	MP	WA	OWU	R^2
I.	.01-.09	E 20	258.0	.031	-.015	-17.199	.019	-.531	.396
	N = 60			(5.2)	(-.1)	(-1.2)	(1.6)	(-2.3)	
		E 26	18.531	.001	-.015	-1.666	.007	-.110	.362
				(.8)	(-.5)	(-.8)	(4.9)	(-3.3)	
		E 28	22.913	.001	-.03	.591	.009	-.142	.529
				(.7)	(-1.3)	(.3)	(7.4)	(-3.9)	
		E 29	No observations						
		E 33	-1.969	-.001	-.009	1.251	.002	.008	.144
				(-.4)	(-.6)	(1.5)	(1.8)	(.5)	
II.	.10-.99	E 20	1519.6	.031	-1.407	15.581	.017	.003	.660
	N = 56			(7.3)	(3.8)	(.9)	(1.5)	(.1)	
		E 26	-366.6	.000	.619	-21.344	.062	-.093	.428
				(.002)	(1.3)	(-.9)	(5.0)	(-3.05)	
		E 28	266.68	.012	-.094	-21.271	.037	-.066	.267
				(2.3)	(-.2)	(-.9)	(2.8)	(-2.2)	
		E 29	-57.753	-.00	.056	1.310	.00	.00	.051
				(-.6)	(1.1)	(.5)	(.2)	(.04)	
		E 33	-622.01	.009	.287	8.501	.051	.004	.469
				(1.4)	(.5)	(.3)	(3.3)	(.1)	
III.	1.0+	E 20	-942.9	.018	.975	15.549	.096	-.069	.892
	N = 25			(7.7)	(1.6)	(.8)	(1.5)	(2.1)	
		E 26	643.23	.007	-.098	-43.479	.191	-.062	.713
				(1.9)	(-.1)	(1.5)	(2.6)	(-1.6)	
		E 28	-41.450	.001	-.163	21.588	.111	-.021	.852
				(.5)	(.4)	(1.7)	(2.4)	(-.6)	
		E 29	-143.15	.001	.235	-14.439	.083	-.037	
				(.6)	(.7)	(-1.3)	(2.16)	(-1.6)	
		E 33	142.13	.010	.021	-25.019	.331	-.175	.979
				(6.7)	(.1)	(-2.4)	(19.3)	(-11.4)	

[a] $\text{WI = water intensity} = \dfrac{\text{mgd of water used by water-oriented manufacturing}}{\text{cfs of minimum low flow X miles of stream}}$

All other notations are the same as in Table 7.

The individual industry equations show the interrelations between water-oriented industries and the increasing importance of these interrelations when water scarcity is increasing.

A water project affecting water availability will affect all five industries. To get the effect on employment in each industry it will be necessary to solve for all five simultaneously. The equation for industry i will be:

$$(15) \quad E_i = b_{oi} + b_{1i}OE + b_{2i}MW + b_{3i}MP + b_{ui}WA + b_{5i}OWU_i \qquad i=1,2,3,4,5$$

The employment in industry i is a function of total employment in other industries, manufacturing wages, market potential, water avilability and water use by the other five industries. Only E and OWU have an index i. All other variables are the same for all the five industries.

Equation (15) can be rewritten in the following way

$$(16) \quad E_i = C_{oi} + C_{1i}WA + C_{2i} \sum_{j \neq i} K_j E_j \qquad i,\ j = 1,\ 2,\ 3,\ 4,\ 5$$

where $C_{oi} = b_{oi} + b_{1i}OE + b_{2i}MW + b_{3i}MP$

and $\sum_{j \neq i} K_j E_j = OWU_i$

and (K_j are the water per employee factors used to calculate water use by other industries).

For the five industries we have five equations and five unknowns (The E_i's) and we can colve the equations simultaneously.

The group of equations to be used will be determined by the water intensity in the area to be affected by the water project.

Conclusions

Water availability is significantly related to employment in the five water-oriented industries: Food, paper, chemicals, petroleum and primary metals. Changes in water availability in areas where water is not available in abundance can be expected to be followed by increases in employment in the water-oriented industries.

Water availability measured in low flow miles represents a combined relationship of water quantity and water quality. This variable takes in account the capacity of streams in an area (county) to accept wastes. Length of stream and low flows are both important in determining the reuse factor of a stream (at a given level of treatment) and the industrial development of the area.

The effect of water resource development projects which add to water availability by increasing low flows (low flow augmentation), on employment in water-oriented manufacturing can be estimated by applying the results of this study. The expected change in employment in all five water-oriented industries together can be calculated by applying the water elasticity coefficient in equation (13). A 1 percent change in water availability will on the average 0.169 percent to employment in the five industries in areas where water is not in abundance.

Structural Unemployment in the Evaluation of Natural Resource Projects*

I. RELATIONSHIP BETWEEN NET MIGRATION AND EXCESS POPULATION

Introduction

The location of a water resource development project in an area affects job opportunities facing different sex and occupational groups differently, depending upon the nature of industries affected by project location and the sex/skill composition of labor demanded in the affected industries. Change in job opportunities affects net migration to or from the area and consequently the population in the area. Further net migration response to changes in job availability varies according to age, sex and type of labor involved. Hence for studying employment and population changes as a consequence of the location of a project in an area, labor must be subdivided into reasonably broad homogeneous groups by sex, occupation and age.

The basic hypothesis advanced is that net migration of a group is related to the excess population of the group, where excess population is defined as the difference between the actual population and the desired population relative to the actual level of employment available to the group. The desired population is defined as that population which relative to the actual jobs available to the group, would have the desired employment participation. Net migration response coefficient of the group is equal to the proportion of the excess population that will net migrate.

Consider a group i and a time interval (0, 1). If the group has survived population P_i^s at the end of the time interval at t = 1,

* G. K. Kripalani, University of Western Michigan, Kalamazoo, Michigan.

employment E_i, the desired employment participation rate λ_i^*, then according to the above hypothesis, the number of net migrants of the group M_i will be given by

$$(1) \qquad M_i = g_i(E_i/\lambda_i^* - P_i^s)$$

$$(2) \qquad = g_i(E_i - E_i^*)/\lambda_i^*$$

where g_i denotes net migration response of the group and is equal to the proportion of the surplus population that will net migrate and $E_i^* = P_i^s \cdot \lambda_i^*$ denotes desired jobs relative to the group population P_i^s.

(1) may also be written as

$$(3) \qquad M_i/P_i^s = g_i(E_i/P_i^s \lambda_i^* - 1)$$

or

$$\mu_i = g_i\left(\frac{E_i - E_i^*}{E_i^*}\right) = g_i\left(\frac{E_i}{E_i^*} - 1\right)$$

where $M_i/P_i^s = \mu_i$, the rate of net migration and $E_i^* = P_i^s \lambda_i^*$ = desired jobs. The real significance of this form of the model is that the net migration rate is proportional to the percentage deficit in jobs, a measure corresponding to potential rate of unemployment in Mazek's model (discussed later).

If P_i is the actual population of the group at t = 1, the relationship between the actual population P_i and survived population P_i^s is given by

$$(4) \qquad P_i = P_i^s(1 + \mu_i).$$

Substituting (3) into (4) we have

$$(5) \qquad P_i = P_i^s(1 - g_i) + g_i E_i/\lambda_i^*.$$

g_i is a measure of the mobility of the group in question. $g_i = 1$ for a perfectly mobile group while $g_i = 0$ for a perfectly immobile group. When $g_i = 1$, $P_i = E_i/\lambda_i^*$ or the population at the end of the time interval at t = 1, is the desired population corresponding to the level of jobs at that time. For a perfectly immobile group $g_i = 0$, and the population at the end of the time interval at t = 1 is $P_i = P_i^s$ or the survived population.

Comments on Jansen's and Mazek's Models

Jansen (1966) investigated and estimated statistically a hypothesized behavior relation, in which net migration is proportional to job deficit where job deficit is defined as the difference between desired employment and actual employment. The relationship was expressed as $M = C(E - E^*)$ where M = number of net migrants (net inmigrants being taken as positive), E is the actual employment available to the group and E* is the desired employment. Desired employment is the product of the population of the group P, multiplied by the desired employment participation rate λ^*. (It is assumed that λ^* is fixed and is not a function of available jobs E). The desired employment participation rate is defined as the maximum supply of labor per unit population and is a concept that replaces the idea of labor force participation rate. Jansen's reasons for the use of the concept of employment participation in preference to labor force participation were that the employment participation concept reflects the equilibration of both supply and demand forces in the labor market and that the measure is based on the objective consideration of whether a person is working and does not depend on how he chose to answer the question of whether he was looking for a job. In Jansen's study the dependent variable represents the number of persons net migrating which represents both the labor force and nonlabor force members of the population, while his independent variable, the job deficit, is a measure related to labor force members only.

Mazek (1966) also investigated a relationship of the same general type. His hypothesis was that net migration rate is a linear function of the rate of potential unemployment. This rate of potential unemployment is defined as the unemployment rate which would exist in a region at the end of the period studied if no migration--in or out--took place during the period. It is important to emphasize that net migration rate in Mazek's study is related to members of the labor force only. Thus in Mazek's work, both the independent and dependent variables relate to labor force members and not to populations. If M_L denotes the number of net migrants of the labor force during a time interval, L is the labor force (supply of labor) at the end of the time interval and E is the actual employment at the end of the time period, then Mazek's model may be written as

$$M_L = b(L - E)$$

ignoring for this exposition the constant term and other variables. The number of labor force members net migrating is thus proportional to the number of potential unemployed workers. The coefficient b represents the proportion of potential unemployed workers who would respond to this situation by net migrating out of the area.

In an important sense, the models of Jansen and the modified form of Mazek (in which the constant term and the independent variables are assumed absent) are similar. The labor force L of Mazek takes the place of E*, the desired employment of Jansen. In both the models the independent variable represents the difference between the supply of and demand for labor at the going wage. The dependent variable of Mazek is the number of labor force members net migrating while that in Jansen's model represents the population net migrating including the job seekers and the associated non-job seeker component. Thus, the relationship between the dependent variables is $M_L \neq M\lambda^*$. Consequently, the response coefficient is given by

$$C = b/\lambda^*.$$

In a certain sense, there is consistency in Mazek's model in that both the dependent and independent variables relate to members of the labor force. This is not so in the case of Jansen, whose independent variable relates to the members of the labor force only but whose dependent variable relates to both labor force and nonlabor force members of the population. On the other hand, it is considered that the concept of labor force as observed used by Mazek is not very satisfactory and that the recorded labor force data do not correctly represent the supply of labor. Besides the subjective nature of the concept, there is what is referred to as the discouraged worker hypothesis. According to this hypothesis, as unemployment rises, a portion of the labor force becomes discouraged with its employment seeking efforts and withdraws from the labor force. Also there is the secondary worker hypothesis in terms of which when the rate of unemployment for primary workers rises, the secondary workers e.g. married women enter the labor force, augmenting the observed supply of labor in relevant categories and groups.

For the reasons adduced above, the proposed model is based on the hypothesis that the number of net migrants is proportional to excess population. In this formulation, the labor force concept of Mazek is given up in favor of Jansen's concept of desired employment participation. On the other hand, Jansen's inconsistency (of a sort) of relating population (consisting of labor force and nonlabor members) to a measure that refers only to labor force is remedied by introducing populations on both sides of the relation.

A Property of the Proposed Model

A property of the proposed model relationship between net migration and surplus population (equation 1) or between net migration and job deficit (equation 2) is that when job opportunities available to a group increase as a result of the location of the project, net out-migration of the group is reduced or net inmigration increased; and hence in a net out-migration area a proportion of the new jobs is appropriated by those net migrants whose net outmigration has been withheld due to increased job opportunities in the area. An increase in job opportunities by ΔE_i would mean that the resultant net inmigration or withheld net out-migration is equal to

(6) $$\Delta M_i = g_i \, \Delta E_i / \lambda_i^*.$$

It is reasonable to assume that these net outmigrants who would have net outmigrated from the area in the without-project situation but who now stay in the area in the with-project situation would do so only when they have the desired employment participation in the area. In other words, the jobs appropriated by these net migrants out of the ΔE_i addition jobs are given by

(7) $$\Delta M_i \lambda_i^* = g_i \Delta E_i.$$

This means that the net reduction in the number of the unemployed in the area (assuming it to be a net outmigration, i.e. a job deficit area) is given by $(1 - g_i)\Delta E_i$ or the net reduction in the number of the unemployed per unit new job is given by

(8) $$1 - g_i.$$

Now g_i is the coefficient of net migration response of a group. A perfectly mobile group has $g_i = 1$ and a perfectly immobile group has $g_i = 0$. The impact on the unemployed per unit job is thus greater for relatively more immobile groups than for these groups which are relatively more mobile.

Relationship Between the Parameters in the Three Models

The three models discussed above are:

(9) Proposed $$M = g(E/\lambda^* - P^S)$$

(9') $$= g(E - E^*)/\lambda^*$$

(10) Jansen $$M = C(E - E^*)$$

(11) Mazek $$M_L = b(E - L^S)$$

It is proposed to express b and C in terms of g, the parameter of the proposed model. Since it is not proposed to make direct estimates of g from actual data for use in the present study, it is necessary to make use of estimates of parameters b and C already available in Mazek's and Jansen's studies and to convert them to corresponding g estimates, and to use g estimates thus obtained in the present study.

Considering the equations (9) and (10) since $E^* = P^S \lambda^*$, it can easily be seen that

(12) $$g = C \lambda^*.$$

Mazek's model relates labor force net migrants to potential unemployed in the labor force, a measure analogous to job deficit of Jansen. Mazek used observed survived labor force as the measure of labor supply or desired jobs, while Jansen makes an estimate of desired jobs by using the concept of desired employment participation. Obviously the two concepts are fundamentally different and no means can exist for reconciling the two and further no fixed relationship can or may be expected to exist between the two.

To express a relationship between g and b (or between C and b), two relations are required, one connecting L^S with E^* and the other connecting M_L and M.

Consider for example males aged 30-34 in a depressed area. If the discouraged worker hypothesis holds, the observed labor force of males aged 30-34 L_S in such an area will be lower than the number of desired jobs E^*, which represents the maximum supply of labor at the going wage. On the other hand in such an area, if the secondary worker hypothesis holds, the observed labor force of females in certain relevant age groups may be higher than the number of desired jobs. If areas of net inmigration or area of small net outmigration are considered, one would generally expect that the observed labor force and the desired jobs would be fairly close to each other. For the sake of simplicity and practical needs of the situation, let the differences between L^S and E^* be disregarded and let $L^S = E^*$. Let the relationship between M_L and M be taken as:

(13) $$(1 - k)M = M_L.$$

This means that the nonworking force complement of net migrant persons is taken at k times the migrant population, the remaining being workers or members of the labor force. Hence

(14) $$(1 - k)M = b(E - E^*).$$

Eliminating $(E - E^*)$ from equations (9') and (14) we have:

(15) $$g = b\,\lambda^*/(1 - k).$$

Consider the quantity $(1 - k)$ in the relation $(1 - k)M = M_L$. We are concerned with net migrants and it may seem a reasonable argument that the associated nonworking component of these net migrants may be expected to be significantly lower than that in the parent population out of which the net migrants outmigrate. Hence a reasonable assumption may be that in general

(16) $$1 - k > \lambda^*.$$

Hence we have

(17) $$g < b.$$

Equation (15) provides a basis for converting Mazek's b estimation into corresponding g estimates. Having regard, however, to the nature of the problem of this study, it may be recognized that a basis for extreme precision does not exist and that as a practical measure it may be reasonably accurate to disregard the differences between $(1 - k)$ and λ^* and to assume

(18) $$g = b.$$

What are the implications of assuming g = b in place of g < b in the case of depressed areas? The consequence is that the quantity (1 - g) is underestimated and hence the net employment effects of location of a project in a depressed area are underestimated, i.e. we will have obtained rather conservative estimates of the impact of the project, a feature which may be considered a desirable one.

Comments on the Use of Net Migration Response Coefficients Estimated by Mazek

Jansen (1966) estimated parameter c by age groups 25-34, 35-44, 45-64 and 65+ for unskilled males, based on analysis of 1950-60 data for (a) all SMSA's (b) SMSA's with inmigration and (c) SMSA's with out-migration. Jansen's work did not provide c estimates for females and for age groups under 25 years in the case of unskilled males. Due to this reason, it was not possible to make use of Jansen's c estimates for estimation purposes in this study.

Mazek (1966) obtained estimates for the parameter b by occupation groups with all age groups combined. He also estimated parameter b for males and females separately subdivided by age groups 20-24, 25-29, 30-34, 35-44, 45-54 and 55-64. In the case of males, therefore, Mazek's work did not provide estimates of net migration response coefficient by age for the group of occupations included in the unskilled male category.

Our requirements in terms of the proposed model are to know estimates of net migration response coefficients by age in respect of (a) low-skill males and (b) females. Since Mazek provided estimates for all males (not for low-skill males) and females separately, it was decided to use Mazek's estimates. This decision necessitated that consideration be given to the problem whether some means could be devised to adjust Mazek's estimates for all males to obtain corresponding estimates in respect of low-skill males. b for high-skill males is unity and hence b for all males is a weighted average of b's for high-skill and low-skill males, the weights depending upon the distribution of actual employment and of labor force in high-skill male and low-skill male categories. Hence, there is no simple general relationship by which b for all males could be translated into corresponding b for low-skill males, taking b for high-skill males equal to unity. All that can be said with certainty is that b for low-skill males is lower than b for all males. The use of Mazek's b for all males without any adjustment would therefore tend to underestimate net employment benefit effects in our calculations.

Back Migration Effects

If there are people who grew up in a depressed area and who have moved elsewhere to work but would like to return to the area, they may return from growing centers and take jobs leaving those originally unemployed in depressed areas still unemployed. Some case studies do provide evidence for this type of phenomenon. Such back-migration effects may however, not generally be of significant dimension for an area which is really a depressed area. Families incur substantial costs, both pecuniary and nonpecuniary when they migrate out, and similar costs have to be incurred again if they back-migrate. The fact that the area is a depressed area is important from the point of future of the children and unless the person or the family in question is unable to adjust itself to the new environment or faces a prolonged period of unemployment in its new location, such back-migration to a depressed area may indeed be small unless the project in question is of such large dimensions that it is likely to have pronounced long-term effects on the character of the area to convert it from a depressed to a growing area. This means that if the area with the project will still be an area of substantial positive labor supplies, the effects of back-migration of past migrants may be small, if not negligible. Whether such effects are, in fact, small or significant cannot obviously be decided on a priori arguments but will be decided on the basis of survey in the area concerned or on the basis of surveys of similar projects in similar type of areas. A view is sometimes held that back-migration is an observed phenomenon and that there are persons who are so "attached" to the area where they were brought up that they would migrate back to the area whenever they could do so and give up their jobs in their present location. Even if such a phenomenon did exist, no further adjustment in our calculations is called for since our study of the relationship between population and employment is based on net migration and in view of our assumption that population (or net migration) changes induced by the project have the desired employment participation (or net migration) change whether such change consists entirely of withheld outmigration or partly of withheld outmigration and partly of backmigration is immaterial as long as the total net population change is the same, since both the back migrants and withheld outmigrants are assumed to have the same desired employment participation. $(\Delta M)\lambda^*$ $= (\Delta M_1 + \Delta M_2)\lambda^*$ since $\Delta M = \Delta M_1 + \Delta M_2$ where ΔM_1 is withheld outmigration and ΔM_2 is backmigration. The need for adjustment will arise if and only if the assumption is made that λ^* for withheld outmigrants and λ^* for back migrants are not the same. On a pirori reasons such an assumption does not seem to be justified and further, to do so would involve introduction of complications in the calculations unwarranted by the degree of precision that can be achieved in these calculations. We will therefore disregard any such differences in λ^*.

The above discussion has assumed that the net migration response coefficient has been estimated on the basis of data pertaining to areas in the case of which backmigration is a significant phenomenon. But when such is not the case, i.e. the estimate of the net-migration response coefficient is based on observation units in the case of which back migration phenomenon is nonexistent or negligible, then it could not be said that the use of such parameter estimates takes care of back-migration effects simply because the analysis is based on net migration data. For example, in Jansen's study, most of 70 SMSA's used in the estimation of C were located outside depressed areas and hence $\hat{C}$ may not be the proper estimate to use when a depressed area is involved. Mazek's sample of SMSA's included in his study included a total of 47 SMSA's "within an area which is roughly the north-eastern quadrant of the United States." Fifteen of these SMSA's had average unemployment rate exceeding 6 percent in 1954-1960, while only 10 SMSA's had unemployment rate exceeding 7 percent in the same period. Hence it may be that the use of Mazek's net-migration response coefficient to the Youghioghenny River Reservoir Project area does not make adequate provision for back-migration effects. Somers (1954) in his study of "the employment histories, over a 12-year period, of 1,015 persons hired by a chemical manufacturer" in Morgantown, West Virginia, in 1951-52 found that "one-fourth returned to Morgantown because the reopening of the chemical plant made 'good' jobs available at home."

Back-migration effects are automatically provided for if net-migration response coefficient is estimated on the basis of data relating to areas of the same type. When such is not the case, rough adjustment on judgment basis may be made for loss of a proportion of additional jobs to such backmigrants.

II. BASIS FOR ALLOCATING ADDITIONAL JOBS DUE TO THE PROJECT AMONG AGE GROUPS

Preliminaries

The discussion in Chapter I on estimating the impact on the area's unemployed measured by the reduction in the number of the unemployed due to increase in job opportunities assumes that the additional jobs for the group as a result of the project is known. It is assumed here that total additional employment ascribable to the project (defined as the difference between employment in the with and without-project situations) is given in terms of three main categories viz. high-skill male jobs, relatively low-skill male jobs and female jobs. The high-skill male category is defined to include (i) professional, technical and kindred workers, (ii) farmers and farm managers and (iii) managers, officials and proprietors, excluding farm. It is assumed that the high-skill male category is perfectly mobile and that in the nation as a whole there is no excess supply of this type of labor. Hence the net employment effects of additional high-skill male jobs due to the project are assumed to be nil.

This study assumes that net migration response to job deficit or surplus population varies by age and sex and hence the net employment effects of additional employment due to project will depend on how additional jobs for relatively low-skill males and for females are distributed among age groups. The problem therefore is to find a suitable basis for dividing total additional jobs among age groups for each of these categories of labor.

Before the problem of an appropriate basis for allocating additional jobs is tackled, some further observations on the precise setting of the situation are necessary. Net employment effects of a project may vary over project's life time even if total additional employment of low-skill males and females due to the project is held constant over time. This is because the distribution of additional jobs over age groups within a category will vary over time as population distribution among age groups changes over time. This makes it necessary that the estimation procedure be outlined for more than one point of time. The discussion in this chapter is with reference to a past year, viz. 1960. (1960 has been selected because it coincides with the year of the last Census of Population, which is the main source of the statistical data required.) Some special problems no doubt arise in calculations pertaining to a point of time in the future, and these are discussed in Part IV. In a general sense, however, the estimation procedure for a point of time in the past (say, 1960) and in the future (say, 1980) follows along identical lines. But it can easily be seen that while for 1960 much of the statistical material is available, that is not the position for 1980 and one will be constrained to use projections for relevant variables. Obviously, projections, e.g. of employment in the project area in the without-project situation, are not easy to make nor do we think that detailed consideration of these problems is an important focus of this research.

The procedure outlined hereunder takes the 1960 employment, population, etc. situation in the project area as representing the without-project situation. The project is then superimposed on the area and its impact on variables of interest estimated.

Three Alternative Bases

Three quantities may be considered as likely candidates for serving as the basis viz:

(A) Job deficit in the without-project situation (denoted by $(E_i^* - {}^oE_i)$ where ${}^oE_i$ represents actual employment of the age group in the without-project situation in 1960.

(B) Desired job vacancies or the number of job seekers (denoted by $S_i) = E_i^* - E_i^h$ where E_i^h represents the number of persons in the age group i in continuous employment over 1950-60 decade.

(C) Actual distribution of job vacancies among age groups during 1950-60 decade in the without-project situation (denoted by ${}^oE_i^v = {}^oE_i - E_i^h$).

Use of Basis A would mean that (a) all qualified job seekers in the without-project situation have the same chance of becoming employed, irrespective of age. In this case, actually employed persons in 1960 ${}^oE_i$, consisting of those in continuous employment since 1950 viz E_i^h, and those who filled vacancies that arose in 1950-60 decade in the without-project situation, ${}^oE_i^v$, are excluded from desired jobs. Use of Basis B also assumes equal chance of becoming employed by qualified job seekers irrespective of age, but it does not exclude those who filled vacancies during the decade. If the equal chance assumption is valid, then 1950-60 vacancies also should have been filled on that basis, i.e. in proportion to S_i. If the area is a net deficit job area in the without-project and with-project situations, then additional jobs due to the project should also be distributed on the same basis since it makes no difference how the job vacancies arose, out of jobs in the without-project situation or due to the project.

When calculations are, however, made for a past year as in the present case, viz. 1960, we in fact have the information as to how the actual job vacancies in the without-project situation were distributed among age groups. This distribution is given by the quantities ${}^{o}B_i^v$. In general, one may expect that the observed distribution would not tally with the distribution on Basis A or Basis B both of which are based on the assumption of equal chance of getting employment by job-seekers of all age groups. It is obviously desirable that use should be made of the available information regarding actual allocation of jobs among age groups in respect of vacancies that arose in 1950-60 decade. This means that equal probability assumption of Basis A or Basis B which leads to distribution on basis of job deficit $(E_i^* - {}^{o}E_i)$ or desired job vacancies, i.e. job seekers S_i be given up and modified so as to make use of actual experience of 1950-60 decade. To handle this problem, we associate with each age group a relative probability coefficient p_i such that the ratio of the probability of becoming employed of a job seeker in age group i to that of a job seeker in age group j is given by p_i/p_j. On this basis (Basis C) the distribution of jobs among age groups will be based on the quantities $p_i \cdot S_i$.

The quantities p_i are not known and may be estimated. Since p_i's are relative, they may be so defined that

(1) $$\sum_i p_i S_i = \sum_i S_i.$$

We have

$$ {}^{o}E_i^v \cdot \sum_i {}^{o}E_i^v = p_i S_i \cdot \sum_i p_i S_i .$$

Hence since by definition $\sum_i p_i S_i = \sum_i S_i$, we have

(2) $$p_i S_i = \sum_i S_i \cdot {}^{o}E_i^v / \sum_i {}^{o}E_i^v.$$

p_i's can now be calculated since all the quantities in the above relationship are known.

A look at equation (2) will show that distribution of jobs on the basis of $p_i S_i$ is the same as the distribution on the basis of job vacancies filled in 1950-60 decade in the without-project situation viz. ${}^{o}E_i^v$ since the quantity $\sum_i S_i / \sum_i {}^{o}E_i^v$ is the same for all age groups and does not affect the distribution over age groups. Thus, a basis which drops the equal probability assumption of Basis A and Basis B leads to Basis C viz. that the distribution of additional jobs due to the project should be on the basis of the actual experience of 1950-60 decade in the matter of distribution of job vacancies. (Our calculations for the Youghiogheny River Reservoir project area showed considerable differences between the distribution on Basis A or Basis B and the actual 1950-60 experience Basis C and provide a strong evidence that the equal probability assumption is not supported by actual experience in the case of this project area.)

Choice of Appropriate Basis in Different Cases

Which of the three bases discussed above is appropriate in a given problem depends upon whether the area is an area of overall deficit (+) or surplus (-) of jobs for each category (viz. low-skill males and females) in the without-project and with-project situations. The following possibilities may arise.

(a) <u>The area is an area of overall job surplus in the without-project situation</u>. Normally, in such an area, the job seekers ($S_i = E_i^* - E_i^h$) during the decade out of the area's own labor supply should be positive. Use of S_i basis (Basis B) would mean that vacancies are allocated among age groups in proportion to area's own job seekers during the decade. Since the area is an area of overall job surplus in the without-project situation, it may be a reasonable practical assumption to make that all qualified job seekers of the area are already employed in the without-project situation. Obviously, therefore, S_i is not a valid basis since the additional jobs due to the project are to be filled not out of area's own labor supply but by net immigrants from the rest of the nation. The ${}^oE_i^v$ basis (Basis C) has the defect that it combines area's job seekers with the net immigrants both of which became employed during the decade; and makes no distinction between these two groups of job seekers. The additional employment due to the project is to be shared entirely by net immigrants of various ages groups, and hence the appropriate basis seems to be that given by $({}^oE_i^v - S_i)$. It may be noted that ${}^oE_i^v - S_i = ({}^oE_i - E_i^h) - (E_i^* - E_i^h) = ({}^oE_i - E_i^*)$ and shows how the overall job surplus in the without-project situation was allocated among various age groups of net immigrants during the decade.

In most cases, one would expect that ${}^oE_i^v - S_i$ for all i would be greater than or equal to zero and have the same positive sign. In some age groups, e.g., the highest age groups 55-64 and/or 65+, ${}^oE_i^v - S_i$ may be zero. The problem of allocation in these cases is straightforward and simple. But if some age groups display a job deficit, some arbitrary decisions to handle the problem on judgment basis become unavoidable. For example:

(i) If the area is an area of relatively substantial overall job surplus and the job deficits revealed by some age groups are relatively small, then (a) if such deficits occur in the case of older age groups, such deficits may be replaced by zero on the assumption that these represent 'unqualified' persons, and (b) if such deficits occur in young or middle-age groups and are relatively small in magnitude, they may be regarded as having arisen from 'error' and either replaced by zero with no further adjustment or replaced by zero with the deficit allocated equally among the adjacent age groups.

(ii) The real difficult problem, however, would arise in the case of an area which is an area of relatively small overall job surplus in the without-project situation. In such a case one may encounter positive and negative job surpluses, each of relatively small magnitude, distributed over age groups in a haphazard manner. Probably a practical and realistic way to handle this situation is to assume that the area is just in balance in the without-project situation, i.e. there is neither overall surplus nor deficit of jobs and that observed deficits and surpluses are due to 'error.' In such a case, additional jobs due to the project are assumed to be shared entirely by net immigrants of various age groups and since we have no experience to go by, the allocation may be according to net migration rates for the nation as a whole.

(b) <u>The area is an area of overall deficit of jobs in the without-project and with-project situations</u>. In this situation the additional jobs due to the project are to be filled by area's own labor supply. All the three bases are relevant since they all relate to the area's own labor force. The defect of Basis A (job deficit $E_i^* - {}^oE_i$ basis) and Basis B (S_i basis) is that they are based on equal chance assumption and in doing so, they disregard the evidence of 1950-60 decade of how actually the job vacancies were distributed among age groups. Basis C (${}^{o}E_i^v$ basis) therefore provides the appropriate basis for allocation.

In most cases, one would expect ${}^{o}E_i^v$ quantities to be all positive. In certain marginal cases and sometimes in the case of old age groups, one may come across negative values for some age groups. These cases may be handled on judgment basis by regarding them as due to 'error' with no further adjustment or with allocating the negative value equally among adjacent age groups.

(c) <u>The area is an area of overall job deficit in the without-project situation and an area of overall job surplus in the with-project situation</u>.

One may come across a case of this type when the project area is an area of small job deficit in the without-project situation or a case in which the project is a major project of great dimensions.

(i) If the overall job deficit in the without-project situation is small in relation to total jobs $\sum_i {}^{o}E_i$ or in relation to additional jobs due to the project, then such an area may be considered as an area in balance, with neither surplus nor deficit and all additional jobs allocated entirely to net inmigrants as in (a) above.

(ii) If the overall job deficit in the without-project situation is large in relation to additional jobs due to the project, such that the overall surplus in the with-project situation is small in relation to overall deficit in the without-project situation, then for practical purposes, the area may be handled as an area of type (b) and ${}^{o}E_i^v$ basis

(Basis C) be used for allocation purposes. (For evaluating net national employment benefits, however, the formula will be $\sum_i (1 - g_i)\Delta E_i^{**}$ where ΔE_i^{**} is the smaller of ΔE_i (obtained on Basis C) and the job deficit in the without project situation given by $S_i - {}^{o}E_i^{v}$).

(iii) If the area is an area intermediate between (i) and (ii) above, then a reasonable practical assumption may be that in the with-project situation, all the area's job seekers in the without-project situation ($\sum_i (S_i - {}^{o}E_i^{v})$ get employment and that the balance of the jobs due to the project given by $\Delta E - \sum_i (S_i - {}^{o}E_i^{v})$ are appropriated by net immigrants. The allocation of jobs among net immigrants may be made on the basis of net immigration rates for the nation as a whole, a procedure suggested in (a) above in the case of areas of overall job surplus, since we have no relevant experience to guide us. (Note ${}^{o}E_i^{v}$ basis relates to area's own labor supply.) The allocation of that part of the additional employment due to project that is assumed to neutralize area's overall job deficit is simple and is equal to each group's job deficit if all the age groups display positive job deficits. But if some age groups display negative job deficits or job surpluses in the without-project situation, then the problem posed by these age groups may be handled on judgment basis. For example, if job surpluses are small, they may be assumed to arise from error and ignored.

III. ESTIMATION OF NET NATIONAL EMPLOYMENT BENEFITS IN TERMS OF DOLLARS OF SALARIES AND WAGES

For the Year of Calculation

It has been shown in Chapter I that if as a result of location of the project in an area, ΔE_i additional jobs are allocated to a group i whose net migration response coefficient is g_i, then the number of persons employed who in the absence of the project would have been unemployed in the area is given by $(1 - g_i)\Delta E_i$. Let us denote the categories high-skill males by (hsm), low-skill males by (lsm) and females by (f). Then if ΔE = total additional jobs due to the project, we have

$$\Delta E = \Delta E(hsm) + \Delta E(lsm) + \Delta E(f)$$

$$\Delta E(hsm) = \sum_i \Delta E_i(hsm)$$

$$\Delta E(lsm) = \sum_i \Delta E_i(lsm)$$

$$\Delta E(f) = \sum_i \Delta E_i(f)$$

where ΔE_i refers to additional jobs allocated to age group i.

The total number of additional jobs taken up in the with-project situation by those who otherwise would have been unemployed in the area is given by

(1) $$\sum_i \{\Delta E_i(lsm) \cdot [1 - g_i(lsm)]\} + \sum_i \{\Delta E_i(f) \cdot [1 - g_i(f)]\}$$

where g_i(lsm) refers to net migration response coefficient of age group i of the low-skill male category and similarly g_i(f) refers to female category. In the above expression, no term for high-skill male category has been included on the assumption that $g_i(hsm) = 1$ for all i.

If W(lsm) and W(f) represents the average annual wage and salary of a low-skill male and female respectively, the equivalent annual net national wage and salary benefit may be calculated as

$$(2)\ B = W(lsm) \cdot \sum_i \{\Delta E_i(lsm) \cdot [1 - g_i(lsm)]\} + W(f) \cdot \sum_i \{\Delta E_i(f) \cdot [1 - g_i(f)]\}.$$

Over the Life of the Project

Let n denote the life of the project (in numbers of years) and let B_t denote the net national wage and salary benefit for the year t. Then the present value at t = 0 of the benefit over project's life time is given by

$$(3) \qquad \sum_{t=1}^{n} B_t(1 + r)^{-t}$$

where r is the rate of interest to be used in the calculations.

Net employment effects of a project may normally be expected to vary over the project's life time even if total additional employment of low-skill males and females is held constant over time on account of changes in the composition of the population and consequent changes in the distribution of additional employment over age groups. Besides, over the project's life time the character and the extent of job deficit in the area may change due to reasons independent of the location or otherwise of the project in the area. Where reasonably sound and valid bases can and do exist for such adjustments, they may be made. In general, the forecasting of possible future developments and their relevant magnitudes is a matter which cannot be extricated from serious uncertainties and wide margins of error and it may perhaps be best not to attempt to put numerical values on such changes.

If it is considered a reasonable assumption that the net employment benefit of the project is constant over the life time of the project and equal to the effect for the year of calculation and if further one assumes that wages and salary (representing labor productivity) will rise at 100a percent per year, then the net benefit over the life time of the project will be given by

$$(4) \qquad B \sum_{t=1}^{n} (1 + r)^{-t} (1 + a)^{t}$$

$$= B \sum_{t=1}^{n} (1 + r - a)^{-t} \text{ appr.}$$

$$= B\, a_{\overline{n}|}^{(r - a)\text{ percent}}$$

where $a_{\overline{n}|}^{(r-a)\text{percent}}$ is the present value of annuity of one payable for n years calculated at rate of interest (r - a).

IV. SOME ESTIMATION PROBLEMS WHICH WOULD ARISE IN CALCULATIONS PERTAINING TO A POINT OF TIME IN THE FUTURE

When net employment benefit calculations are made with respect to a point of time in the future, some new estimation problems arise. These are: (i) estimating total employment in each category of labor viz. high-skill male, low-skill male and female in the project area in the without-project situation (ii) estimating dependent net migration of wives and children and (iii) estimating transfers from low-skill male category to high-skill male category and vice versa. These are discussed below.

Problem of Estimating Employment in the Project Area in the Without-Project Situation

One of the inputs in the evaluation procedure for a point of time in the future, say 1980 via an intermediate point of time also in the future, say 1970 is the employment by category (viz. high-skill male, low-skill male and female) in the without-project situation in 1970 and 1980. This information is required (a) for ascertaining the in or out net migration character of the area, i.e. whether the area is an area of net deficit or net surplus of jobs in the without-project situation and the extent of overall net deficit on surplus, (b) for estimating 'dependent' net migration of children and of wives during 1960-70 decade and (c) for obtaining population projections for the project area in the with and without-project situations in 1970 which serve as base for 1980 calculations.

Insofar as we are aware, no employment projections by sector and by county or by sex sub-classified by occupation are available by county. Hence normally such projections will have to be made either for each county in the project area or for the project area as a whole, taking into account (i) changes in employment by sector in the recent past by county, for the project area and in the nation as a whole, (ii) changes in relevant technological and other factors affecting industry location and (iii) the anticipated national rate of growth of employment in each sector. Since projected employment by sector in 1970 and 1980 is to be subdivided into employment by sex and skill, likely changes in the state of technology relevant to each industry in the future affecting the sex and skill mix of employment must also be considered. We recognize that these detailed investigations into the choice of appropriate bases for employment projection and of anticipated changes in sex-skill mix of employment in different industries are outside the primary focus of this research.

For the purpose of demonstrating the working of an evaluation procedure therefore it will be assumed that the sex and skill mix of employment in each sector in the future will remain as in 1960. For projecting employment, the following five alternative bases may, among other suitable bases, be considered.

(a) Assuming that the 1950-60 <u>rate</u> of change of employment by sector in each county will apply in 1960-70 and 1970-80 decades.

(b) Assuming that the 1950-60 <u>amount</u> of change in employment by sector in each county will apply in 1960-70 and 1970-80 decades.

(c) Assuming that the 1960 total employment-population ratio by sex for each county will apply in 1970 and 1980. (Note this approach requires population projection by sex for each county for 1970 and 1980.)

(d) Assuming that the 1950-60 national rate of change of employment by sector will apply to the project area in 1960-70 and 1970-80 decades.

(e) Assuming that the 1950-60 project area's rate of change of employment by sector will apply in 1960-70 and 1970-80 decades.

Problem of Estimating 'Dependent' Net Migration of Wives and Children

In the case of females, the estimation of survived female population by age at time t + 10 (say in 1970) requires the consideration of two adjustments for 'dependent' female net migration (a) as wives and (b) as children of married men who net migrate. Rigorous treatment of these adjustments, particularly the former is a very complicated affair requiring the use of age distribution of wives by age of husband. The gain in accuracy likely to be achieved by adopting very complicated procedures to allow for the impact of these adjustments on the results of the estimation procedure will in general not be very significant and hence we do not consider that for our purposes such very complicated procedures have any justification. A simple procedure using 1960 census of population material is outlined below:

Let $(mm)_i$ denote the proportion of men aged i who are married. If $M(m)_i$ is the number of male net migrants aged i, then the number of married men net migrating is given by $M(m)_i \cdot (mm)_i$. This quantity also represents the number of 'dependent' net migration of married women. For this purpose, we propose to use information given in the U. S. Census of Population, 1960, Subject Reports, Marital Status Table 1 (Vol. PC (2)-4E) regarding proportion married by age and sex for the United States in 1960.

As regards the average age of wives of net immigrating married men, reference is invited to Appendix G which gives the age of husband and wife for married couples for the United States in 1960 and to the following table giving median difference between ages of husband and wife for the United States in 1960.

Table 1. Median difference between ages of husband and wife for the United States, 1960

Age	Median Age Difference (years)
All classes	2.7
Husband under 35	1.8
Husband under 35-54	2.8
Husband under 55+	4.2

Source: U. S. Census of Population, Subject Reports, Marital Status, Vol. PC (2)-4E (Table 9).

It will be observed that the average age of wife is lower than the age of husband by 2.7 years for all age groups and by 1.8 years when the husbands are aged under 35. In view of the fact that we are dealing with broad age groups, problems will arise of the basis for allocating dependent net migration of wives to appropriate age groups. To get over these problems and for the reason that the role of adjustment for dependent migration in our calculations is very minor, it is considered sufficiently accurate that age difference between net migrating married men and their wives may be disregarded and the average age of wives may be assumed the same as that of the net migrating husbands.

Adjustment for dependent net migration of children is necessary in the estimation of survived males and females for young ages relating to new entrants into the labor force during the decade t, t + 10 (e.g., during 1960-70) and in following years. For example, persons aged 14-19 in 1970 are not equal to the survivors of persons aged 4-9 in 1960 since some of the latter net migrated as children of net migrating married men in the decade. A simple method for making this adjustment is to use the information given in the 1960 Census of Population for the State of Pennsylvania. Table 110 of the State of Pennsylvania Volume No. PC (1)-Part 40 gives information about the average family size (number of family members) for all families as 3.61 persons and for husband-wife families as 3.69 persons. In these calculations, it may be assumed that the dependent net migration of children is related to husband-wife married families and that the average number of children per net migrating family is 1.69 children. It may further be assumed that half of these children are male and half of them are female.

The total number of children net migrating as 'dependent' children of net migrating married men may be calculated as follows. The number of married men aged i net migrating is given by $M_i(m) \cdot (mm)_i$. Assuming that net migrating married men age 45 and over do not have dependent children, the total number of net migrating children is given by

$$(1) \qquad M(ch) = d \sum_{(20-24)}^{(40-44)} M_i(m)(mm)_i$$

where d = 1.69, the summation extending over males in the age groups from 20-24 to 40-44.

Fifty percent of that net migrating children will be assumed male and 50 percent female.

$$(2) \qquad M(ch/m) = (.845) \sum_{20-24}^{40-44} M_i(m)(mm)_i$$

$$(3) \qquad M(ch/f) = (.845) \sum_{20-24}^{40-44} M_i(m)(mm)_i$$

where M(ch/m) and M(ch/f) denote net migrating male children and female children respectively.

The distribution of M(ch/m) and M(ch/f) over individual age groups may be made on proportional basis related to the number of unadjusted survived children in each age group for each sex.

Problem of Transfers from Low-Skill Male Category to High-Skill Male Category and Vice Versa

It is a reasonable assumption that members of the high-skill male category are perfectly mobile and that over the nation as a whole, there is no positive excess supply of this type of labor. Hence it is safe to assume that there are no transfers from the high-skill male category to low-skill male category. The crucial question is whether it is a reasonable assumption that over time there are no transfers from the low-skill male category to the high-skill male category. It may be argued that at the middle and older ages such transfers may not be very significant in relation to numbers involved. Hence it may, for practical purposes, be assumed that desired high-skill jobs in age group (i + 10) at time (t + 10) are the survivors (survival against mortality) of high-skill males aged i at time t for age groups 35-44, 45-54, 55-64 and 65+ at time (t + 10). But can it be said of young age groups that all low-skill males aged 20-24 at time t (say 1960) will survive as low-skill males aged 30-34 at time (t + 10) (say 1970)? It is considered that low-skill males in young age groups are subject to two causes of decrement viz. mortality and transfer to high-skill male category, besides the third cause of net migration. We are not aware of any studies directed towards the quantitative measurement of this cause of change by transfer from low-skill male category to high-skill male category by age. It is proposed to handle this question by an arbitrary adjustment on judgment basis.

The transfer from the low-skill male (lsm) category to the high-skill male (hsm) category is important for the younger age groups at the start of the decade and must be allowed for. In the case of these age groups, the proportion which high-skill males make of the total males at the start of the decade at age i is very significantly different from the corresponding proportion at the end of the decade at age i + 10. To adjust for this transfer in the case of age groups 14-15, 16-17, 18-19, 20-24, 25-29, and 30-34 at the end of the decade (t + 10), the total desired male jobs at time (t + 10) may be distributed in the two categories on the basis of the ratio of actual high-skill male jobs to total jobs in the same age group at time t.

V. RECOMMENDATIONS

We have used net migration response coefficients estimated by Mazek to carry out necessary calculations in the demonstration of the evaluation procedure developed in this study. The limitations of Mazek's estimates have been dealt with at relevant points in the preceding chapters. First, in the case of males, the estimates relate to all males and not to low-skill males only. Secondly, being based on 47 SMSA's in the northeastern quadrant of the United States, of which only nine had average annual unemployment rate greater than 7 percent during the period of the study, it may well be that Mazek's estimates of net migration response coefficient may not adequately reflect the full impact of back migration effects to which chronically depressed areas are subject. Hence the need for special separate adjustment for back migration may exist when evaluation procedure is applied to a depressed area. Finally, an important implication of Mazek's model is that the contribution of potential unemployment rate to the rate of net migration is zero only when the potential unemployment rate is zero. In a net migration model a correct assumption would be that the contribution of potential unemployment rate to net migration rate is zero when the potential unemployment rate for the group in the nation as a whole. This important underlying implication of the proposed model is its principal merit. The use of national employment participation rate as the desired employment participation rate ($\lambda_i^* = \lambda_i^N$) implies that net migration is zero when the actual employment participation rate of the group in the area is the same as that of the similar group in the nation as a whole. Hence, it is recommended that net migration response coefficients be estimated afresh for low-skill males and females by desired age groups viz. 14-15, 16-17, 18-19, 20-24, 25-29, 30-34, 35-44, 45-54 and 55-64 (at the end of the decade), using observation units (SMSA's for example) which may broadly be classified in the same category on the basis of areas' rates of unemployment, e.g. by high, average and low rates of unemployment.

In Chapter IV some of the problems which would arise in calculations pertaining to a future date have been discussed. In particular it may be felt that reasonably firm bases for projecting future employment by sector do not exist. In such a situation, it is recommended that net national benefit in a future year may be estimated on the assumption that the benefit is directly proportional to the net out-migration rate for the project area in the without-project situation. This procedure is obviously arbitrary but it may roughly be a good approximation in cases in which a seemingly more accurate procedure involves the use of projections which are highly conjectural.

REFERENCES

Battelle Memorial Institute, Columbus Laboratories. 1966. Final Report on a dynamic model of the economy of the Susquehanna River Basin.

Industrial Relations Research Institute. 1966. Retraining and migration as factors in regional economic development, (September), Madison, University of Wisconsin.

Jansen, E. F., Jr. 1966. Employment participation behavioral relationships. Unpublished Ph.D. thesis, Department of Economics, North Carolina State University, Raleigh, North Carolina.

Mazek, W. F. 1966. The efficacy of labor migration with special emphasis on depressed areas. Working Paper CWR 2, June. Institute for Urban and Regional Studies, Washington University, St. Louis, Missouri.

Somers, Gerald G. 1954. Labor supply for manufacturing in a coal area. Monthly Labor Review, (December), 1327-1330.

INCOME AND EDUCATION *

Introduction

One type of investigation, concerned with local government behavior, has attempted to estimate effects on education expenditures of a number of independent variables including income in an area. Another type of investigation which has proceeded separately has been concerned with the contribution of education to earnings. If both directions of causality exist, then studies which consider only one direction may overestimate effects. The association between education and income is attributed entirely to one direction of influence instead of recognizing that the association results from the combined action of two influences. The likelihood that there is a two way causality between income and education has been mentioned frequently, but there has apparently been no serious attempt to take account of the simultaneity.

Part 1 of this paper considers the determinants of local government expenditures on primary and secondary education, and presents results from traditional least-squares regressions with income as one of the independent variables. Part 2 concerns use of the same data in traditional least-squares regressions when education variables are among those used to explain income. Part 3 is concerned with the inter-relationship between education expenditures and income in light of the identification problem, namely that education influences income and income influences education. Estimates of the two influences are obtained from a simultaneous equations framework. The results are compared with traditional least-squares estimates. Part 4 is concerned with how to use the estimates in evaluating an action, e.g. building a water resource project, which affects education expenditures.

* F. Olson and G. S. Tolley, Dept. of Economics, The Univ. of Chicago.

1. Least-Squares Explanation of Education Expenditures

The traditional approach to explaining the relationship between education expenditures and income has been to use variations in income to explain variations in local education expenditures. This traditional approach might be called the short-run model. The effect of income on education expenditures is immediate as opposed to the process of education raising income which takes a longer time due to the lag between time of education and increased earnings.

These short-run models have ordinarily included several variables to explain education expenditures in addition to income. For comparison, findings for income, urbanization, density and intergovernmental transfers from previous studies are presented in Table 9. All these studies have used state data, except Weicher's which used a sample of Standard Metropolitan Areas.

As a first step in the present study, state data for 1960 were used in regressions of the kind reported in Table 10 including several further variables not used in previous studies. The results for arithmetic and logarithmic form are the first and third regressions presented in Table 11. These are denoted as type 1 regressions, indicating that the dependent variable is education expenditure and one of the independent variables is income. These will be the sole concern in the remainder of this section.

The type 1 regression equation is

(1) $E = a_1 + b_1Y + c_1U + d_1D + e_1N + f_1A + g_1P + r_1$

with the variables defined as follows:

Table 9. Regressions from previous studies with dollars of educational expenditure per pupil as the dependent variable

Independent variable	Fabricant	Fisher	Sacks and Harris, 1960a	Sacks and Harris, 1960b	Woo	Weicher
			Regression Coefficients			
Per capita income	.024**	.024*	.034**	.037**	.024***	.070***
Percent urbanization	-.02	-.25	-.14	-.152	-.183*	
Population density per square mile	-.01**	-.04*	-.05*	-.32**		.327***
Dollars of intergovernmental transfers per capita				.517*	.379	
			Multiple R^2			
	.59	.62	.60	.72	.77	.49

***Significant at the .99 level.

**Significant at the .95 level.

*Significant at the .90 level.

Table 10. Regressions using state data, 1960

Equation type	Coefficient	Arithmetic 1	Arithmetic 2	Logarithmic 1	Logarithmic 2	Logarithmic 3	Logarithmic 4
		Regression Coefficients					
Per pupil expenditures, E	b	D.V.	2.60***	D.V.	.432***	D.V.	--
Per capita income, Y	b	.063***	D.V.	.590***	D.V.	--	D.V.
Urbanization, U	c	-.599	1.24*	.015	.018	.078	.051
Population Density, D	d	.049	-.03	-.026*	.001	.029	.006
Percent nonwhite, N	e	-2.39**	-.83	-.042*	.026	.045	.000[b]
Intergovernmental transfers per capita, A	f	.454***	--	.052*	--	.082**	.055[a]
Pupil-population ratio, P	g	-1093.1**	--	-.689***	--	-.800***	-.161*
Years of schooling per adult, S	h	--	19.86**	--	.854***	.670*	1.121***
Capital per capita, K	k	--	-588.6	--	.023	.035	.059
Labor per capita, L	m	--	5790.3**	--	.304[a]	.131	.359
		Multiple R^2					
		.788	.866	.828	.872	.759	.828

D.V. Dependent variable.

*** Significant at .99 level.

** Significant at .95 level.

* Significant at .90 level.

[a] Significant at .80 level.

[b] rounds to zero.

E = per pupil education expenditures, 1959,

Y = median income of individuals 14 years or older, 1959,

U = percent of the population living in urban areas, 1960,

D = population per square mile, 1960,

N = percent of the population which was nonwhite, 1960,

A = intergovernmental transfers from higher government units to local governments in dollars per person, 1960,

P = ratio of average daily pupil attendance to population, 1960,

r_1 = random variable.

The source of all data is the 1960 Census of Population, except for E and A which are from the 1962 Census of Governments.

The units of the coefficients for the arithmetic type 1 regression in Table 11 are comparable to those in Table 10. Only the signs and significance levels for the type 1 logarithmic regression in Table 11 should be compared with the Table 10 results.

The results from Table 11 generally agree with those in Table 10. Consistent with previous studies, income is the most highly significant explanatory variable. While the sign of the urbanization coefficient is in doubt, it is not significant. A rationale for including urbanization and density is that these are associated with lower costs per pupil due to reduced transportation costs and larger classroom size. The one significant sign among the coefficients of these two variables in the type 1 regressions of Table 11 agrees with the negative signs found for the state analyses reported in Table 10. This is suggestive that the demand for education is price-inelastic in that lower costs reduce expenditure.[1/]

[1/] The one positive sign for population density in Table 1 pertains to a study using Standard Metropolitan Area data where the economies from density have probably been fully realized for every observation.

The positive significant coefficient for intergovernmental transfers as found in both tables is expected since transfers augment resources available to communities. However, because many intergovernmental transfers are from state to local level, this variable does not measure net additions to availabilities for states (the jurisdictions whose education expenditures are being explained in the present analysis). To some extent, the intergovernmental transfers are simply a mechanism by which decisions to make expenditures are implemented rather than having independent causal significance. The transfers undoubtedly depend partly on income and may be taking variation away from income. The variable may also tend to be a catch-all reflecting effects of other shifters of expenditures not included in the regressions.

Two variables remain for which comparisons are not available from Table 1. These are both statistically significant. The negative sign for the nonwhite coefficient may reflect a taste for discrimination of the white population combined with lack of political power of nonwhites. It could also reflect a lower demand for education by nonwhites due to job discrimination resulting in lower rates of return to education for nonwhites. The negative coefficient of the pupil-population ratio indicates the extent to which expenditures per pupil are reduced with increases in the total number of pupils that must be supported by a given population.

2. Least-Squares Estimates of Income

The more recent interest in the relationship between education and income is concerned with what might be called the long-run approach. While over short periods of time income may determine the levels of local education expenditure, over longer time periods education is viewed as one of the determinants of income. As Schultz, Denison, Becker and others have viewed it, at least a part of education expenditures is investment in human capital with an associated return in the form of future income.

The two basic types of analysis have been the macro-approach of Denison and Schultz, who have attempted to estimate that portion of the increase in national income which is attributable to education expenditures and the micro-approach of Becker, Hannoch and also Schultz who have tried to estimate the increase in individual income attributable to education expenditures. Between these two approaches is the analysis of Welch which is a disaggregated approach attempting to distinguish between quantity and quality in schooling.

The approach followed here is similar to that of Welch, whose empirical focus on rural farm males by detailed age and years of schooling groups is quite different from the present study. The relation explaining income per person is the result of two underlying relations, one of which is an aggregate production function:

$$Y = Y(K, L, Q).$$

On the right-hand side are the per capita inputs of nonhuman capital K, raw labor L as reflected in number of persons in the labor force, and human capital due to education Q. The latter variable is in turn a function of the number of years of schooling and factors explaining the quality of the schooling:

$$Q = Q(S, E, U, D, N)$$

where S is years of schooling per capita. Expenditure per pupil E affects quality since, other variables unchanged, this variable indicates inputs applied to the given number of years of schooling. Urbanization U and population density D are cost-lowering factors that are expected to be positively associated with amount of quality that can be attained with a given level of expenditure. The proportion of the population nonwhite reflects the effects of discrimination on education quality.

A regression equation is obtained by substituting the expression for the quality of schooling into the production function assuming a linear or linear logarithmic approximation:

(2) $Y = a_2 + b_2E + c_2U + d_2D + e_2N + h_2S + k_2K + m_2L + r_2.$

Y, E, D, U and N were defined in connection with equation (1). The remaining variables are:

S = median years of schools completed by population 25 years or older, multiplied by number of persons who are 25 years or older, divided by population 1960,

K = value of manufacturing property divided by population, 1960,

L = labor force divided by population, 1960,

r_2 = random variable.

The sources for S and L are the 1960 Census of Population, and the source for K is the 1962 Census of Government.

Least-squares results are given in the second and fourth columns of Table 11. These are called type 2 regressions denoting dependent variable is income and one of the independent variables is education expenditures. The arithmetic and logarithmic forms give similar results, and most of the signs are as expected. The negative sign for capital per capita in the arithmetic version and the positive sign for proportion nonwhite in the logarithmic version are the only two signs not expected, and these coefficients are not significant.

While years of schooling and labor per capita have significant coefficients, the most important explanatory variable is per pupil expenditures. A basic question of this study pertains to direction of causality in the relationship between income and per pupil expenditures. Because of the probability of mutual causality, the coefficient indicating effects of education expenditures on income obtained by direct estimation of equation (2) may have an upward bias. The association between education expenditures and income that was attributed entirely to income in the type 1 regressions is now attributed entirely to education expenditures in the type 2 regressions.

One way to attempt to avoid this bias in explaining income is to use lagged data for education. The income of the present working population may be viewed as a function of the amount of their education as determined by educational inputs at the time of their schooling. To carry out this approach the 1960 population of each state was divided into 10-year age intervals. It was assumed that the expenditure per pupil applicable to each adult in 1960 was the deflated value at the time he was 5-15 years of age. The wholesale price index was used to convert previous years to 1957-59 values. The weighted average of these deflated values was used in place of current expenditures per pupil. This attempts to correct for the fact that the unlagged variable overstates amount of schooling by failing to take account of the lower education expenditures per pupil prevailing when the present adult population was in school. In both the arithmetic and logarithmic forms based, this approach gave results very similar to columns 2 and 4 of Table 11 already discussed.

The similarity of the regressions reflects the high correlation between lagged and unlagged expenditures. Those states which had high education expenditures per pupil in 1960 also had high expenditures relative to other states in the past. Since these high expenditure states tended to be relatively higher income states in the past, the present higher education expenditures may well have been caused by the higher income. Thus because of their correlation between present and past values, the use of lagged data apparently does not solve the problem of the direction of causality.

3. Simultaneous Estimation

The bias in regression parameters from the approach in the preceding sections is traditional least-squares bias encountered if a single equation estimation approach is applied to a situation where the variables are determined by simultaneous interaction. An approach to avoid this problem is to estimate the reduced form of the simultaneous equations and to use combinations of the reduced form coefficients to obtain unbiased estimates of the parameters.

To restate the behavioral hypothesis, the relationship of income and education may be viewed as forming a two-equation system explaining how education expenditure and incomes are mutually determined:

(1) $E = a_1 + b_1Y + c_1U + d_1D + e_1N + f_1A + g_1P + r_1$

(2) $Y = a_2 + b_2E + c_2U + d_2D + e_2N + h_2S + k_2K + m_2L + r_2.$

Solving (1) and (2) for the two endogenous variables E and Y in terms of the remaining variables, gives the reduced form equations:

(3) $E = a_3 + c_3U + d_3D + e_3W + h_3S + f_3A + g_3P + k_3K + m_3L + r_3$

(4) $Y = a_4 + c_4U + d_4D + e_4W + h_4S + f_4A + g_4P + k_4K + m_4L + r_4$

where

$$c_3 = \frac{c_1 + b_1c_2}{1 - b_1b_2}, \qquad c_4 = \frac{c_2 + b_2c_1}{1 - b_1b_2},$$

$$d_3 = \frac{d_1 + b_1d_2}{1 - b_1b_2}, \qquad d_4 = \frac{d_2 + b_2d_1}{1 - b_1b_2},$$

$$e_3 = \frac{e_1 + b_1e_2}{1 - b_1b_2}, \qquad e_4 = \frac{e_2 + b_2e_1}{1 - b_1b_2},$$

$$f_3 = \frac{f_1}{1 - b_1b_2}, \qquad f_4 = \frac{b_2f_1}{1 - b_1b_2},$$

$$g_3 = \frac{g_1}{1 - b_1b_2}, \qquad g_4 = \frac{b_2g_1}{1 - b_1b_2},$$

$$h_3 = \frac{b_1h_2}{1 - b_1b_2}, \qquad h_4 = \frac{h_2}{1 - b_1b_2},$$

$$k_3 = \frac{b_1k_2}{1 - b_1b_2}, \qquad k_4 = \frac{k_2}{1 - b_1b_2},$$

$$m_3 = \frac{b_1m_2}{1 - b_1b_2}, \qquad m_4 = \frac{m_2}{1 - b_1b_2}.$$

If equations (3) and (4) are estimated by least-squares, the resulting regression coefficients are free of the type of bias encountered in direct least-squares estimation of (1) or (2).

The denominator $1 - b_1 b_2$ of each reduced form coefficient reflects the fact that any change affecting the system has a magnified impact due to the reinforcing interaction of education and income. In considering the numerators, note that a shifter such as urbanization affects education directly because it is a cost factor ($c_1 \neq 0$), and the direct effect is reinforced by an indirect effect operating through b_1 because urbanization raised income ($c_2 \neq 0$).[1/] Conversely, the direct effect of urbanization on income via c_2 is reinforced by the indirect effect through b_2 because of higher education expenditures. The sum of direct and indirect effects can be seen in the numerators of the reduced form coefficients c_3 and c_4, but from this structure there is no way to disentangle the direct and indirect effects. The same remarks apply to D and N, the other variables which appear in both equations (1) and (2).

Now consider years of schooling S which has a direct effect on income but not on education. The numerator of h_4 reflects only the direct effect of S on income as there is no indirect effect of S operating through education expenditures since S is excluded from equation (1). For the same reason, the numerator of h_3 reflects only the indirect effect on education expenditures of S acting through its effect on income. The only reason that h_3 and h_4 differ is that the effect of S on education must act through the marginal propensity to spend on education b_1. For instance, if the marginal propensity to spend on education were one-half, the effect of S on education would be one-half its effect on income.

[1/] Urbanization was included in equation (2) due to its effect on the quantity of education, but it may also be correlated with income because it stands in for other things affecting income, e.g. higher wages for comparable labor due to rural-urban labor immobilities. This does not affect the analysis of the present study.

As can be seen from the formulas, dividing h_3 by h_4 gives an estimate of b_1. Since K and L are also excluded from equation (1), additional estimates of b_1 can be obtained from the ratios of their reduced form coefficients.

The same type of reasoning applies to the pupil-population ratio P and intergovernmental transfers A, both of which are excluded from equation (2). In the reduced form equations, the difference in their effects is due to the fact that they have a direct effect only on education expenditures and influence income only as through their effect on education expenditures. The ratios g_4/g_3 and f_4/f_3 thus give estimates of b_2.

A type 3 regression has per pupil expenditures as the dependent variable and excludes per capita income, giving an estimate of the reduced form equation (3). The dependent variable in a type 4 regression is income, with per pupil expenditures excluded, giving an estimate of the reduced form equation (4). The coefficients from the logarithmic forms are shown in the last two columns of Table 11. Because of elimination of the association between income and education expenditures in the reduced form regressions, the multiple R^2's are reduced but not drastically. The signs in both equations agree, and the significant variables in both are years of schooling, pupil-population ratio and intergovernmental transfers.

Given b_1 and b_2, the remaining structural parameters can also be estimated from the reduced form coefficients. For instance, from the expressions that were given for the reduced form coefficients from equations (3) and (4) in terms of the parameters of the original system, it can be seen that multiplying g_4 by b_1 and subtracting the result from g_3 gives as estimate of the effect of pupil population ratio on education, g_1. The first column of Table 12 gives the formula for each structural coefficient in terms of the reduced form coefficients, and the second column applies these formulas using the reduced form estimates from the last two columns of Table 11. Thus b_1 is .670 divided by 1.121, b_2 is -.161 divided by -.800, and so on.

Two estimates of b_1 are available which are not given in Table 12. namely k_3/k_4 (= .59) and m_3/m_4 (= .37). These are not too disparate from the value of .60 given in the table but are not preferred because they are based on coefficients which are not significant.

An additional estimate of b_2 is available from the reduced form coefficients of intergovernmental transfers, f_4/f_3 (= .66). This is more than three times the estimate based on the reduced form schooling coefficients and is even half again as large as the direct least squares estimates. Problems connected with the intergovernmental transfer variable were discussed in presenting the type 1 regressions, the basic difficulty being that in analysis of state data intergovernmental transfers may be partly explained by income rather than being solely an explainer of education. This would impart an upward bias to the intergovernmental transfer coefficient in the type 4 regression where income is the dependent variable, leading to a similar upward bias in the estimate of b_2. Therefore, the estimate of b_2 (= .20) given in Table 12 is strongly preferred.

The test for whether a ratio differs significantly from zero is the same as the test of whether its numerator differs from zero, so

Table 12. Structural parameters estimated from state data logarithmic regressions

Elasticity	Formula in terms of reduced form coefficients	Estimate implied by reduced form regressions	Direct least-squares estimate
Effect of income on education, b_1	h_3/h_4	.60*	.59***
Effect of education on income, b_2	g_4/g_3	.20*	.43***
Effect of urbanization			
on education, c_1	$c_3 - b_1c_4$	.06	.02
on income, c_2	$- b_2c_3 + c_4$	.02	.02
Effect of population density			
on education, d_1	$d_3 - b_1d_4$	.04	-.03
on income, d_2	$- b_2d_3 + d_4$	.01	.00[b]
Effect of percent population nonwhite			
on education, e_1	$e_3 - b_1e_4$	-.06	-.04*
on income, e_2	$- b_2e_3 + e_4$	.03	.03
Effect of intergovernmental transfers on education, f_1	$f_3 - b_1f_4$	-.06	.05**
Effect of pupil-population ratio on education, g_1	$g_3 - b_1g_4$	-.70	-.68***
Effect of quantity of schooling on income, h_2	$- b_2h_3 + h_4$	.68	.85***
Effect of capital per capita on income, k_2	$- b_2k_3 + k_4$	.04	.02
Effect of labor per capita on income, m_2	$- b_2m_3 + m_4$	.27	.30[a]

***Significant at .99 level.

**Significant at .95 level.

*Significant at .90 level.

[a]Significant at .80 level.

[b]Rounds to zero.

that the significance of the estimate of b_1 from the reduced form regression is the same as the significance of h_3. Similarly, the significance of the reduced form estimate of b_2 is the same as the significance of g_4. None of the other estimates in the column are based on significant coefficients. However, the estimates are generally reasonable, particularly for the larger parameters g_1, h_2 and m_2.

To apply two-stage least-squares estimation to equations (1) and (2), instead of using the observed values of the second endogenous variable [Y in equation (1) or E in equation (2)], purged values are used which are the values of the variable predicted from reduced form regression. This gives results which are essentially the same as the reduced form estimates of the parameters presented in Table 3.

The estimate of effect of income on education, b_1, implied by the reduced form regression is practically identical to the direct least-squares estimates. On the other hand, the estimate of effect of education on income, b_2, from the direct least-squares estimate is about twice the estimate from the reduced forms.

The least-squares bias results from correlation of an independent variable with the random variable. In equation (1), Y is correlated with R_1; and in equation (2), E is correlated with R_2. The higher are these correlations the greater will be the biases in the estimation of the coefficients of these variables. Y is correlated with R_1 in equation (1) because a shift in R_1 leads to higher education expenditures which in the other equation leads to higher income. Similarly, E is correlated with R_2 in equation (2) because R_2 leads to a higher income which leads to greater education. The magnitude of the indirect effect depends on how important the residual shifts in one equation are as a contributor to the total variation in the variable being expressed as the dependent variables in the other equation. The residual shifts in equation (1) operate in equation (2) through expenditures per pupil, which accounts for a relatively small proportion of total variation in income. Thus one expects a relatively low correlation between Y and R_1 with consequence small least-squares bias in b_1. However, income is the predominant variable explaining education so that R_2 through affecting income explains a high percentage of the variation in E leading to a large correlation between E and R_2 in equation (2) and hence a large least-squares bias in the estimate of b_2.

4. Value of Education Investments Induced by Regional Growth

The education of people living in Appalachia and other low-income areas has been a topic of much concern. The oft-cited fact, that per pupil expenditures tend to be low in these areas, is consistent with the preceding sections of this paper in which income was found consistently to be the most important variable explaining education expenditures.

Are the low expenditures symptomatic of malallocation in educational investment? If so, will actions aimed at increasing incomes in depressed areas lead to benefits from education? Such benefits have not been included in measured benefits from projects (e.g. dams, roads) aimed at promoting regional development.

One of four income streams is project costs. For a federal resource development project these are borne partly by taxpayers at large and partly by persons in the vicinity of the project insofar as they are required to repay costs. Another income change is the national income benefits resulting from the project. These include project outputs, such as increased production due to flood control or the value of road use. The benefits also include employing resources more productively than in the absence of the project, either through employing the unemployed or through employing persons so as to increase their marginal product.

Two other income streams are exactly offsetting. These are due to spatial redistribution of activities. The increased output at the project site is likely to lead to more people living there, with some of their demands partly localized as for trade and services. To fulfill these demands, resources will be used which would otherwise find employment in the rest of the economy.

The nomenclature used in benefit-cost analyses varies, but the four income streams that have been mentioned tend to coincide with: (1) primary costs which are costs of constructing and maintaining the project, (2) primary benefits which are national income benefits of the project, (3) secondary benefits which are increases in income presumably at or near the project exactly offset by decreases in other areas, and (4) secondary costs which are the decreases just mentioned.

If the marginal propensity to spend on education were the same for all areas and the rate of return on education expenditures were the same as rate of return on alternative investments, there would be no net benefits connected with changes in education expenditures. The discount rate used for the future benefits would make the present value of the education just equal to the costs, so that the benefits of any education expenditures induced by the project would be achieved only by foregoing an equal amount of costs. Suppose, however, that the rate of return on education is higher than alternative rates of return due to such factors as lack of knowledge on the part of those deciding on public school expenditure and to impediments in equating marginal education benefits to costs in view of financial and taxing arrangements. Then every dollar of education expenditure induced by the project would be a cost from which would be obtained a future return greater than the cost.

Let B_t be the value in year t of one of the four income streams. Let the marginal propensity to spend from this income on education of pupils in the n^{th} grade be m_n so that the marginal propensity to spend on education is M $(= \sum_{n=1}^{n=N} m_n)$. The spending will lead to a stream of later returns $x_{nt'}$, where t' is the number of years after the expenditure that returns accrue. Positive values of $x_{nt'}$ occur from the year that increased earnings from the education begin to accrue until the end of the time when earnings are affected. The value in year t for each dollar of the spending on the n^{th} grade is $p_n = -1 + \Sigma x_{nt'}/(1 + i)^{t'}$ where i is the rate of discount. The total value of the induced education connected with the j^{th} income stream is then $B_t \; \Sigma m_n p_n$.

There is little information available on the p_n returns from increased expenditure different grades of education. The p_n values for different grades are undoubtedly related. For instance, an increase in investment in one grade alone probably would have less effect than if associated with increased investments in other grades. The estimation of returns to expenditures on different grades is further complicated by the fact that these may influence drop-out rates. The return to not dropping out might be estimated, but it would be difficult to estimate how expenditures influence the drop-out rate.

Studies are available indicating returns to investing in education in ways other than increased spending on various grade levels. Most notably effort has been diverted to estimating returns to completing different number of years of schooling [1], [7], [8]. The studies of returns to education uniformly indicate rates of return above 10 percent. One way to proceed is to assume these rates of return are applicable to the type of increased spending on education under consideration here. Then a problem is to find the value of the spending in view of the rate of discount i and the lag between the spending and the increased earnings. If the increased spending is spread uniformly over the various grades of school, then the average lag before returns begin to accrue is on the order of five or six years, depending on how long students stay in school. The present value of a dollar invested in education is $-1 + \sum_{t=L}^{t=L+40} X_t/(1+i)^t$ where X_t is increased earnings in year t due to the education, L is the lag before increased earnings begin (presumably at least five or six years on the average) and post-education earnings are assumed to endure for 40 years. Suppose the general shape of the time stream remains the same with different types of educational investments. Past studies then indicate a set of ratios $R_{tj} = X_t/X_{t+j}$ between increased earnings in different time periods.[1/] The internal rate r must be such that $-1 + \sum_{t=L}^{t=L+40} X_t/(1+i)^t$ is zero, or $-1 + X_L \sum_{t=L}^{t=L+40} R_{Lt}/(1+r)^t$ is

[1/] There are 39 independent ratios. If for convenience one works with the first year, the ratios are the 39 R_{1t}'s$(t = 2 \ldots 40)$.

zero. Given r and the R_{Lt}'s from previous studies one can solve this expression for X_L and making use of $X_t = R_{Lt}X_L$ solve for the remaining X_t's. If the increase in earnings due to the education investment is the same in every year (all R_{Lt}'s equal to 1), X ($=X_L$) is found from the expression for internal rate of return to be $1/[\sum_{t=L}^{t=L+40} 1/(1 + r)^t]$ in which case the expression for present value of a dollar invested in education becomes $-1 + [\sum_{t=L}^{t=L+40} 1/(1 + i)^t]/[\sum_{t=L}^{t=L+40} 1/(1 + r)^t]$.

To this point, the concern has been with the value of the induced education expenditures in one year t. That value would need to be discounted and summed over all t periods to obtain present value. A more general approach not based on the average lag assumption is to build in the time stream of expenditures leading to the future returns. The present value summing over all years is obtained in the process of doing this. This present value of the education expenditures considering all years is $\sum_{t=1}^{t=H} \sum_{n=1}^{n=N} B_t m_n p_n/(1 + i)^t$, where H is the time horizon. For federal projects H has typically been assumed to be 50 or 100 years. The present value can be rewritten as

$$\sum_{u=0}^{u=H-N} \sum_{t=1}^{t=N} B_{t+u} m_t p_t/(1+i)^{t+u} + \sum_{u=1}^{u=N} \sum_{t=1}^{t=N-u} B_t m_{g+u} p_{t+u}/(1+i)^t + \sum_{u=H-N+1}^{u=H-1} \sum_{t=1}^{t=H-u} B_{t+u} m_t p_t/(1+i$$

The first of the three terms in this expression is for returns from streams of costs incurred through all grades of training. The second term pertains to the streams of costs just after the project is completed when expenditures are increased for children who have already had some education. The third term pertains to stream of costs that have begun to be incurred but are not completed by the time the horizon is reached. These are for children who begin their education within N years of the time horizon. The third term is likely to be small (it will be zero if H is infinite), and the second term will require knowledge of returns to a stream of education expenditures that were increased in the middle of the training period. Most likely these would be estimated according to a proportionality assumption given the estimated returns to completed education streams at the higher level of expenditures. Here attention will center on the completed streams comprising the first term. If the benefits are the same in different years as is the usual assumption in working with average annual benefits (B_{t+u} the same for all years), the first term becomes

$$\left[\sum_{t=0}^{t=H-N} 1/(1+i)^t\right][B]\left[\sum_{t=1}^{t=N} m_t p_t/(1+i)^t\right].$$

B times the bracket containing the $m_t p_t$ terms is the value of one completed stream of education expenditures. The total value resulting from these streams is analogous to the present value of an annuity beginning in the present and being paid each year through H-N. Multiplying by the other brackets gives the present value of such an annuity.

The p_t's are returns to spending on each grade and as already mentioned are not likely to be independent of one another. The increase in expenditures will be distributed over all grades, and a future return will result. To allocate the return among the grades, practically, would be arbitrary. An advantage of the present approach not assuming an average lag is that one does not need to be concerned with the p_t's. The value of a completed stream is:

$$\sum_{t=0}^{t=N} m_t p_t/(1+i)^t = -\sum_{t=0}^{t=N} m_t/(1+i)^t + \sum_{t=N+1}^{t=N+41} Y_t/(1+i)^t.$$

The Y_t's are the future returns to the stream of education expenditures and are similar to the X_t's considered earlier, except that the Y_t's are returns resulting from the m_t's whereas the X_t's are returns resulting from a one dollar increase acting with an average lag L. If the time shape of the returns were assumed identical (the same R_{tj}'s applying to the Y's as to the X's) and if the internal rate of return were the same, each Y_t would be an exact multiple of X_t. Given the rate of return, the time shape of returns, and the marginal propensities to spend on each grade, the condition must be satisfied that

$$0 = -\sum_{t=0}^{t=N} m_t/(1+r)^t + Y_{N+1} \sum_{t=N+1}^{t=N+41} R_{1t}/(1+r)^t,$$

which can be solved for Y_{N+1}. Using the R's, this can then be used to find the present value of the education stream in the expression given preceding this one. Assume that expenditures are increased equally on all grades (m_t = M[1/N]) and that the effect on future incomes is equal in all years (all R's equal to 1). Then the present value of one of the education streams is

$$M\{-\frac{1}{N}\sum_{t=1}^{t=N} 1/(1+i)^t + \frac{1}{N}[\sum_{t=1}^{t=N} 1/(1+r)^t][\frac{\sum_{t=N+1}^{t=N+41} 1/(1+i)^t}{\sum_{t=N+1}^{t=N+41} 1/(1+r)^t}]\}.$$

The value of a dollar spent on education implied by this expression is similar to that when the same assumptions were made using the average lag approach.

More complicated time stream assumptions could be used in estimating the Y's in the general formula above. For instance, it has been contended that education increases returns to on-the-job training affecting the time at which returns accrue to the initial education [2]. A greater deferral of benefits could raise the income consistent with a given rate of return and might not much affect present value.

Two further questions might be raised about the procedure for estimating the value of a dollar spent on education. First, does it not vary with the area where the spending occurs? Second, can one obtain estimates of the value of spending on education from the behavioral estimates of the first part of this paper?

The preferred estimate of b_2 from the earlier sections of the paper indicates that a 1 percent change in education expenditures will lead to a .2 percent increase in income. The results imply complementarity between educational expenditures and other factors contributing to income. It might therefore be thought that a dollar spent in a high income area returns more than in a lower income area. The marginal return from a dollar spent on education is the elasticity of income with respect to education (b_2) times the ratio of income to education expenditures. Because the demand for education is income inelastic, this ratio tends to be higher for high income states. At the U.S. average ratio of median income to expenditures per pupil, income rises $2.50 for every dollar increase in spending on education. For the low income state of Mississippi, the figure is $1.82, and for the high income state of California the figure is $2.60. Suppose we view these figures as reflecting a process whereby the $1 of education expenditures is part of a 12-year stream of spending on a student while he is in school, whereas the increase in income is part of a 40-year stream of earnings while the person is in the labor force. Then the rate of return implied by these figures is such that

$$- \sum_{t=1}^{t=12} \frac{1}{(1+r)^t} + \sum_{t=13}^{t=53} \frac{b_2 \; E/Y}{(1+r)^t} = 0.$$

At the U.S. average, the implied internal rate of return is 10.9 percent. For Mississippi the implied rate of return is 8.8 percent, and for California it is 11.3 percent. One reason these figures may give underestimates is that the present adults are earning as a result of the lower real education expenditures that prevailed when they were being educated. To use the present expenditures per pupil in calculating the ratio Y/E tends to make the ratio too low.

Even if the rate of return calculated by this method could be based on accurate measurement of the level of education expenditures, the question remains, what is the effect on income of an increase in education expenditures under conditions of factor mobility? The estimate of b_2 shows how education contributes to income in an explanation of income based on existing factor proportions in each state. To use the state average ratio E/Y would imply people being educated in each state would as adults work at the factor ratios now existing in those states. People now being educated in low income states may migrate out to high income states, and nonhuman capital may migrate into low income states. These types of factor movements are now going on. Furthermore, increasing education may speed them up. More highly educated people are more mobile, and a better educated labor force is known to be an attractor of capital. Rather than attempt to differentiate returns to be expected in different localities, a better judgment might be to assume that all additional education expenditures will have the same rate of return regardless of where made. If the estimates of this paper were being used, one might take the estimates for the U.S. average.

Suppose the value of a dollar of spending on education has been established from previous studies or from the estimates earlier in this paper, and consider using the regression results from the first part of this paper to estimate the marginal propensity to spend on education. The regressions indicate that a logarithmic form is preferred to arithmetic form and that a good estimate of the elasticity of spending on education with respect to income is .6. The elasticity is the marginal propensity to spend times expenditures divided by income. Assume that primary benefits and secondary benefits accrue in the project area and that primary costs and secondary costs accrue in areas which on net have national average per capita income and per pupil expenditures. Then, to estimate the induced change in spending on education, primary and secondary benefits would be multiplied by .6 times (E_i/Y_i) and primary and secondary costs would be multiplied by .6 times (E_{US}/Y_{US}), where i refers to project area and US refers to U.S. average. These magnitudes would be multiplied by the value of a dollar spent on education to arrive at the net education benefits. Because E/Y tends to be higher than the national average for low income areas, building a project in a low income area will be conducive to obtaining high education benefits.

REFERENCES

1. BECKER, GARY S. Human Capital. New York: Columbia University Press, 1965.

2. BOWMAN, MARY JEAN. "The Assessment of Human Investments as Growth Strategy," (forthcoming in Joint Economic Committee of the U.S. Congress; Compendium on Human Resources) February, 1968.

3. FABRICANT, SOLOMON. The Trend in Government Activity Since 1900. New York: National Bureau of Economic Research, 1952.

4. FISHER, GLENN W. "Interstate Variation in State and Local Government Expenditure," National Tax Journal, XVII (March, 1964).

5. KEE, WOO SIK. "Central City Expenditures and Metropolitan Areas," National Tax Journal, XVIII (December, 1965).

6. SACKS, SEYMOUR and HARRIS, ROBERT. "The Determinants of State and Local Government Expenditures and Intergovernmental Flows of Funds," National Tax Journal, XVII (March, 1964).

7. SCHULTZ, THEODORE W. Economic Value of Education. New York: Columbia University Press, 1963.

8. _________. "Resources for Higher Education: An Economist's View," Investment in Human Capital Paper No. 68:1 (unpublished), Department of Economics, The University of Chicago, (February, 1968).

9. U.S. BUREAU OF THE CENSUS, 1962 Census of Government, Vols. 2 and 4 (Chapter C). Washington: U.S. Government Printing Office, 1964.

10. _________. U.S. Census of Population: 1960, Vol. 1, Chapters A, B, and C. Washington: U.S. Government Printing Office, 1962.

11. U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE. Statistics of State School Systems: 1959-60. (Office of Education). Washington: U.S. Government Printing Office, 1962.

12. WEICHER, JOHN. "Urban Renewal and Municipal Expenditures." Unpublished Ph.D. Dissertation, Department of Economics, The University of Chicago, 1967.

13. WELCH, FINIS. "The Determinants of the Return to Schooling in Rural Farm Areas: 1959." Unpublished Ph.D. Dissertation, Department of Economics, University of Chicago, 1966.

Refinement of Regional Employment Multiplier Estimates *

I. Introduction

The main purpose of the research project on hand is to develop and apply techniques for making improved estimates of the magnitude and area distribution of secondary or indirect benefits, usable in Corps' project planning and evaluation.

In Corps project, as well as in any planning for regional development, secondary benefits are an important factor to be taken under consideration, especially for depressed areas, because they are indirect returns added as a result of developing, maintaining, and operating a project. They indicate the effect on location of induced activities such as retailing and attraction of related industrial activities in addition to that directly attributable to a project.

Indirect project effects or multiplier effects might be estimated by using input-output analysis for each project. However, the accuracy of the estimates obtained from input-output model remains in doubt.

As a step towards improving estimates of regional multipliers, a study was completed in the earlier stage of the present research work covering the investigation of the input-output model, the theoretical and empirical problems which it ignores and which should be taken into account to obtain accurate estimates, and the manner in which that model should be used as a tool for regional analysis.[1] This study reveals the fact that the immediate problems involved in the estimates are connected with the definition of a sector or a region (i.e., industrial composition of a sector and

---

* A. S. Daghestani, San Angelo State College, San Angelo, Texas, and G. S. Tolley, University of Chicago, Chicago, Illinois.

[1] G. S. Tolley and F. M. Goode, "The Input-Output Model As a Tool For Regional Analysis." North Carolina State University. Unpublished paper.

its size), and with the accuracy of data, both of which may be reflected in changes in coefficients and, consequently, errors in the multipliers. A second undertaking towards improving estimates of regional multipliers is, therefore, to analyze mathematically and statistically the sensitivity of multiplier estimates to errors in the data and to errors introduced by aggregation of sub-industries into larger industrial groups. Such analysis was initiated at the conceptual level in the earlier part of the research work and, as a result, a separate paper[2] has been written on preliminary thinking on the general problem of how to investigate the effects on multipliers of errors in coefficients due to inaccuracy of data and to aggregation.

The present study, which is intended to indicate the work completed and the progress made on the project, attempts to deal with the conceptual analysis and empirical investigation developed regarding effects of coefficient errors on multipliers. For the purpose of providing an integrated study, part of the earlier work done in this respect will be repeated here.

---

[2] F. M. Goode and G. S. Tolley, "Refinement of Regional Multiplier Estimates," paper given at Kerr Lake Conference, (N.C.) February, 1965.

I. Causes of Differences in Regional Multiplies Values

Basically, estimated values of regional or area multipliers depend on various factors among which are the type of multiplier analysis, techniques and assumptions employed in deriving them, and component economic structure of a region or an area under study. Hence, some of the main reasons for variation in regional or area multipliers can be revealed through examining some of the area empirical studies made and the methods they employed in deriving these multipliers.

One type of regional multiplier analysis which has been used by regional analysts is that associated with economic base studies. Most of these studies deal with aggregate economic activities, that is, with a gross industrial classification, in a given region or an area, and, consequently, derive a single average regional or area multiplier.

The economic base type of analysis makes a distinction between export-oriented (basic or primary) activities and locally-oriented (non-base or service) activities. This distinction is based on the premise that economic existence and growth of a region, whatever size it may be, are caused by its export-oriented activities on which locally-oriented activities themselves are dependent. Accordingly, the economic base of a region is the group of industries which produce locally but sell their productions out of the borders of that region, i.e. economic base of a region or an area is made up of its export-oriented activities.

An empirical economic base-regional multiplier is determined by observing the historical relationship between changes in export-oriented activity and changes in total economic activity or between changes in the former and changes in locally-oriented activity in the region or area under study. In the latter case, the multiplier is obtained by adding unity to the ratio of export-oriented to locally-oriented activity. The multiplier so derived is a single average regional (or area) multiplier which is used to estimate a region's economic activity increases resulting from increases in its export-oriented activity.

Application of economic base approach to derive a regional multiplier, thus, requires separate data on each of a region's export-oriented activity and locally-oriented activity for each industry included. However, since such data are not readily available for most regions or areas, different methods have been used by different analysts to divide total economic activity of each industry in a region or an area under study into export-oriented and locally-oriented components. It is in such differences in methods where one of the main reasons for variation in regional multiplier lies.

With regard to industrial and commercial firms, the method followed by most studies has been to separate them into those which are mixed, those which are exclusively export-oriented, and those which are exclusively locally-oriented. To estimate export-oriented and locally-oriented categories of mixed industries (or sectors), different techniques have been employed among which are some form of location quotient and the empirical firm-by-firm method.[1]

---

[1]W. Isard and Associates. "Methods of Regional Analysis An Introduction to Regional Science," (4th Printing, May, 1966). Cambridge: Massachusetts Institute of Technology Press, 1966, pp. 195-197.

Still other factors contributing to differences in regional (or area) multiplier are revealed from several empirical economic base type of multiplier studies examined for the purpose of this study, some of which are discussed in the following sections. Such factors, as well as that stated above, will be explained later when a comparison is made between the studies examined below.

One of the empirical studies dealing with economic base-regional multiplier is that which was conducted by G. E. Thompson[1] to estimate employment multiplier in Lancaster County, Nebraska. His main concern was with the relationship of changes in export-oriented employment to changes in total employment in that county. Income multiplier in the county, however, was not measured because the data necessary to derive it was not available. The base year he selected for his study was 1950. The market orientation of each of the industries in the county was determined by using location quotient (or concentration ratio) method.[2] For this purpose, four market areas were assumed to exist, namely, Lancaster County itself, the thirteen counties of Southeast Nebraska, the state of Nebraska, and the U.S. The county was considered as a subject economy. The other three areas were used as benchmark economies, i.e. economies with which the subject economy is being compared. Then, the market

---

[1] G. E. Thompson. "An Investigation of the Local Employment Multiplier," The Review of Economics and Statistics, vol. XLI(February, 1959), pp. 61-67.

[2] Location quotient was defined in this study as "a ratio of the employment in a given industry as a percent of total employment in that economy to employment in the same industry as a percent of total employment in another economy." Ibid., pp. 62.

areas which are served by industries in the county were estimated by comparing the concentration of employment by industry of these economic areas, i.e. by computing three location quotients. First, Lancaster County was used as a subject economy and the U.S. as a benchmark economy:

$$\frac{\text{Lancaster County}}{\text{U.S.}} - \text{Lancaster County} = q_1.$$

Secondly, the state of Nebraska was considered as a benchmark:

$$\frac{\text{Lancaster County}}{\text{Nebraska}} - \text{Lancaster County} = q_2.$$

Finally, Southeast Nebraska was used as a benchmark:

$$\frac{\text{Lancaster County}}{\text{Southeast Nebraska}} - \text{Lancaster County} = q_3.$$

The values of these quotients were calculated for each of the industries using the data of the base year 1950. Such values were assumed to indicate the market area which supported any specialization that existed in the county. In those industries in which q_1 was the largest, the county was considered to be serving the U.S. market and was assumed in relation to the U.S. economy as the benchmark in determining the percentage of export-oriented employment. Where q_3 was the largest, Southeast Nebraska was considered as the benchmark, and where q_2 showed to be the highest in value, the state of Nebraska was assumed as a benchmark.

To obtain the percentage of export-oriented employment in each of the industries, Thompson used a different method "from that employed by Hildebrand, Mace, and others."[1] The method used in this study included the estimation of a "specialization ratio" for each industry, i.e.

---

[1]Ibid, pp. 64.

$$\text{Specialization ratio for each of the industries} = \frac{n_s - \dfrac{n_b + n_s}{N_b + N_s}(N_s)}{n_s}$$

where n_s = industry employment in the subject economy, n_b = industry employment in the benchmark economy, N_s = total employment in the subject economy, N_b = total employment in the benchmark economy, $\frac{n_b + n_s}{N_b + N_s}(N_s)$ = industry employment if the two economies were self-sufficient.

The resulting specialization ratios, which were computed on the basis of the 1950 data, were used to separate monthly employment of each industry in the county from 1953 through 1955 into export-oriented and locally-oriented components. The estimated components then were aggregated, and their totals were used to derive the employment multiplier in Lancaster County. This multiplier was determined by computing the regression of changes of locally oriented employment on changes in export-oriented employment in the county. The resulting regression coefficient was found to be 1.31, and the employment multiplier of 2.31 was derived by adding unity to this coefficient which is actually the ratio between the two employment components in the county.

Two points are worth mentioning with regard to the methods employed in this study. The first point is that the specialization ratio method assumes implicitly that income and expenditure patterns are the same in the county and in the benchmark economy. Since such an assumption is not true in the actual practice, any attempt made to introduce possible differences in the structure of those patterns into the study would result in differences in the computed ratios and, thus, in any derived multiplier value.

The second point to be made is that the method used to estimate the multiplier implies the assumption of a linear functional relationship between the export-oriented and locally-oriented employment and excludes the consideration of time lags. However, any other methods employed in this regard which would take into account time lags (such as that discussed below) and/or would be based on a different assumption would result in a different value of the multiplier obtained.

Another empirical study was completed by K. Sasaki.[1] His main purpose was to estimate the economic impact of changes in military expenditures on the Hawaiian economy. The measure of such impact is the multiplier. Hence, he estimated the employment multiplier in the state of Hawaii which indicates the change in the total employment of this state as a result of a change in employment of the defense sector (or industry). Income multiplier was not determined because certain data required to estimate it were unavailable in the state.

In this study, he assumed a close functional relationship between the total and export-oriented employment and took into account the time lags by setting the locally-oriented employment at time 't' as a function of the total employment at "-t' periods, i.e.

(1) $T = L + X$

(2) $L = f(T_{-1})$ $\qquad (i = 0, 1, 2, 3)$

---

[1]Kyohei Sasaki, "Military Expenditures and the Employment Multiplier in Hawaii," The Review of Economics and Statistics, vol. XLV(August, 1963), pp. 298-304.

where T = total employment, L = locally oriented employment, X = export-oriented employment and i = a given unit of time.

Assuming the functional relationship between T and L to be linear, he then derived from equations (1) and (2) the following linear difference equation:

$$(3) \quad T = m + b_i \sum_{i=1}^{3} T_{-i} + vx$$

where m, b, and v are constants.

Solving this latter equation and taking the derivative of T with respect to x, yielded:

$$(4) \quad \frac{d\bar{T}}{dx} = \frac{1}{1 - \sum_{i=1}^{3} b_i} v$$

which is the employment multiplier that indicates the changes in the equilibrium level of the total employment as a result of a change in the level of the export-oriented employment.

Sasaki used three different methods to divide the total employment in Hawaii into locally-oriented and export-oriented components. (1) He estimated the ratio between the value of exports and the value of total output for each industry for which empirical data on values of exports and output were available. Then, he used this ratio to separate the employment in an industry into locally-oriented and export-oriented categories. (2) In the case of industries for which such data were not available, the division of the total employment was based on location quotient. The ratio of employment in a given industry in Hawaii to the population of Hawaii, P, was compared with the ratio of total employment in that industry in the U.S. to the population of the U.S., P'. When the former exceeded the latter (that is, when the difference (P-P') times the local population was positive) by some fraction, this fraction was assumed to indicate the degree by which this industry is export-oriented. On the other hand, when the former was less than the latter (i.e. (P-P') < 0), the industry was considered to be locally-oriented. (3) Federal government and hotel industry were treated as export-oriented industries (or sectors), and finance and local government were considered as locally-oriented industries.

Using these methods, Sasaki divided Hawaii's employment into its components for the period of 9 years, from 1947 through 1955. The resulting data for this period were then used to derive the employment multiplier in the state.

The method employed to estimate the multiplier value was based on the assumption of a linear relationship between locally-oriented and total employment. The structural equation obtained from equations (1) and (2) was then:

$$(5) \quad L = a + b_i \sum_{i=1}^{3} T_{-i} + u$$

$$(6) \quad T_o = L + X$$

$$(7) \quad X = \text{exogenous}$$

where T_o = total employment at time o, and u = random disturbance.

A substitution of equation (5) into (6) yields the model that was used to compute the multiplier, i.e. yields:

$$(8) \quad T_o = k_o + k_1 T_{-1} + k_2 T_{-2} + k_3 T_{-3} + k_4 X + k_4 u$$

The data used to estimate this regression equation were those obtained for the nine year period as shown above. The employment multiplier in the state of Hawaii, which was found to be 1.28, was computed from the coefficients of determination, R^2 of this equation (8).

Still another study of economic base type of regional multiplier was carried out by Federal Reserve Bank of Kanaas City.[1] Briefly stated, the purpose of this study was to derive employment multiplier in terms of the effect of changes in export-oriented employment on locally-oriented employment, for the city of Wichita, Kansas. Data used for this purpose were total employment by industry for 1940 and 1950. The methods used to separate total employment in each industry into export and locally-oriented categories varied depending on availability of empirical data on sales and markets. With respect to industries for which such data were not available, the division of employment into its components was based on location quotient. The per capita employment in a given industry in the city of Wichita was divided by the per capita employment in that industry in the U.S. If the ratio obtained was less than one, the industry was assumed as locally-oriented. If this ratio was greater than one by some amount, this amount was taken to indicate the extent to which the employment in that industry was export-oriented.

Furthermore, some of the industries in Wichita, such as textile mill products, construction, furniture, and public administration, were considered as exclusively locally-oriented. Other industries, such as petroleum, aircraft, machinery, apparel, and metals, were treated as exclusively export-oriented.

---

[1] Federal Reserve Bank of Kansas City, "The Employment Multiplier in Wichita," Monthly Review, vol. 37 (September 1952).

The method used to derive the multiplier desired was very simple. The change in the total export-oriented employment in 1940-50 was divided by the change in the total locally-oriented employment in the same period. The resulting ratio was 1:16 or 1.6. This ratio indicates that every change in export-oriented employment of one causes a change in locally-oriented employment of 1.6. Thus, the multiplier derived in this study was expressed in terms of the effect of changes in export-oriented employment on locally-oriented employment, rather than on total employment.

A close examination of the empirical studies discussed above reveals the reason for variation in the employment multiplier value derived in each case.

One of these reasons lies in the size of the area chosen. The region (or area) covered in the second of these studies was the largest in size since it included the whole state of Hawaii, whereas that examined in the first analysis was limited within the boundaries of Lancaster County. The region considered in the last study was the smallest of all for it comprised the City of Wichita. The importance of and the problems involved in the selection of regions have been extensively analyzed in the regional science literature and do not need to be repeated here. The point to be made in this context, however, is that the definitions of employment components, and thus the export-locally-oriented employment ratio, are greatly affected by the base area boundaries chosen. Many of the industries which are considered as export-oriented in small areas become locally-oriented as larger areas of analysis are defined. Accordingly, the multiplier value varies depending on the size of the area (or a region) included in the analysis.

Another reason for difference in the computed employment multiplier value lies in the base year which was selected. In the Lancaster County analysis, the base year used was 1950, i.e. total employment data of 1950 were used as a basis for separating total employment into export and locally-oriented components. In the Wichita study, on the other hand, the 1940 and 1950 employment data were used for similar purposes. In the study which was made to compute the employment multiplier in Hawaii, there was no specific base year mentioned. However, the analysis indicated that for the period 1947-55, employment for each year was separated into export and locally-oriented components by the methods explained earlier in this paper.

Such differences in the base year chosen yield different ratios and, hence, different values of multiplier. This is, basically, due to the fact that data for different years are not the same.

A third reason for variation in the multiplier lies in the various methods which were employed in these studies to divide total employment into the two components. Thompson used, in this respect, the so-called "specialization ratio" method which was discussed above and which was indicated to be different from the method used by other analysts. Sasak, on the other hand, employed three methods, in this regard all of which were different from that used by Thompson. Although the methods used in the Wichita study were similar in principle to those employed by Sasaki, the results obtained were different. This is specially true with respect to the number of industries which were considered as exclusively export and exclusively locally-oriented in each of these studies.

It is evident that variation in such techniques would cause variation in the computed multiplier value.

The length of time period which is used to derive the multiplier is also found to be a cause of variation. As it was indicated earlier, the studies examined here used different periods and lengths of time. The first study estimated the multiplier value for the period of 36 months (1953 through 1955). The second and third studies determined that value for the periods of nine years (1947 through 1955) and ten years (1940-1950), respectively. The Wichita study indicates that even the same multiplier would have different values for different periods in length.[1] This fact was also evidenced by computing the employment multiplier in Lancaster County, Nebraska, for only 12 months (1953). In this attempt, both the method and the data given in the first study discussed above were used. The value of this multiplier for the period of 12 months was found to be about 1.00 which is obviously different from 1.31 shown in that study.

One more reason for variation in the multiplier value is embodied in the methods used to derive the multiplier itself. The earlier discussion shows that each of the studies analyzed used a different method for this purpose. Both Thompson and Sasaki employed regression analysis. However, the latter took into consideration time lags whereas the former did not, or at least he did not indicate that he did. Furthermore, the model developed by Sasaki measured total employment multiplier, while the model used by Thompson estimated local employment multiplier. The

---

[1]Federal Reserve Bank of Kansas City, op.cit., pp. 4-7.

Wichita study employed the simple export-oriented--locally-oriented ratio method. This method measures local employment multiplier but is different from that used by Thompson.

In addition to these explained here, there are still other factors that contribute to differences in regional multiplier values which are not revealed in the empirical studies examined. States briefly, such factors include:

(1) Unit of measurement. Evidently, different multipliers would result from different units of measurement used.

(2) Changes in technology, which tends to cause changes in export-local ratio and, hence, in the multiplier for any region over time.

The other type of regional or area multiplier analysis, which is in fact the primary concern of this paper, is that associated with the use of the input-output, or interindustry relations, technique which has been increasingly used as a tool of regional and small-area development analysis.

The main purpose of the input-output analysis is to show the inter-industrial structure of product. In this analysis, a given economy is divided into a relatively large number of identifiable sectors, each of which produces goods and services which are sold to other sectors and, at the same time, purchases goods and services from other sectors. An economy is represented by the closed input-output model if all the sectors within it are considered to be interdependent with functionally related outputs and inputs. If, however, some sectors are not functionally interrelated to others, then the economy is represented by the open input-output model. In such a model the final demand (appearing in the columns and the corresponding rows) is exogenously determined by factors outside the system.

The static, open models, which are commonly used in regional development studies, are based on at least three assumptions. These are that: (1) The level of output of each sector determines uniquely the level of input required by that sector;(2) returns to scale are constant; and (3) each group of commodities and services is supplied by a single production sector.

According to this type of model, a region is classified into (n + 1) sectors. n of these are the processing, or endogenous, sectors and structural interrelationships can be established among them. The remaining is composed of the exogenous sectors which represent the final demand.

The total gross output of any one sector, during a specified period of time is set to be equal to the sum of the amounts of products sold to the endogenous sectors for further processing within the system plus the amounts of output sold to the exogenous sector for final use by its components (typically, household, investors, local government, and other regions). Such identity can be expressed by the following equation:

$$(9)\quad X_i = \sum_{2=1}^{n} X_{ij} + Y_i$$

where X_i = total output of the sector i (i = 1, 2, ..., n); X_{ij} = output of sector i used as input by sector j, both being endogenous sectors; and Y_i = final demand by the exogenous sector for output of sector i.

The model of the whole region can be represented by the following set of linear equations:

$$(10)\quad X_1 = X_{11} + \ldots + X_{1j} + \ldots + X_{1n} + Y_1$$

$$X_i = X_{i1} + \ldots + X_{ij} + \ldots + X_{in} + Y_i$$

$$X_n = X_{n1} + \ldots + X_{nj} + \ldots + X_{nn} + Y_n$$

For each column in the exogenous sector which represents final demand, there is a corresponding row representing a component in the payments, or primary inputs, sector which is also exogenous. The rows which appear in this exogenous sector (typically, household income, government income in terms of taxes, savings, and imports) reflect the payment and receipts for the primary inputs which are not produced by the endogenous sectors of the region. Also, in this model, the total outlays made by each endogenous sector (i.e., column) is equal to the total gross output of that sector (i.e. row). The only condition concerning the exogenous sectors, however, is that the total of all columns in the final demand sector should be equal to the total of all rows in the factor payments sector.

As assumption (1) indicates, the demand for part of the output of one endogenous sector X_i by another endogenous sector X_j is considered as a unique function of the level of output in X_j. This means that:

$$(11)\quad X_{ij} = d_{ij}\, X_j.$$

From this equation we can compute the input coefficients for the processing sectors. That is:

$$a_{ij} = \frac{X_{ij}}{X_j} \qquad (i, j = 1, 2, \ldots, n)$$

Thus, the input-output coefficients of each sector are obtained by dividing each element of a given column by the total of that column. These coefficients show the direct purchases by each producing sector from every other sectors per dollar of output.

It should be stated at this point that, in the open input-output model, each column in the input coefficient table add up to less than unity, the difference being the values of the coefficients of the exogenous sector components which are excluded from input coefficient table because they are determined outside the system. The exception to this is where the coefficients of the latter are zeros, meaning that a given endogenous sector does not purchase labor input, does not pay taxes, does not save, and does not import.

This point is important to make because it shows that although the exogenous sector components are not a part within the structure, the coefficients of the processing or endogenous sectors are dependent in their magnitudes upon the magnitude of one or more elements included in the rows of that exogenous sector.

From these coefficients we can derive the combined direct and indirect effects of changes in final demand upon the production of the regional endogenous sectors as follows. Substituting equation (11) into equation (9), rearranging the terms, and setting the system in a matrix from yields.

$$(12) \quad \begin{bmatrix} (1-a_{11}) & \cdots & -a_{1j} & \cdots & -a_{1n} \\ -a_{i1} & \cdots & (1-a_{ij}) & \cdots & -a_{in} \\ -a_{n1} & \cdots & -a_{nj} & \cdots & (1-a_{nn}) \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_i \\ \vdots \\ Y_n \end{bmatrix}$$

(13) <u>or</u> $(I - A) X = Y$

Since the main goal is to obtain the effect of final demand, Y, upon the output of producing sectors, X, i.e., to express the latter as a function of the former, we multiply both sides of equation (13) by the inverse of (I - A) and obtain the desired expression:

(14) $X = (I - A)^{-1} Y.$

The elements of the inverse matrix $(I - A)^{-1}$ represent the direct and indirect effect of one dollar change in the final demand upon the output of all endogenous sectors within the region. These elements are what we call regional multipliers. Several methods are used to compute them, one of which is to express a given multiplier, say α_{LK}, in terms of the coefficients, (a_{ij}), as follows:

(15) $\alpha_{LK} = \dfrac{A_L^K}{|I-A|}$

where A_L^K is the cofactor of the coefficient a_{KL} in the (I-A) matrix and $|I-A|$ is the determinant of that matrix.

After this rather brief explanation of the input-output technique and how the multipliers are derived, we turn to the main subject matter of this study.

One of the primary and important sources of variations in regional multipliers, obtained by using the interindustrial relations technique, is variations in the input coefficients, a_{ij}'s. This is so because of the functional relationship between the multipliers and the coefficients.

There are various factors which cause variations in the coefficients and, consequently, in the multipliers. One reason for variation can be attributed to differences in the sources and the methods used to collect data necessary for empirical regional studies. Detailed data for an input-output table are not available on regional or small area basis. Different analysts use somewhat different methods in an attempt to obtain data for the regions they select to study. Consequently, the data they obtain are different in quantity and quality, and the regional multipliers computed from them are also different. The early regional analysts based their regional tables on the national input coefficients table. This table is built on the assumption that regional input patterns are identical to national input patterns. However, the regional multipliers obtained by various studies were not identical due to other reasons for variations which are explained in this paper. To give only a few examples, the method used by Moore and Peterson[1] in their Utah study involved the estimation of gross output for each sector from published sources. The interindustry flows were estimated from the national input coefficients and modified in the light of differences in regional productive process, product-mix, and marketing practices. The modification was based on estimates obtained from employment and income data and from technical data. In the study of Roger Mills County, Oklahoma, Jansma[2] used a different method. His single source of data was a local bank from which he obtained information available on micro-filmed checks which cleared the bank in 1960. Then he developed a

---

[1]F. Moore and J. W. Peterson, "REgional Analysis: An Interindustry Model of Utah," Review of Economica and Statistics, vol. 37 (November, 1955).

[2]J. D. Jansma, "Estimating Accounts and Economic Structure for an Oklahoma County," Regional Development Analysis, North Carolina State and The Great Plains Resource Economic Committee, Oklahoma State University, May, 1963.

sampling scheme which provided a "composit month" of 24 days from the total of 306 banking days. The date, amount, and the purpose of each check written in the bank during the 24 sample days were used as a basis for constructing the transactions table. Some assumptions and arbitrary adjustments were made to remedy the discrepancies that appeared in the data and in the resulting table. Still a different approach was used by Hirsch[1] in his study of the St. Louis Metropolitan Area. Published sources were used to obtain sector output totals. The input-output transactions data, however, were obtained directly from medium-sized companies which constructed their own input-output tables. Then he compared the aggregated results from these tables with the control totals he had estimated from published data and made the necessary modifications.

Another reason why coefficients and the multipliers computed from them vary stems from variations in industrial classification. Isard[2] states some of the factors that cause differences in the choice of a set of sectors to be employed as being the quantity and quality of existing data, costs of and resources available for data collection and processing, type of regional situation, the inclinations of the researcher, and purposes of study. Depending on these and other factors, the industrial classification may vary from region to region because of various reasons, such as the different relative importance of some

---

[1] W. Z. Hirsch. "Interindustry Relations of a Metropolitan Area," The Review of Economics and Statistics, vol. 41(1959), pp. 360-369.

[2] W. Isard. Methods of Regional Analysis: An Introduction to Regional Science. Cambridge: The M.I.T. Press, 1960, pp. 319-322.

sectors in different regions, or "the different degree of technical linkage between these sectors in the several regions, or the different degrees of interlocking corporate and management structures,"[1] or different limitations of data availability for disaggregation. Whatever the reason is, however, many of the regional input-output tables vary in the number of sectors they contain. The input-output table constructed for Utah includes the total of seven sectors which are consolidations of 75 producing industries for which data were collected. The table developed for St. Louis, on the other hand is much more disaggregated and contains the total of 33 sectors although this area is much smaller in the size than Utah. Still different is the table that made up for Roger Mills County which comprises of 12 sectors.

Variation in industrial compositions of different models cause differences in both the input coefficients and the multipliers. Much has been written on the aggregation problems[2] and how it introduces errors and variations in the multipliers. However, sufficient and workable techniques to correct errors or changes in the multipliers resulting from aggregation or disaggregation have not yet been fully developed.

Still another reason for differences in the regional coefficients, and multipliers, is variation in the regional boundaries selected by different studies. If a region selected is relatively self-sufficient and does not have significant ties to other regions, its imports and exports will be relatively very small. Consequently the input-output table for such a region will show a strong interdependence among its

---

[1] Ibid., pp. 321.

[2] See for example John Green, "

sectors, meaning that the coefficients as well as the multipler are relatively large. On the other extreme if the region chosen by another analyst is significantly dependent on "the rest of the world," in its imports and exports, then the coefficients of the input-output table constructed for this region will be relatively low. The reason is that considerable portions of the total columns are embodied in the import entries which are excluded from coefficients table. As a result of this situation the multipliers also will be relatively small. Between these two extremes, there are many other differences between different regions. Hence variations in the choice of a region is one of the main factors that cause variation in regional multipliers. In the empirical studies examined in this study, St. Louis was found to be dependent on other regions, and the Utah study showed that changes in Utah's import requirements lead to rather small changes in the demand for its exports. Another point to be made is that generally input-output models of small areas are more open than those which apply to broad geographic areas. There is more specialization and exchange among smaller areas than among sufficiently large areas so that small area imports and exports account for a substantial proportion of total transactions.

One additional reason for variation is differences in the extent to which the input-output models are open to include or exclude the final demand sectors. There is no fixed rule to exclude or include any specific sector in the payments sector or final demand sector. Some studies might close the system with respect to only one or two of the components if exogenous sectors; others might exclude more than that; and still others might shift one or more of the endogenous sectors into final demand or primary input sectors. To decide what sector to include or to exclude from the structure depend, to a large extent, upon the purposes of the studies. Such variations in decisions will no doubt affect the regional multipliers. In the typical studies, all the components which are commonly known as belonging to the exogenous sectors have been excluded from the system. As was stated earlier, a model of this type yields multipliers which are defined as "direct and indirect effect." Some regional economists and analysts, however, who were interested in measuring the income and employment multipliers in terms of change included household in the structure table. Among them were Hirsch, and Moore and Peterson. The main purpose of treating household as endogenous sector was to measure the direct and indirect effects plus the induced change in income which results from a change in household spending. In both studies, two types of income multipliers were computed. The multipliers obtained after including household in the system were found to be greater than those computed by excluding household from the structure, the difference being the induced change in income. Moreover, employment multipliers of both types were measured by Moore and Peterson for only some sectors and were found to be different, the second type (with household included) being greater in magnitude than the first.

Still another factor causing differences in the regional multipliers is found in different models used for determining the volumns of exports and imports in a region. Irving Hoch[1] explained two different models in this regard. This regard shows how regional multipliers vary (under

[1]Irving Hoch. "Inter-Industry Factors for the Chicago Region," Regional Science Associations, Papers, 1959, pp. 217.

various assumptions) in these two models. One is called "variable trading model," in which a given row's entries show output a given sector flowing into the sectors of the regional economy. The difference between regional output and factor input requirements appears as export (+) or import (-) in the foreign trade column. The equality of row and column is obtained by adding export or import to a given row so it seems to the corresponding column. In the other model, "constant trading pattern," a given row's entries show output of a given sector used as inputs by the sectors of the regional economy, where the outputs are only those produced by the regional sector. Any additional outputs needed from that sector appear as components of an import row. A particular entry in the import row shows all the imports flowing into the corresponding sector, regardless of the sector source of those imports. This means that all the imports into a given sector are added together.

Besides these two models, the model constructed by Hirsch[1] in his study of St. Louis area can be considered as a third model in which all export and import flows are given in detail by sector. In this model, there is an export table appearing to the right of the input-output table and a similar import table appearing below the input-output table. This table shows the interindustry transactions within the region and also the detailed interindustry transactions between this region and the "rest of the world." Variations in the volume of imports, as well as of other final demand components cause variations in the values of coefficients because of the fact that the latter are derived by dividing the elements of each column by the total of that column.

---

[1]Hirsch, op. cit.

One more reason for differences in regional multipliers stems from the fact that in some studies data on sector output totals as well as the data used in input-output tables may be collected in terms of consumer's prices. In other studies, producer's values are used. It is also possible to measure the production by physical unit. The coefficients, and, consequently, the multipliers will vary depending on what measurement is used in the regional studies.

Also, the values of regional multipliers differ depending on the types of models used to measure them. Different from the method explained earlier, was the approach followed by Moore and Peterson to measure the multipliers of both types of ratios. The multipliers of type (9) (both income and employment) were computed by dividing each element of direct and indirect changes (in income or in employment) by the direct changes. The multipliers of type (10) were derived by taking the ratios of direct, indirect, and induces changes (in income or employment) to direct change. It should be noted that the income multiplier of type (9) is not the Keynesian-type multiplier which is dependent on effects generated by change in local income itself. The income multiple of type (9) reveals only the interindustry linkage effect.

Hirsch adopted Moore's and Peterson's method and applied it on St. Louis area. However, he measured income multipliers (of both types) of all the sectors included in the system, whereas Moore and Peterson measured the multipliers of only a few sectors in Utah. Furthermore, Hirsch derived only the type (9) of the employment multipliers. In both studies, consumption-income function as well as employment-production function were assumed to be linear. Hirsch himself admitted that these assumptions have caused overestimates in the multiplier values he computed. The values of the regional multipliers would have been different had he or Moore and Peterson assumed non-linear functional relationships between these variables. This point reveals another reason for variations in regional multipliers. The model used by Jansma to measure income multipliers of Roger Mills County, Oklahoma was essentially based on "From-to" model which had been developed by Leven. His main interest was with agricultural sectors and the multipliers he derived were different in their values from those derived by Hirsch as well as those measured by Moore and Peterson.

In addition to the factors analyzed thus far, there are other factors which cause variations in both the input coefficients and the multipliers in the long-run. These can be outlined as follows: (1) Changes in the relative prices of the factor inputs in the long-run will lead to change in the value of production and as a result will require us to revise the input coefficient tables in order that they be consistent with the prevailing economic situations; (2) As the time passes new industries might appear in one or more regions. Such industries cannot be taken account of in the current input coefficients table of that region(s) unless the table is replaced by a new and more consistent one; (3) The effects of technological changes on the input coefficients in the long-run should be counted for by bringing the input coefficients table up-to-date. Such effects are important to consider because some industries might, through time, substitute capital for labor or vice-versa depending on various factors one of which might be a development of new machines which would reduce the cost of producing a given product(s) in a given sector(s) of one or more regions. In a case such as this, an industry might become capital-intensive. In difference cases labor might be substituted for capital. If coefficient tables are to be consistant these effects cannot be ignored.

Since these and other related problems have been extensively analyzed in the regional science literature a repetition can be eliminated here.

III. Effects on Multipliers of Errors in Coefficients

The objective now is to develop mathematical and statistical tools for measuring the sensitivity of multiplier estimates to errors in data, i.e., changes in coefficients. It would be appropriate, however, to set forth an input-output model, first, and to show how the multipliers are derived. This will indicate the functional relationships between the coefficients and the multipliers and will make the following analysis more understandable.

Suppose, for illustrative purposes, the existence of an economy with n number of sectors (or activities). Let X_i measure the total output (or employment) of the i^{th} sector, where i = 1, 2, ..., and n; and X_{ij} measure the output (or employment) of the i^{th} sector which is used by the j^{th} sector as an input, where j = 1, 2, ..., and n. Furthermore, assume that Y_i = final demand for the product of the i^{th} sector (or the employment exported from the i^{th} sector). Based on these assumptions, the input-output table can be formed as follows:

Sectors	1 . . j . . n	Final Demand Y_i	Total Output X_i
1	$X_{11} \cdots X_{1j} \cdots X_{1n}$	Y_1	X_1
.	. . .	.	.
.	. . .	.	.
.	. . .	.	.
i	$X_{i1} \quad X_{ij} \quad X_{in}$	Y_i	X_i
.	. . .	.	.
.	. . .	.	.
.	. . .	.	.
n	$X_{n1} \quad X_{nj} \quad X_{nn}$	Y_n	X_n

The data of this table can be represented as a system of simultaneous equations which is:

$$X_{11} + \ldots + X_{1j} + \ldots + X_{1n} + Y_1 = X_1$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$X_{i1} + \ldots + X_{ij} + \ldots + X_{in} + Y_i = X_i$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$X_{n1} + \ldots + X_{nj} + \ldots + X_{nn} + Y_n = X_n$$

Since in this case i = j, the above system can be rearranged and set as follows:

$$X_1 - X_{11} - \dots - X_{1i} - \dots - X_{1n} = Y_1$$

$$\dots\dots\dots\dots\dots\dots\dots$$

$$- X_{i1} - \dots + X_i - X_{ii} - \dots - X_{in} = Y_i$$

$$\dots\dots\dots\dots\dots\dots\dots$$

$$- X_{n1} - \dots - X_{ni} - \dots + X_n - X_{nn} = Y_n$$

The input coefficients, a_{ij}, which shows the amount of output from the i^{th} producing sector required per unit of output of the j^{th} using sector, is equal to the output of i^{th} sector used by j^{th} sector divided by the total output of j^{th} sector, <u>i.e.</u>,

$$a_{ij} = \frac{X_{ij}}{X_j}$$

From this definition, we find that $X_{ij} = a_{ij}X_j$. Substituting X_{ij} by $a_{ij}X_j$, the system can be rewritten as follows:

$$(1 - a_{11})X_1 - \dots - a_{1i}X_i - \dots - a_{1n}X_n = Y_1$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$- a_{i1}X_1 - \dots + (1 - a_{ii})X_i - \dots - a_{in}X_n = Y_i$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$- a_{n1}X_1 - \dots - a_{ni}X_i - \dots + (1 - a_{nn})X_n = Y_n$$

Putting these equations in matrix form, we get:

$$\begin{bmatrix} (1 - a_{11}) \dots - a_{1i} \dots - a_{1n} \\ \dots\dots\dots\dots\dots \\ - a_{i1} \dots\dots (1-a_{ii}) \dots - a_{in} \\ \dots\dots\dots\dots\dots \\ - a_{n1} \dots\dots - a_{ni} \dots (1-a_{nn}) \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ \dots \\ X_i \\ \dots \\ X_n \end{bmatrix} = \begin{bmatrix} Y_1 \\ \dots \\ Y_i \\ \dots \\ Y_n \end{bmatrix}$$

which can be written as:

$$(I - A)X = Y$$

where A = matrix of input-output coefficients a_{ij}, X = the vector of total outputs, and Y = the vector of final demands for output of each sector. Since the main goal of input-output analysis is to relate output of one sector to quantities of final demand for output of other sectors, i.e., to express output of producing sectors as a function of final demand, we multiply both sides of the above equation by the inverse of (I - A) and obtain the desired expression:

$$X = (I - A)^{-1} Y$$

The elements of the inverse matrix $(I - A)^{-1}$, which are called multipliers, are of our concern in this study. Let the multiplier α_{LK} refer to the element in the L^{th} row and the K^{th} column of the $(I - A)^{-1}$ matrix, where L = 1, 2, ..., and n; and K = 1, 2, ..., and n. It indicates the amount by which output (or employment) of the L^{th} producing sector will change as final demand for output of the K^{th} sector (or employment export) is changed by one unit. The total effect on output (or employment), i.e., the total regional multiplier, of a change in final demand for output of the K^{th} sector is, then the sum over L of α_{Lk}.

Any given multiplier, say α_{Lk}, can be expressed in terms of the coefficients a_{ij} as follows:

$$\alpha_{Lk} = \frac{A_L^k}{|I - A|}$$

where A_L^k is the co-factor of the coefficient a_{kL} in the (I - A) matrix and $|I - A|$ is the determinant of the coefficient matrix (I - A).

Having set forth the functional relationship between the multipliers and coefficients, the problem on hand, now, is to trace the effect of the change in coefficients of the matrix A on an element of the inverse matrix $(I - A)^{-1}$, i.e., on the multiplier α_{Lk}. To do that, we first differentiate α_{Lk} with respect to some coefficient a_{ij} in the (I - A) matrix which gives:

$$\frac{\partial \alpha_{Lk}}{\partial a_{ij}} = \alpha_{Lk}\alpha_{ji} - \frac{A^{ki}_{Lj}}{|I-A|}$$

$$= \alpha_{Lk}\left(\alpha_{ji} - \frac{A^{ki}_{Lj}}{A^{k}_{L}}\right)$$

where A^{ki}_{Lj} is the cofactor of the elements in the K^{th} and i^{th} rows and the L^{th} and j^{th} columns of the coefficient matrix (I - A), and A^{k}_{L} is the cofactor of the element in the k^{th} row and L^{th} column of the (I - A) matrix. The ratio of the two cofactors, $\frac{A^{ki}_{Lj}}{A^{k}_{L}}$, can be considered as an expression of a new multiplier, different from α_{Lk}, which we will call a sub-multiplier and denote by $\alpha^{(ji)}_{Lk}$. Interpreted in economic terms, $\alpha^{(ji)}_{Lk}$ indicates the effect on output of the L^{th} sector as a result of one unit change in final demand for output of the k^{th} sector in an economy where the i^{th} sector is not an input within the system (but has output which is equal to its total production) and the j^{th} sector is not an output in the system (but it is an input sector and therefore it imports).

Substituting $\alpha^{(ji)}_{Lk}$ into the derivative equation yields:

$$\frac{\partial \alpha_{Lk}}{\partial a_{ij}} = \alpha_{Lk}\ (\alpha_{ji} - \alpha^{(ji)}_{Lk})$$

When k = i and/or L = j the sub-multiplier $\alpha_{Lk}^{(ji)}$ would be zero and the above equation would become:

$$\frac{\partial \alpha_{Lk}}{\partial a_{ij}} = \alpha_{Lk}\alpha_{ji}$$

This expression indicates the change in the multiplier α_{Lk} resulting from a change in a coefficient a_{ij} in terms of the multipliers themselves which are easy to compute. However, a derivative of α_{Lk} with respect to only one coefficient is only the first step in solving our problem because all the a_{ij}'s vary simultaneously rather than one at a time.

A second step towards identifying the sensitivity of multipliers to coefficient errors is to derive the variances and coefficients of variation of multipliers in terms of specific variations in coefficients. Let

$[a_{ij}]$ = A = the matrix of coefficients obtained from accurate data, <u>i.e.</u>, true coefficients.

$[\alpha_{Lk}]$ = the matrix of true multipliers, <u>i.e.</u>, multipliers obtained from true coefficients.

$[\hat{a}_{ij}]$ = $\hat{A}$ = the matrix of estimated coefficients.

$[\hat{\alpha}_{Lk}]$ = the matrix of estimated multipliers.

Since the multipliers are functions of the coefficients, as was shown earlier, i.e., $\alpha_{Lk} = \frac{A^k_L}{|I-A|}$, then the variation between estimated and true values of any one of the multipliers, say, α_{Lk} can be expressed as follows:

$$(\hat{\alpha}_{Lk} - \alpha_{Lk}) = \sum^{ij} \frac{\partial \alpha_{Lk}}{\partial a_{ij}}(\hat{a}_{ij} - a_{ij})$$

where the right side of the equation indicates the total amount of change in α_{Lk} resulting from specific changes (or errors) in the coefficients of matrix A. Squaring both sides of the equation gives:

$$\text{variance of } \alpha_{Lk} = E(\hat{\alpha}_{Lk} - \alpha_{Lk})^2 = E\left[\sum^{i,j} \frac{\partial \alpha_{Lk}}{\partial a_{ij}} (\hat{a}_{ij} - a_{ij})\right]^2,$$

which measures the variance of a multiplier α_{Lk} in terms of variances and co-variances of the coefficients, $\left[a_{ij}\right]$. Expanding this formula yields:

$$\text{var}(\alpha_{Lk}) = \sum^{i,j} \sum^{r,s} \left(\frac{\partial \alpha_{Lk}}{\partial a_{ij}}\right)\left(\frac{\partial \alpha_{Lk}}{\partial a_{rs}}\right) \sigma a_{ij,rs}$$

$$\text{since } \frac{\partial \alpha_{Lk}}{\partial a_{ij}} = \alpha_{Lk}\left(\alpha_{ji} - \frac{A^{ki}_{Lj}}{A^k_L}\right) = \alpha_{Lk}\left(\alpha_{ji} - \alpha^{(ji)}_{Lk}\right)$$

$$\text{and } \frac{\partial \alpha_{Lk}}{\partial a_{rs}} = \alpha_{Lk}\left(\alpha_{sr} - \frac{A^{kr}_{Ls}}{A^k_L}\right) = \alpha_{Lk}\left(\alpha_{sr} - \alpha^{(sr)}_{Lk}\right), \text{ then}$$

$$\text{var}(\alpha_{Lk}) = \sum^{i,j} \sum^{r,s} \left[\alpha^2_{Lk}\left(\alpha_{ji} - \alpha^{(ji)}_{Lk}\right)\left(\alpha_{sr} - \alpha^{(sr)}_{Lk}\right) \sigma a_{ij,rs}\right],$$

or

$$\text{var}(\alpha_{Lk}) = \alpha^2_{Lk} \sum^{i,j} \sum^{r,s} \left(\alpha_{ji} - \alpha^{(ji)}_{Lk}\right)\left(\alpha_{sr} - \alpha^{(sr)}_{Lk}\right) \sigma a_{ij,rs}$$

It can be easily seen that if variations of all the coefficients are equal to zero, then the variance as well as the variation of α_{Lk} will be zero.

To develop these formulas further, assume that the co-variances of coefficients are zero, i.e., that $\sigma a_{ij,rs} = 0$ except when $i = r$ and $j = s$, then

$$\text{var}(\alpha_{Lk}) = \sum^{i,j} \left(\frac{\partial \alpha_{Lk}}{\partial a_{ij}}\right)^2 \left(\hat{a}_{ij} - a_{ij}\right)^2$$

$$= \sum^{i,j} \left(\frac{\partial \alpha_{Lk}}{\partial a_{ij}}\right)^2 \sigma a_{ij,ij}, \text{ or}$$

$$= \alpha^2_{Lk} \sum^{i,j} \left(\alpha_{ji} - \alpha^{(ji)}_{Lk}\right)^2 \sigma a_{ij,ij}$$

Let us further assume that all the coefficients of variation of the coefficients, $\left[a_{ij}\right]$, are the same and equal to some constant C, i.e.,

$\frac{\sigma a_{ij}}{a_{ij}} = \frac{\sigma a_{kh}}{a_{kh}} = C$, then $\sigma a_{ij,ij} = C^2 a_{ij}^2$; $\sigma a_{kh,kh} = C^2 a_{kh}^2$, and:

$$\mathrm{var}(\alpha_{Lk}) = C^2 \overset{i,j}{\Sigma} \left[\frac{\partial \alpha_{Lk}}{\partial a_{ij}}\right]^2 a_{ij}^2, \text{ or:}$$

$$= C^2 \alpha_{Lk}^2 \overset{i,j}{\Sigma} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)}\right]^2 a_{ij}^2$$

The coefficient of variation of the multiplier α_{Lk} derived from these last equations can be expressed in elasticity form as follows:

$$\sqrt{\frac{\mathrm{var}(\alpha_{Lk})}{\alpha_{Lk}^2}} = \sqrt{\frac{C^2}{a_{Lk}^2} \overset{i,j}{\Sigma} \left[\frac{\partial \alpha_{Lk}}{\partial a_{ij}}\right]^2 a_{ij}^2}$$

$$\frac{\sqrt{\mathrm{var}(\alpha_{Lk})}}{\alpha_{Lk}} = C \sqrt{\overset{i,j}{\Sigma} \left[\frac{\partial \alpha_{Lk}}{\partial a_{ij}} \cdot \frac{a_{ij}}{Lk}\right]^2}$$

This formula is of practical importance in that it indicates the degree of responsiveness of α_{Lk} to changes in the coefficients. Now, substituting $\frac{\partial \alpha_{Lk}}{\partial a_{ij}}$ by its expression gives:

$$\frac{\sqrt{\mathrm{var}(\alpha_{Lk})}}{\alpha_{Lk}} = \sqrt{\frac{C^2}{\alpha_{Lk}^2} \alpha_{Lk}^2 \overset{i,j}{\Sigma} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)}\right]^2 a_{ij}^2}$$

$$= C\sqrt{\overset{i,j}{\Sigma} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)}\right]^2 a_{ij}^2}$$

It is worth mentioning that in the equations developed above both variance and coefficient of variation of any one of the multipliers, say, α_{Lk} are expressed in terms of the assumed known coefficients, $\left[a_{ij}\right]$ and multipliers and of the constant C. This means that the effect on α_{Lk} resulting from changes or errors in coefficients can be measured even if the values of estimated coefficients $\left[\hat{a}_{ij}\right]$ are not known in detail. Only the value of C should be known in order to compute the actual magnitude of such effect.

The variance and coefficient of variation of the total regional multiplier,[4] i.e., $\overset{L}{\Sigma}\,\alpha_{Lk}$, can be measured as follows:

$$\text{var}\left[\overset{L}{\Sigma}\,\alpha_{Lk}\right] = E\left[\overset{L}{\Sigma}\,\hat{\alpha}_{Lk} - \overset{L}{\Sigma}\,\alpha_{Lk}\right]^2$$

$$= E\left[\overset{L}{\Sigma}\,(\hat{\alpha}_{Lk} - \alpha_{Lk})\right]^2$$

$$= \left[\overset{L}{\Sigma}\left(\overset{i,j}{\Sigma}\,\frac{\partial\alpha_{Lk}}{\partial a_{ij}}\,(\hat{a}_{ij} - a_{ij})\right)\right]^2$$

Assuming that $\sigma a_{ij,rs} = 0$, except for $i = r$ and $j = s$, yields:

$$\text{var}\left[\overset{L}{\Sigma}\,\alpha_{Lk}\right] = \overset{i,j}{\Sigma}\left[\overset{L}{\Sigma}\left(\frac{\partial\alpha_{Lk}}{\partial a_{ij}}\right)^2 \sigma a_{ij,ij}\right]$$

$$\text{since} \quad \left[\overset{L}{\Sigma}\,\frac{\partial\alpha_{Lk}}{\partial a_{ij}}\right]^2 = \overset{L,h}{\Sigma}\,\frac{\partial\alpha_{Lk}}{\partial a_{ij}}\cdot\frac{\partial\alpha_{hk}}{\partial a_{ij}}\,, \text{ then:}$$

$$\text{var}\left[\overset{L}{\Sigma}\,\alpha_{Lk}\right] = \overset{i,j}{\Sigma}\left[\overset{L,h}{\Sigma}\,\frac{\partial\alpha_{Lk}}{\partial a_{ij}}\cdot\frac{\partial\alpha_{hk}}{\partial a_{ij}}\cdot\sigma a_{ij,ij}\right], \text{ \underline{or}:}$$

---

[4]The total regional multiplier here is defined as the total effect on regional output, i.e., output of all the producing sectors in the region or economy, resulting from a change in final demand for output of one sector, K. Such multiplier is obtained by summing over L of α_{Lk}, i.e., $\overset{L}{\Sigma}\,\alpha_{Lk}$.

$$= \sum^{L,h} \left[\sum^{i,j} \left(\frac{\partial \alpha_{Lk}}{\partial a_{ij}} \right) \left(\frac{\partial \alpha_{hk}}{\partial a_{ij}} \right) \sigma a_{ij,ij} \right]$$

Assuming further that $\frac{\sigma a_{ij,ij}}{a_{ij}^2} = \frac{\sigma a_{kr,kr}}{a_{kr}^2} = C^2$, where C^2 is constant, gives:

$$\text{var} \left[\sum^{L} {}_{Lk} \right] = C^2 \sum^{L,h} \left[\sum^{i,j} \left(\frac{\partial \alpha_{Lk}}{\partial a_{ij}} \right) \left(\frac{\partial \alpha_{hk}}{\partial a_{ij}} \right) a_{ij}^2 \right]$$

The coefficient of variatinn of $\left[\sum^{L} \alpha_{Lk} \right]$ is indicated by the following equation:

$$\sqrt{\frac{\text{var} \left[\sum^{L} \alpha_{Lk} \right]}{\left[\sum^{L} \alpha_{Lk} \right]^2}} = \frac{C}{\sum^{L} \alpha_{Lk}} \sqrt{\sum^{L,h} \left[\sum^{i,j} \left(\frac{\partial \alpha_{Lk}}{\partial a_{ij}} \right) \left(\frac{\partial \alpha_{hk}}{\partial a_{ij}} \right) a_{ij}^2 \right]}$$

$$= \frac{C}{\sum^{L} \alpha_{Lk}} \sqrt{\sum^{L,h} \left[\sum^{i,j} \alpha_{Lk}(\alpha_{ji} - \alpha_{Lk}^{(ji)}) \cdot \alpha_{hk}(\alpha_{ji} - \alpha_{hk}^{(ji)}) \cdot a_{ij}^2 \right]}$$

In These formulas, also, the variance and coefficient of variation of the total regional multipliers, $\left[\sum^{L} \alpha_{Lk} \right]$, are measured in terms of specific variations in the coefficients.

IV. Results of Empirical Study

The mathematical expressions developed in the preceding section were applied to an empirical input-output matix in order to obtain an impression of the actual magnitude of the variances and the coefficients of variation of multipliers. The empirical data used in this respect are estimated employment coefficients of six counties in North Carolina for the year 1960 as shown below:[5]

$$\left[a_{ij}\right] = A = \begin{bmatrix} .466560 & 0 & .021088 & 0 & .016104 & 0 \\ 0 & .397388 & .031045 & .114756 & .007471 & .010359 \\ .009582 & .002612 & .430060 & 0 & .006486 & 0 \\ 0 & .048902 & 0 & .426217 & .000447 & .012962 \\ .051984 & .013892 & .036318 & .002490 & .488077 & .011736 \\ 0 & .014447 & 0 & .013691 & .009617 & .495203 \end{bmatrix}$$

where a_{ij} indicates the amount of employment in county i used per one unit of employment (per employee) in county j.[6]

In order to compute the values of the multipliers, α_{Lk}, we first set the matrix of (I - A) which is:

---

[5]These coefficients have been estimated by Mr. Paul Stone, Economics Department, North Carolina State University, and constitute part of his Ph.D. thesis which is being written now.

[6]The coefficients in this particular example are independent of each other.

$$(I - A) = \begin{bmatrix} .533440 & 0 & -.021088 & 0 & -.016104 & 0 \\ 0 & .602612 & -.031045 & -.114756 & -.007471 & -.010359 \\ -.009582 & -.002612 & .569940 & 0 & 0 & 0 \\ 0 & -.048902 & 0 & .573783 & -.000447 & -.012962 \\ -.051984 & -.013892 & -.036318 & -.002490 & .511923 & -.011736 \\ 0 & -.014447 & 0 & -.013691 & -.009617 & .504797 \end{bmatrix}$$

The multipliers, which are the inverse matrix $(I - A)^{-1}$, are found to be equal to the values shown in the following matrix:[7]

$$\left[\alpha_{Lk}\right] = \begin{bmatrix} 1.88181130 & .00179421 & .07356054 & .00065468 & .06018380 & .00145284 \\ .00432721 & 1.68903660 & .09388978 & .33897291 & .02709795 & .04399494 \\ .03386175 & .00832323 & 1.75777280 & .00178469 & .02347332 & .00076235 \\ .00060623 & .14519177 & .00823085 & 1.77303780 & .00470376 & .04861629 \\ .19370188 & .04853377 & .13488726 & .01934913 & 1.96283170 & .04712658 \\ .00383054 & .05320174 & .00548007 & .05815780 & .03829744 & 1.98446980 \end{bmatrix}$$

where α_{Lk} measures the effect on employment (goods and services in terms of employment) of county L as a result of one unit change in employment of county k which is exported out of the region, <u>i.e.</u>, out of the six counties.

Variances of $\left[\alpha_{Lk}\right]$ can be computed by applying the formula:

$$\text{var}\ (\alpha_{Lk}) = \left[\sum^{ij} \frac{\partial \alpha_{Lk}}{\partial a_{ij}} \left(\hat{a}_{ij} - a_{ij}\right)\right]^2$$

---

[7] $\left[\alpha_{Lk}\right]$ refers to the matrix of all the multipliers, whereas α_{Lk} represents an element located in the L^{th} row and the k^{th} column of this matrix, where $L = 1,2,\ldots, n$; and $k = 1,2,\ldots,n$.

$$= \alpha_{Lk}^2 \left[\sum^{i,j} \left(\alpha_{ji} - \alpha_{Lk}^{(ji)} \right) \sigma a_{ij} \right]^2 \text{, where } \sigma a_{ij} \text{ is a}$$

variation of a_{ij}

$$= \alpha_{Lk}^2 \sum^{i,j} \sum^{r,s} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)} \right] \left[\alpha_{sr} - \alpha_{Lk}^{(sr)} \right] \sigma a_{ij,rs}$$

Due to the fact that the values of σa_{ij}'s are not known in our case, we assume that all of them are equal and that $\sigma a_{ij,rs} = .05$. Then the equation becomes:

$$\text{var}(\alpha_{Lk}) = \sigma a_{ij,rs} \, \alpha_{Lk}^2 \left[\sum^{i,j} \left(\alpha_{ji} - \alpha_{Lk}^{(ji)} \right) \right]^2$$

$$= \sigma a_{ij,rs} \, \alpha_{Lk}^2 \sum^{i,j} \sum^{r,s} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)} \right] \left[\alpha_{sr} - \alpha_{Lk}^{(sr)} \right]$$

Using this last expression, which indicates that $\text{var}(\alpha_{Lk})$ equals to the Constant $\sigma a_{ij,rs} = .05$ multiplied by the square of the total average rate of change in α_{Lk} resulting from one unit change in each one of the entire Coefficients, we obtain the numerical values of variances of multipliers as shown below in matrix form:

$$\begin{bmatrix} .914846 & .772186 & .876964 & .979701 & .913529 & .921996 \\ 1.083084 & .914278 & 1.038247 & 1.159893 & 1.081505 & 1.091579 \\ .747912 & .631367 & .716974 & .800977 & .746850 & .753688 \\ .854502 & .742664 & .843358 & .942140 & .878496 & .886684 \\ 1.299167 & 1.096575 & 1.245258 & 1.391171 & 1.297147 & 1.309229 \\ 1.030412 & .869988 & .987944 & 1.103711 & 1.029115 & 1.038701 \end{bmatrix}$$

Then, taking the square roots of these variances we get variations of the multipliers which are shown in the following matrix:

$$\begin{bmatrix} .956 & .878 & .936 & .989 & .955 & .960 \\ 1.040 & .956 & 1.019 & 1.077 & 1.040 & 1.045 \\ .865 & .794 & .846 & .895 & .864 & .868 \\ .924 & .862 & .918 & .970 & .937 & .941 \\ 1.140 & 1.047 & 1.116 & 1.180 & 1.139 & 1.144 \\ 1.015 & .933 & .994 & 1.051 & 1.014 & 1.019 \end{bmatrix}$$

The number in the first row and the first column of this matrix, for example, indicates that if each one of the entire coefficients of the A matrix changes (i.e. deviates from its true value) by $\sqrt{.05} = 0.223$, then the resulting change which is introduced in the multiplier α_{11} would be 0.956. From this information, then, we can compute the value of the new multiplier α_{11} (i.e. of estimated multiplier, $\hat{\alpha}_{11}$) as follows:

$$(\hat{\alpha}_{11} - \alpha_{11})^2 = \left[\sum^{i,j} \frac{\partial\alpha_{11}}{\partial a_{ij}} (\hat{a}_{ij} - a_{ij})\right]^2$$

Under the given assumptions, we have found that the right side of this equation is equal to 0.914846 or to $(.956).^2$

Substituting this value gives:

$$(\hat{\alpha}_{11} - \alpha_{11})^2 = (.956)^2\text{, or:}$$

$$(\hat{\alpha}_{11} - \alpha_{11}) = |.956| \quad \text{in absolute term}$$

The value of α_{11}, as shown in the data above, is 1.8818, hence

$$(\hat{\alpha}_{11} - 1.8818) = |.956|$$

If the number .956 indicates a decrease, then:

$$\hat{\alpha}_{11} = 1.8818 - .956 = .9258$$

If the number .956 indicates an increase, then

$$\hat{\alpha}_{11} = 1.8818 + .956 = 2.8378$$

The same way can be applied regarding the other multipliers.

So far, we have been assuming that the coefficients are interdependent or interrelated to each other, i.e., that any change occurs in one of them affects the values of the others. In our empirical example, however, all the coefficients are independent, i.e., a change in one does not affect the others.[9] The case being that, we set $\sigma a_{ij,rs} = 0$, except for $i = r$ and $j = s$, and compute the variances of $\left[\alpha_{Lk}\right]$ using the equation:

$$\mathrm{var}(\alpha_{Lk}) = \sigma a_{ij,ij}\, \alpha_{Lk}^{2} \sum^{i,j} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)}\right]^{2}$$

where $\sigma a_{ij,ij}$ is assumed to be constant and equal to 0.05. The values of $\mathrm{var}(\alpha_{Lk})$'s obtained this way are:

$$\begin{bmatrix} .635468 & .511038 & .554267 & .576947 & .685040 & .700208 \\ .533250 & .428890 & .465100 & .468792 & .575318 & .587603 \\ .553346 & .445053 & .482632 & .503466 & .596527 & .609654 \\ .566887 & .455953 & .494437 & .516599 & .611124 & .624673 \\ .700450 & .562823 & .610933 & .637968 & .755105 & .771843 \\ .706142 & .566989 & .615997 & .643631 & .761371 & .778247 \end{bmatrix}$$

---

[9] This fact is due to the assumptions set and the method applied in estimating the coefficients, the analysis of which is immaterial at the present stage of the study.

And the variations of the multipliers, i.e., $\left[\sqrt{\mathrm{var}(\alpha_{Lk})}\right]$, are:

$$\begin{bmatrix} .7971 & .7148 & .7444 & .7595 & .8276 & .8367 \\ .7302 & .6549 & .6820 & .6846 & .7585 & .7665 \\ .7432 & .6671 & .6947 & .7095 & .7723 & .7808 \\ .7522 & .6752 & .7031 & .7187 & .7817 & .7903 \\ .8370 & .7502 & .7816 & .7987 & .8690 & .8785 \\ .8403 & .7530 & .7848 & ,8022 & .8725 & .8822 \end{bmatrix}$$

where, for example, $\sqrt{\mathrm{var}(\alpha_{11})} = .7971$ indicates the same thing as explained on page 15 above. Now, let us drop the assumption of constant $\sigma a_{ij,ij}$, and replace it by the assumption that not the variances but the coefficients of variation of a_{ij}'s are equal to the same constant C, i.e,

$$\frac{\sigma a_{ij}}{a_{ij}} = \frac{\sigma a_{kh}}{a_{kh}} = C$$

Then the modified equation of $\mathrm{var}(\alpha_{Lk})$ would be:

$$\mathrm{var}(\alpha_{Lk}) = C^2 \sum^{i,j} \left[\frac{\partial\alpha_{Lk}}{a_{ij}}\right]^2 a_{ij}^2$$

$$= C^2\alpha_{Lk}^2 \sum^{i,j} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)}\right]^2 a_{ij}^2$$

and of $\dfrac{\sqrt{\mathrm{var}(\alpha_{Lk})}}{\alpha_{Lk}}$ would be:

$$\frac{\sqrt{\mathrm{var}(\alpha_{Lk})}}{\alpha_{Lk}} = C \sqrt{\sum^{i,j} \left[\frac{\partial\alpha_{Lk}}{\partial a_{ij}} \cdot \frac{a_{ij}}{\alpha_{Lk}}\right]^2}$$

$$= C \sqrt{\sum^{i,j} \left[\alpha_{ji} - \alpha_{Lk}^{(ji)}\right]^2 a_{ij}^2}$$

$$\frac{\sqrt{\mathrm{var}\left[\sum_{L=1}^{6}\alpha_{L2}\right]}}{\sum_{L=1}^{6}\alpha_{L2}} = 0.6838k$$

$$\frac{\sqrt{\mathrm{var}\left[\sum_{L=1}^{6}\alpha_{L3}\right]}}{\sum_{L=1}^{6}\alpha_{L3}} = 0.7659k$$

$$\frac{\sqrt{\mathrm{var}\left[\sum_{L=1}^{6}\alpha_{L4}\right]}}{\sum_{L=1}^{6}\alpha_{L4}} = 0.7620k$$

$$\frac{\sqrt{\mathrm{var}\left[{}_{L=1}^{6}\ {}_{L5}\right]}}{\sum_{L=1}^{6}\alpha_{L5}} = 0.9682k$$

$$\frac{\sqrt{\mathrm{var}\left[\sum_{L=1}^{6}\alpha_{L6}\right]}}{\sum_{L=1}^{6}\alpha_{L6}} = 0.9925k$$

V. Effects of Coefficient Errors on Multipliers In Specific Input-Output Systems

The purpose of the present section is to develop hypothetical input-output systems, showing different stages of production, and to derive the derivatives, variances, and coefficients of variation of multipliers in each system. This undertaking is intended to provide for investigation of the sensitivity of multipliers to changes in coefficients under various assumptions and conditions.

A. Stages of Production With No Common Factor

Let us assume another system which is given by:

$X_1 = Y_1$

$X_2 = a_{21}X_1$

$X_3 = a_{32}X_2$

$X_4 = a_{43}X_3$

this indicates that all output of sector 1 goes to final demand, that output of sector 2 is used in sector 1, that output of sector 3 is used up by sector 2, and that all output of sector 4 is purchased by sector 3. Now, rearranging the system gives:

$$\begin{aligned}
(1 - 0)\ X_1 \qquad\qquad\qquad\qquad\qquad &= Y_1 \\
- a_{21}X_1 + (1 - 0)\ X_2 \qquad\qquad\qquad &= 0 \\
- a_{32}X_2 + (1 - 0)\ X_3 \qquad &= 0 \\
- a_{43}X_3 + (1 - 0)\ X_4 &= 0
\end{aligned}$$

In matrix form, these equations are set as follows:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ 0 & -a_{32} & 1 & 0 \\ 0 & 0 & -a_{43} & 1 \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} Y_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

<u>or</u>

$$[I - A] \cdot [X] = [Y]$$

In this case, only α_{L1}'s have meaning because all Yi's except Y1, are assumed to be zero.

1. <u>The Determinant and Cofactors</u>:

$|I - A| = \underline{1}$

$A_1^1 = 1$ $\quad A_2^1 = a_{21}$ $\quad A_3^1 = a_{21}a_{32}$ $\quad A_4^1 = a_{21}a_{32}a_{43}$

$A_1^2 = 0$ $\quad A_2^2 = 1$ $\quad A_3^2 = a_{32}$ $\quad A_4^2 = a_{32}a_{43}$

$A_1^3 = 0$ $\quad A_2^3 = 0$ $\quad A_3^3 = 1$ $\quad A_4^3 = a_{43}$

$A_1^4 = 0$ $\quad A_2^4 = 0$ $\quad A_3^4 = 0$ $\quad A_4^4 = 1$

2. <u>The Multipliers $\left[\alpha_{Lk}\right]$</u>:

$$\alpha_{11} = \frac{A_1^1}{|I-A|} = 1 \quad \alpha_{21} = \frac{A_2^1}{|I-A|} = a_{21} \quad \alpha_{31} = \frac{A_3^1}{|I-A|} = a_{21}a_{32} \quad \alpha_{41} = \frac{A_4^1}{|I-A|} = a_{21}a_{32}a_{43}$$

$$\alpha_{12} = \frac{A_1^2}{|I-A|} = 0 \quad \alpha_{22} = \frac{A_2^2}{|I-A|} = 1 \quad \alpha_{32} = \frac{A_3^2}{|I-A|} = a_{32} \quad \alpha_{42} = \frac{A_4^2}{|I-A|} = a_{32}a_{43}$$

$$\alpha_{13} = \frac{A_1^3}{|I-A|} = 0 \quad \alpha_{23} = \frac{A_2^3}{|I-A|} = 0 \quad \alpha_{33} = \frac{A_3^3}{|I-A|} = 1 \quad \alpha_{43} = \frac{A_4^3}{|I-A|} = a_{43}$$

$$\alpha_{14} = \frac{A_1^4}{|I-A|} = 0 \quad \alpha_{24} = \frac{A_2^4}{|I-A|} = 0 \quad \alpha_{34} = \frac{A_3^4}{|I-A|} = 0 \quad \alpha_{44} = \frac{A_4^4}{|I-A|} = 1$$

Only α_{11}, α_{21}, α_{31}, and α_{41} have meaning under the assumption that there is no common factor, because $Y_2 = Y_3 = Y_4 = 0$.

3. The Derivatives of the Multipliers:

$$\frac{\partial \alpha_{11}}{\partial a_{ij}} = 0 \qquad \frac{\partial \alpha_{22}}{\partial a_{ij}} = 0 \qquad \frac{\partial \alpha_{33}}{\partial a_{ij}} = 0$$

$$\frac{\partial \alpha_{12}}{\partial a_{ij}} = 0 \qquad \frac{\partial \alpha_{23}}{\partial a_{ij}} = 0 \qquad \frac{\partial \alpha_{34}}{\partial a_{ij}} = 0$$

$$\frac{\partial \alpha_{13}}{\partial a_{ij}} = 0 \qquad \frac{\partial \alpha_{24}}{\partial a_{ij}} = 0 \qquad \frac{\partial \alpha_{44}}{\partial a_{ij}} = 0$$

$$\frac{\partial \alpha_{14}}{\partial a_{ij}} = 0$$

$\dfrac{\partial \alpha_{21}}{\partial a_{ij}} = 0$, except when ij = 21.

$$\frac{\partial \alpha_{21}}{\partial a_{21}} = \frac{1}{|I-A|} - (\alpha_{21})(0) = \frac{1}{|I-A|} = \alpha_{21}\,\frac{1}{A_2^1} = \underline{1}$$

$\dfrac{\partial \alpha_{31}}{\partial a_{ij}} = 0$, except when ij = 21 and 32.

$$\frac{\partial \alpha_{31}}{\partial a_{21}} = \frac{a_{32}}{|I-A|} - (\alpha_{31})(0) = \frac{a_{32}}{|I-A|} = \alpha_{31}\,\frac{a_{32}}{A_3^1} = \underline{a_{32}}$$

$$\frac{\partial \alpha_{31}}{\partial a_{32}} = \frac{a_{21}}{|I-A|} - (\alpha_{31})(0) = \frac{a_{21}}{|I-A|} = \alpha_{31}\,\frac{a_{21}}{A_3^1} = \underline{a_{21}}$$

$\dfrac{\partial \alpha_{32}}{\partial a_{ij}} = 0$, except when ij = a_{32}.

$$\frac{\partial\alpha_{32}}{\partial a_{32}} = \frac{1}{|I-A|} - (\alpha_{32})(0) = \frac{1}{|I-A|} = \alpha_{32}\,\frac{1}{A_3^2} = \underline{1}$$

$\frac{\partial\alpha_{41}}{\partial a_{ij}} = 0$, except when ij = 21, 32, and 43.

$$\frac{\partial\alpha_{41}}{\partial a_{21}} = \frac{a_{32}a_{43}}{|I-A|} - (\alpha_{41})(0) = \frac{a_{32}a_{43}}{|I-A|} = \alpha_{41}\,\frac{a_{32}a_{43}}{A_4^1} = \underline{a_{32}a_{43}}$$

$$\frac{\partial\alpha_{41}}{\partial a_{32}} = \frac{a_{21}a_{43}}{|I-A|} - (\alpha_{41})(0) = \frac{a_{21}a_{43}}{|I-A|} = \alpha_{41}\,\frac{a_{21}a_{43}}{A_4^1} = \underline{a_{21}a_{43}}$$

$$\frac{\partial\alpha_{41}}{\partial a_{43}} = \frac{a_{21}a_{32}}{|I-A|} - (\alpha_{41})(0) = \frac{a_{21}a_{32}}{|I-A|} = \alpha_{41}\,\frac{a_{21}a_{32}}{A_4^1} = \underline{a_{21}a_{32}}$$

$\frac{\partial\alpha_{42}}{\partial a_{ij}} = 0$, except when ij = 32 and 43.

$$\frac{\partial\alpha_{42}}{\partial a_{32}} = \frac{a_{43}}{|I-A|} - (\alpha_{42})(0) = \frac{a_{43}}{|I-A|} = \alpha_{42}\,\frac{a_{43}}{A_4^2} = \underline{a_{43}}$$

$$\frac{\partial\alpha_{42}}{\partial a_{43}} = \frac{a_{32}}{|I-A|} - (\alpha_{42})(0) = \frac{a_{32}}{|I-A|} = \alpha_{42}\,\frac{a_{32}}{A_4^2} = \underline{a_{32}}$$

$\frac{\partial\alpha_{43}}{\partial a_{ij}} = 0$, except when ij = 43.

$$\frac{\partial\alpha_{43}}{\partial a_{43}} = \frac{1}{|I-A|} - (\alpha_{43})(0) = \frac{1}{|I-A|} = \alpha_{43}\,\frac{1}{A_4^3} = \underline{1}$$

4. <u>Variances of the Multipliers</u>:

$$\mathrm{var}(\alpha_{Lk}) = \left[\sum^{ij} \frac{\partial \alpha_{Lk}}{\partial a_{ij}} (\hat{a}_{ij} - a_{ij})\right]^2$$

$$\mathrm{var}(\alpha_{11}) = 0$$

$$\mathrm{var}(\alpha_{12}) = 0$$

$$\mathrm{var}(\alpha_{13}) = 0$$

$$\mathrm{var}(\alpha_{14}) = 0$$

$$\mathrm{var}(\alpha_{21}) = \left[(1)(\hat{a}_{21} - a_{21})\right]^2 = \sigma a_{21} = \alpha_{21}^2 \frac{\sigma_{21}}{a_{21}^2}$$

$$\mathrm{var}(\alpha_{22}) = 0$$

$$\mathrm{var}(\alpha_{23}) = 0$$

$$\mathrm{var}(\alpha_{24}) = 0$$

$$\mathrm{var}(\alpha_{31}) = \left[a_{32}(\hat{a}_{21} - a_{21}) + a_{21}(\hat{a}_{32} - a_{32})\right]^2 = \alpha_{31}^2 \left[\frac{(\hat{a}_{21} - a_{21})}{a_{21}} + \frac{(\hat{a}_{32} - a_{32})}{a_{32}}\right]^2$$

$$= \alpha_{31}^2 \left[\frac{\sigma_{21}}{a_{21}^2} + 2\frac{\sigma_{21,32}}{a_{21}a_{32}} + \frac{\sigma_{32}}{a_{32}^2}\right]$$

$$\mathrm{var}(\alpha_{32}) = \left[(1)(\hat{a}_{32} - a_{32})\right]^2 = \sigma a_{32} \qquad = \alpha_{32}^2 \frac{\sigma_{32}}{a_{32}^2}$$

$$\mathrm{var}(\alpha_{33}) = 0$$

$$\mathrm{var}(\alpha_{34}) = 0$$

$$\mathrm{var}(\alpha_{41}) = \left[a_{32}a_{43}(\hat{a}_{21} - a_{21}) + a_{21}a_{43}(\hat{a}_{32} - a_{32}) + a_{21}a_{32}(\hat{a}_{43} - a_{43})\right]^2$$

$$= \alpha_{41}^2 \left[\frac{(\hat{a}_{21} - a_{21})}{a_{21}} + \frac{(\hat{a}_{32} - a_{32})}{a_{32}} + \frac{(\hat{a}_{43} - a_{43})}{a_{43}}\right]^2$$

$$\text{var}(\alpha_{42}) = \left[a_{43}(\hat{a}_{32} - a_{32}) + a_{32}(\hat{a}_{43} - a_{43})\right]^2 = \alpha_{42}^2 \left[\frac{(\hat{a}_{32} - a_{32})}{a_{32}} + \frac{(\hat{a}_{43} - a_{43})}{a_{43}}\right]^2$$

$$\text{var}(\alpha_{43}) = \left[(1)(\hat{a}_{43} - a_{43})\right]^2 = \sigma a_{43} = \alpha_{43}^2 \frac{\sigma_{43}}{a_{43}^2}$$

$$\text{var}(\alpha_{44}) = 0$$

<u>If we assume that</u> $\frac{(\hat{a}_{ij} - a_{ij})}{a_{ij}} = \frac{(\hat{a}_{rs} - a_{rs})}{a_{rs}} = K$, then:

$$\text{var}(\alpha_{21}) = \alpha_{21}^2 K^2$$

$$\text{var}(\alpha_{31}) = \alpha_{31}^2 K^2(2)^2 = 4\alpha_{31}^2 K^2$$

$$\text{var}(\alpha_{32}) = \alpha_{32}^2 K^2$$

$$\text{var}(\alpha_{41}) = \alpha_{41}^2 K^2(3)^2 = 9\alpha_{41}^2 K^2$$

$$\text{var}(\alpha_{42}) = \alpha_{42}^2 K^2(2)^2 = 4\alpha_{42}^2 K^2$$

$$\text{var}(\alpha_{43}) = \alpha_{43}^2 K^2$$

5. <u>Coefficients of Variation of the Multipliers</u>.

$$\frac{\text{var}(\alpha_{Lk})}{\alpha_{Lk}^2} = \frac{1}{\alpha_{Lk}^2}\left[\alpha_{Lk} \sum^{ij} (\alpha_{ji} + \alpha_{Lk}^{(ji)}) (\hat{a}_{ij} - a_{ij})\right]^2 = \left[\sum^{ij} (\alpha_{ji} + \alpha_{Lk}^{(ji)})(\hat{a}_{ij} - a_{ij})\right]^2$$

Under the assumption that $(\hat{a}_{ij} - a_{ij}) = K\, a_{ij}$ and that all k's are equal,

we get:

$$\frac{\mathrm{var}(\alpha_{Lk})}{\alpha_{Lk}^2} = K^2 \left[\sum^{ij} (\alpha_{ji} + \alpha_{Lk}^{(ji)}) \, a_{ij} \right]^2 \text{, and:}$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{Lk})}{\alpha_{Lk}^2}} = K \sqrt{\left[\sum^{ij} (\alpha_{ji} + \alpha_{Lk}^{(ji)}) \, a_{ij} \right]^2}$$

Applying this last equation, we find that:

$$\sqrt{\frac{\mathrm{var}(\alpha_{21})}{\alpha_{21}^2}} = K$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{31})}{\alpha_{31}^2}} = 2K$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{32})}{\alpha_{32}^2}} = K$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{41})}{\alpha_{41}^2}} = 3K$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{42})}{\alpha_{42}^2}} = 2K$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{43})}{\alpha_{43}^2}} = K$$

B. Stages of Production With Common Factor

Suppose we have a system whose data are represented by the equations:

$$X_1 = a_{14}X_4 + Y_1$$
$$X_2 = a_{21}X_1 + a_{24}X_4$$
$$X_3 = a_{32}X_2 + a_{34}X_4$$
$$X_4 = a_{41}X_1 + a_{42}X_2 + a_{43}X_3$$

Rearranging these simultaneous equations yields:

$$\begin{array}{rcl} X_1 \qquad\qquad\qquad\qquad - a_{14}X_4 &=& Y_1 \\ - a_{21}X_1 + X_2 \qquad\qquad - a_{24}X_4 &=& 0 \\ - a_{32}X_2 + X_3 - a_{34}X_4 &=& 0 \\ - a_{41}X_1 - a_{42}X_2 - a_{43}X_3 + X_4 &=& 0 \end{array}$$

Putting them in matrix form gives:

$$\begin{bmatrix} 1 & 0 & 0 & -a_{14} \\ -a_{21} & 1 & 0 & -a_{24} \\ 0 & -a_{32} & 1 & -a_{34} \\ -a_{41} & -a_{42} & -a_{43} & 1 \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} Y_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

or $\quad [I - A] \cdot [X] = [Y]$

Since Y_2, Y_3, and Y_4 are assumed to be zero, only α_{11}, α_{21}, α_{31}, and α_{41} have meaning. Therefore, we will find only α_{L1}'s, their variances and their derivatives.

1. The Determinant of $[I-A]$ and the Cofactors of the First Row:

$$|I-A| = 1 - a_{32}a_{43}a_{24} - a_{24}a_{42} - a_{34}a_{43} - a_{21}a_{32}a_{43}a_{14} - a_{41}a_{14} - a_{42}a_{21}a_{14}$$

$$A_1^1 = 1 - a_{32}a_{43}a_{24} - a_{24}a_{42} - a_{34}a_{43}$$

$$A_2^1 = a_{21} + a_{24}a_{41} - a_{34}a_{43}a_{21}$$

$$A_3^1 = a_{21}a_{32} + a_{34}a_{41} + a_{24}a_{32}a_{41} + a_{34}a_{42}a_{21}$$

$$A_4^1 = a_{21}a_{32}a_{43} + a_{41} + a_{42}a_{21}$$

2. The Multipliers $\left[\alpha_{L1}\right]$:

$$\alpha_{11} = \frac{A_1^1}{|I-A|} = \frac{1 - a_{32}a_{43}a_{24} - a_{24}a_{42} - a_{34}a_{43}}{|I-A|}$$

$$\alpha_{21} = \frac{A_2^1}{|I-A|} = \frac{a_{21} + a_{24}a_{41} - a_{34}a_{43}a_{21}}{|I-A|}$$

$$\alpha_{31} = \frac{A_3^1}{|I-A|} = \frac{a_{21}a_{32} + a_{34}a_{41} + a_{24}a_{32}a_{41} + a_{34}a_{42}a_{21}}{|I-A|}$$

$$\alpha_{41} = \frac{A_4^1}{|I-A|} = \frac{a_{21}a_{32}a_{43} + a_{41} + a_{42}a_{21}}{|I-A|}$$

3. Derivatives of $\left[\alpha_{L1}\right]$:

$$\frac{\partial\alpha_{11}}{\partial a_{11}} = 0 \qquad \frac{\partial\alpha_{11}}{\partial a_{21}} = \alpha_{11}\alpha_{12}$$

$$\frac{\partial\alpha_{11}}{\partial a_{12}} = 0 \qquad \frac{\partial\alpha_{11}}{\partial a_{22}} = 0$$

$$\frac{\partial\alpha_{11}}{\partial a_{13}} = 0 \qquad \frac{\partial\alpha_{11}}{\partial a_{23}} = 0$$

$$\frac{\partial\alpha_{11}}{\partial a_{14}} = \alpha_{11}\alpha_{41} \qquad \frac{\partial\alpha_{11}}{\partial a_{24}} = \alpha_{11}\left[\alpha_{42} - \frac{a_{32}a_{43} + a_{42}}{A_1^1}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{31}} = 0 \qquad \frac{\partial\alpha_{11}}{\partial a_{41}} = \alpha_{11}\alpha_{14}$$

$$\frac{\partial\alpha_{11}}{\partial a_{32}} = \alpha_{11}\left[\alpha_{23} - \frac{a_{43}a_{24}}{A_1^1}\right] \qquad \frac{\partial\alpha_{11}}{\partial a_{42}} = \alpha_{11}\left[\alpha_{24} - \frac{a_{24}}{A_1^1}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{33}} = 0 \qquad \frac{\partial\alpha_{11}}{\partial a_{43}} = \alpha_{11}\left[\alpha_{34} - \frac{a_{32}a_{24} + a_{34}}{A_1^1}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{34}} = \alpha_{11}\left[\alpha_{43} - \frac{a_{43}}{A_1^1}\right] \qquad \frac{\partial\alpha_{11}}{\partial a_{44}} = 0$$

$$\frac{\partial\alpha_{21}}{\partial a_{11}} = 0 \qquad \frac{\partial\alpha_{21}}{\partial a_{21}} = \alpha_{21}\left[\alpha_{12} + \frac{1 - a_{34}a_{43}}{A_2^1}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{12}} = 0 \qquad \frac{\partial\alpha_{21}}{\partial a_{22}} = 0$$

$$\frac{\partial\alpha_{21}}{\partial a_{13}} = 0 \qquad \frac{\partial\alpha_{21}}{\partial a_{23}} = 0$$

$$\frac{\partial\alpha_{21}}{\partial a_{14}} = \alpha_{21}\alpha_{41} \qquad \frac{\partial\alpha_{21}}{\partial a_{24}} = \alpha_{21}\left[\alpha_{42} + \frac{a_{41}}{A_2^1}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{31}} = 0 \qquad \frac{\partial\alpha_{21}}{\partial a_{41}} = \alpha_{21}\left[\alpha_{14} + \frac{a_{24}}{A_2^1}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{32}} = \alpha_{21}\alpha_{23} \qquad \frac{\partial\alpha_{21}}{\partial a_{42}} = \alpha_{21}\alpha_{24}$$

$$\frac{\partial \alpha_{21}}{\partial a_{33}} = 0 \qquad \frac{\partial \alpha_{21}}{\partial a_{43}} = \alpha_{21}\left[\alpha_{34} - \frac{a_{34}a_{21}}{A_2^1}\right]$$

$$\frac{\partial \alpha_{21}}{\partial a_{34}} = \alpha_{21}\left[\alpha_{43} - \frac{a_{43}a_{21}}{A_2^1}\right] \qquad \frac{\partial \alpha_{21}}{\partial a_{44}} = 0$$

$$\frac{\partial \alpha_{31}}{\partial a_{11}} = 0 \qquad \frac{\partial \alpha_{31}}{\partial a_{21}} = \alpha_{31}\left[\alpha_{12} + \frac{a_{32} + a_{34}a_{42}}{A_3^1}\right]$$

$$\frac{\partial \alpha_{31}}{\partial a_{12}} = 0 \qquad \frac{\partial \alpha_{31}}{\partial a_{22}} = 0$$

$$\frac{\partial \alpha_{31}}{\partial a_{13}} = 0 \qquad \frac{\partial \alpha_{31}}{\partial a_{23}} = 0$$

$$\frac{\partial \alpha_{31}}{\partial a_{14}} = \quad 31\ 41 \qquad \frac{\partial \alpha_{31}}{\partial a_{24}} = \alpha_{31}\left[\alpha_{42} + \frac{a_{32}a_{41}}{A_3^1}\right]$$

$$\frac{\partial \alpha_{31}}{\partial a_{31}} = 0 \qquad \frac{\partial \alpha_{31}}{\partial a_{41}} = \alpha_{31}\left[\alpha_{14} + \frac{a_{34} + a_{24}d_{32}}{A_3^1}\right]$$

$$\frac{\partial \alpha_{31}}{\partial a_{32}} = \alpha_{31}\left[\alpha_{23} + \frac{a_{21} + a_{24}a_{41}}{A_3^1}\right] \qquad \frac{\partial \alpha_{31}}{\partial a_{42}} = \alpha_{31}\left[\alpha_{24} + \frac{a_{34}a_{21}}{A_3^1}\right]$$

$$\frac{\partial\alpha_{31}}{\partial a_{33}} = 0 \qquad \frac{\partial\alpha_{31}}{\partial a_{43}} = \alpha_{31}\alpha_{34}$$

$$\frac{\partial\alpha_{31}}{\partial a_{34}} = \alpha_{31}\left[\alpha_{43} + \frac{a_{41} + a_{42}a_{21}}{A_3^1}\right] \qquad \frac{\partial\alpha_{31}}{\partial a_{44}} = 0$$

$$\frac{\partial\alpha_{41}}{\partial a_{11}} = 0 \qquad \frac{\partial\alpha_{41}}{\partial a_{21}} = \alpha_{41}\left[\alpha_{12} + \frac{a_{32}a_{43} + a_{42}}{A_4^1}\right]$$

$$\frac{\partial\alpha_{41}}{\partial a_{12}} = 0 \qquad \frac{\partial\alpha_{41}}{\partial a_{22}} = 0$$

$$\frac{\partial\alpha_{41}}{\partial a_{13}} = 0 \qquad \frac{\partial\alpha_{41}}{\partial a_{23}} = 0$$

$$\frac{\partial\alpha_{41}}{\partial a_{14}} = \alpha_{41}^2 \qquad \frac{\partial\alpha_{41}}{\partial a_{24}} = \alpha_{41}\alpha_{42}$$

$$\frac{\partial\alpha_{41}}{\partial a_{31}} = 0 \qquad \frac{\partial\alpha_{41}}{\partial a_{41}} = \alpha_{41}\left[\alpha_{14} + \frac{1}{A_4^1}\right]$$

$$\frac{\partial\alpha_{41}}{\partial a_{32}} = \alpha_{41}\left[\alpha_{23} + \frac{a_{21}a_{43}}{A_4^1}\right] \qquad \frac{\partial\alpha_{41}}{\partial a_{42}} = \alpha_{41}\left[\alpha_{24} + \frac{a_{21}}{A_4^1}\right]$$

$$\frac{\partial\alpha_{41}}{\partial a_{33}} = 0 \qquad \frac{\partial\alpha_{41}}{\partial a_{43}} = \alpha_{41}\left[\alpha_{34} + \frac{a_{21}a_{32}}{A_4^1}\right]$$

$$\frac{\partial\alpha_{41}}{\partial a_{34}} = \alpha_{41}\alpha_{43} \qquad \frac{\partial\alpha_{41}}{\partial a_{44}} = 0$$

4. <u>The Variances of $\lfloor\alpha_{L1}\rfloor$</u>

$$\mathrm{var}(\alpha_{Lk}) = \left[\sum^{ij} \frac{\partial\alpha_{Lk}}{\partial a_{ij}} (\hat{a}_{ij} - a_{ij})\right]^2 = \alpha_{Lk}^2\left[\sum^{ij} (\alpha_{ji} + \alpha_{Lk}^{(ji)})(\hat{a}_{ij} - a_{ij})\right]^2,$$

where $\alpha_{Lk}^{(ji)} = \dfrac{A_{Lj}^{ki}}{A_L^k}$

$$\mathrm{var}(\alpha_{11}) = \alpha_{11}^2\Bigg[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{12}\sqrt{\sigma_{21}} + \alpha_{42}\sqrt{\sigma_{24}} - \frac{(a_{32}a_{43} + a_{42})}{A_1^1}\sqrt{\sigma_{24}} + \alpha_{23}\sqrt{\sigma_{32}}$$

$$- \frac{a_{43}a_{24}}{A_1^1}\sqrt{\sigma_{32}} + \alpha_{43}\sqrt{\sigma_{34}} - \frac{a_{43}}{A_1^1}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}$$

$$+ \quad {}_{24}\sqrt{\sigma_{42}} - \frac{a_{24}}{A_1^1}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} - \frac{(a_{32}a_{24} + a_{34})}{A_1^1}\sqrt{\sigma_{43}}\Bigg]^2$$

$$\mathrm{var}(\alpha_{21}) = \alpha_{21}^2\Bigg[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{12}\sqrt{\sigma_{21}} + \frac{(1 - a_{34}a_{43})}{A_2^1}\sqrt{\sigma_{21}} + \alpha_{42}\sqrt{\sigma_{24}} + \frac{a_{41}}{A_2^1}\sqrt{\sigma_{24}}$$

$$+ \alpha_{23}\sqrt{\sigma_{32}} + \alpha_{43}\sqrt{\sigma_{34}} - \frac{a_{43}a_{21}}{A_2^1}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}$$

$$+ \frac{a_{24}}{A_2^1}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} - \frac{a_{34}a_{21}}{A_2^1}\sqrt{\sigma_{43}}\Bigg]^2$$

$$\mathrm{var}(\alpha_{31}) = \alpha_{31}^2\Bigg[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{12}\sqrt{\sigma_{21}} + \frac{(a_{32} + a_{34}a_{42})}{A_3^1}\sqrt{\sigma_{21}} + \alpha_{42}\sqrt{\sigma_{24}} + \frac{a_{32}a_{41}}{A_3^1}$$

$$\sqrt{\sigma_{24}} + \alpha_{23}\sqrt{\sigma_{32}} + \alpha_{43}\sqrt{\sigma_{34}} + \frac{(a_{21} + a_{24}a_{41})}{A_3^1}\sqrt{\sigma_{32}}$$

$$+ \frac{(a_{41} + a_{42}a_{21})}{A_3^1}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} + \frac{(a_{34} + a_{24}a_{32})}{A_3^1}\sqrt{\sigma_{41}}$$

$$\left. + \alpha_{24}\sqrt{\sigma_{42}} + \frac{a_{34}a_{21}}{A_3^1}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\right]^2$$

$$\mathrm{var}(\alpha_{41}) = \alpha_{41}^2\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{12}\sqrt{\sigma_{21}} + \frac{(a_{32}a_{43} + a_{42})}{A_4^1}\sqrt{\sigma_{21}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{23}\sqrt{\sigma_{32}}\right.$$

$$+ \frac{a_{21}a_{43}}{A_4^1}\sqrt{\sigma_{32}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} + \frac{1}{A_4^1}\sqrt{\sigma_{41}}$$

$$\left. + \alpha_{24}\sqrt{\sigma_{42}} + \frac{a_{21}}{A_4^1}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} + \frac{a_{21}a_{32}}{A_4^1}\sqrt{\sigma_{43}}\right]^2$$

If $\frac{\sqrt{\sigma_{ij}}}{a_{ij}} = \frac{\sqrt{\sigma_{rs}}}{a_{rs}} = K$, and if all K's are equal, then:

$$\mathrm{var}(\alpha_{11}) = \alpha_{11}^2 K^2\left[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41}\right.$$

$$\left. + \alpha_{24}a_{42} + \alpha_{34}a_{43} + \frac{2A_1^1 - 2 - a_{32}a_{24}a_{43}}{A_1^1}\right]^2$$

$$\mathrm{var}(\alpha_{21}) = \alpha_{21}^2 K^2\left[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41}\right.$$

$$\left. + \alpha_{24}a_{42} + \alpha_{34}a_{43} + 2 - \frac{a_{21}(1 + a_{34}a_{43})}{A_2^1}\right]^2$$

$$\mathrm{var}(\alpha_{31}) = \alpha_{31}^2 K^2 \left[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41} + \alpha_{24}a_{42} + \alpha_{34}a_{43} + 3 - \frac{a_{21}a_{32} + a_{34}a_{41}}{A_3^1}\right]^2$$

$$\mathrm{var}(\alpha_{41}) = \alpha_{41}^2 K^2 \left[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41} + \alpha_{24}a_{42} + \alpha_{34}a_{43} + 3 - \frac{2a_{41} + a_{42}a_{21}}{A_4^1}\right]^2$$

5. <u>Coefficients of Variations of $\lfloor \alpha_{L1} \rfloor$:</u>

$$\sqrt{\frac{\mathrm{var}(\alpha_{11})}{\alpha_{11}^2}} = K\left[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41} + \alpha_{24}a_{42} + \alpha_{34}a_{43} + 3 - \frac{(3 + a_{24}a_{42} + a_{34}a_{43})}{A_1^1}\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{21})}{\alpha_{21}^2}} = K\left[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41} + \alpha_{24}a_{42} + \alpha_{34}a_{43} + 3 - \frac{(2a_{21} + a_{24}a_{41})}{A_2^1}\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{31})}{\alpha_{31}^2}} = K\Bigg[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41} + \alpha_{24}a_{42} + \alpha_{34}a_{43} + 3 - \frac{(a_{21}a_{32} + a_{34}a_{41})}{A_3^1}\Bigg]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{41})}{\alpha_{41}^2}} = K\Bigg[\alpha_{41}a_{14} + \alpha_{12}a_{21} + \alpha_{42}a_{24} + \alpha_{23}a_{32} + \alpha_{43}a_{34} + \alpha_{14}a_{41} + \alpha_{24}a_{42} + \alpha_{34}a_{43} + 3 - \frac{(2a_{41} + a_{42}a_{21})}{A_3^1}\Bigg]$$

C. No Internal Sales or Purchases Except for a Common Factor.

The following equations represent data of third hypothetical system:

$$X_1 = a_{14}X_4 + Y_1$$

$$X_2 = a_{24}X_4 + Y_2$$

$$X_3 = a_{34}X_4 + Y_3$$

$$X_4 = a_{41}X_1 + a_{42}X_2 + a_{43}X_3 + Y_4$$

Here, outputs of the sectors 1, 2, and 3 are distributed or used by sector 4 and by the final consumers, whereas output of sector 4 goes to the first 3 sectors and to final demand.

$$
\begin{aligned}
X_1 \qquad\qquad\qquad\qquad\qquad\qquad - a_{14}X_4 &= Y_1 \\
+ \quad X_2 \qquad\qquad\qquad - a_{24}X_4 &= Y_2 \\
+ \quad X_3 - a_{34}X_4 &= Y_3 \\
- a_{41}X_1 - a_{42}X_2 - a_{43}X_3 + \quad X_4 &= Y_4
\end{aligned}
$$

Setting the system in matrix form gives:

$$
\begin{bmatrix} 1 & 0 & 0 & -a_{14} \\ 0 & 1 & 0 & -a_{24} \\ 0 & 0 & 1 & -a_{34} \\ -a_{41} & -a_{42} & -a_{43} & 1 \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix}
$$

<u>or</u> $\qquad [\underset{-}{I} - \underset{-}{A}] \cdot [\underset{-}{X}] = [\underset{-}{Y}]$

1. <u>Determinant and Cofactors</u>:

$$|\underset{-}{I} - \underset{-}{A}| = 1 - a_{24}a_{42} - a_{34}a_{43} - a_{14}a_{41}$$

$A_1^1 = 1 - a_{24}a_{42} - a_{34}a_{43}$ $\qquad$ $A_2^1 = a_{24}a_{41}$

$A_1^2 = a_{14}a_{42}$ $\qquad$ $A_2^2 = 1 - a_{14}a_{41} - a_{34}a_{43}$

$A_1^3 = a_{43}a_{14}$ $\qquad$ $A_2^3 = a_{24}a_{43}$

$A_1^4 = a_{14}$ $\qquad$ $A_2^4 = a_{24}$

$A_3^1 = a_{34}a_{41}$ $\qquad$ $A_4^1 = a_{41}$

$A_3^2 = a_{34}a_{42}$ $\qquad$ $A_4^2 = a_{42}$

$A_3^3 = 1 - a_{14}a_{41} - a_{24}a_{42}$ $\qquad$ $A_4^3 = a_{43}$

$A_3^4 = a_{34}$ $\qquad$ $A_4^4 = 1$

2. The Multipliers $\left[\alpha_{Lk}\right]$:

$$\alpha_{11} = \frac{1 - a_{24}a_{42} - a_{34}a_{43}}{|I - A|} \qquad \alpha_{21} = \frac{a_{24}a_{41}}{|I - A|}$$

$$\alpha_{12} = \frac{a_{14}a_{42}}{|I - A|} \qquad \alpha_{22} = \frac{1 - a_{14}a_{41} - a_{34}a_{43}}{|I - A|}$$

$$\alpha_{13} = \frac{a_{43}a_{14}}{|I - A|} \qquad \alpha_{23} = \frac{a_{24}a_{43}}{|I - A|}$$

$$\alpha_{14} = \frac{a_{14}}{|I - A|} \qquad \alpha_{24} = \frac{a_{24}}{|I - A|}$$

$$\alpha_{31} = \frac{a_{34}a_{41}}{|I - A|} \qquad \alpha_{41} = \frac{a_{41}}{|I - A|}$$

$$\alpha_{32} = \frac{a_{34}a_{42}}{|I - A|} \qquad \alpha_{42} = \frac{a_{42}}{|I - A|}$$

$$\alpha_{33} = \frac{1 - a_{14}a_{41} - a_{24}a_{42}}{|I - A|} \qquad \alpha_{43} = \frac{a_{43}}{|I - A|}$$

$$\alpha_{34} = \frac{a_{34}}{|I - A|} \qquad \alpha_{44} = \frac{1}{|I - A|}$$

3. Derivatives of $\left[\alpha_{Lk}\right]$:

$$\frac{\partial\alpha_{Lk}}{\partial a_{11}} = \frac{\partial\alpha_{Lk}}{\partial a_{12}} = \frac{\partial\alpha_{Lk}}{\partial a_{13}} = \frac{\partial\alpha_{Lk}}{\partial a_{21}} = \frac{\partial\alpha_{Lk}}{\partial a_{22}} = \frac{\partial\alpha_{Lk}}{\partial a_{23}} = \frac{\partial\alpha_{Lk}}{\partial a_{31}} = \frac{\partial\alpha_{Lk}}{\partial a_{32}} = \frac{\partial\alpha_{Lk}}{\partial a_{33}} = \frac{\partial\alpha_{Lk}}{\partial a_{44}} = 0$$

$$\frac{\partial\alpha_{11}}{\partial a_{14}} = \alpha_{11}\alpha_{41} \qquad \frac{\partial\alpha_{12}}{\partial a_{14}} = \alpha_{12}\left[\alpha_{41} + \frac{a_{42}}{A_1^2}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{24}} = \alpha_{11}\left[\alpha_{42} - \frac{a_{42}}{A_1^1}\right] \qquad \frac{\partial\alpha_{12}}{\partial a_{24}} = \alpha_{12}\alpha_{42}$$

$$\frac{\partial\alpha_{11}}{\partial a_{34}} = \alpha_{11}\left[\alpha_{43} - \frac{a_{43}}{A_1^1}\right] \qquad \frac{\partial\alpha_{12}}{\partial a_{34}} = \alpha_{12}\,\alpha_{43}$$

$$\frac{\partial\alpha_{11}}{\partial a_{41}} = \alpha_{11}\alpha_{14} \qquad \frac{\partial\alpha_{12}}{\partial a_{41}} = \alpha_{12}\alpha_{14}$$

$$\frac{\partial\alpha_{11}}{\partial a_{42}} = \alpha_{11}\left[\alpha_{24} - \frac{a_{24}}{A_1^1}\right] \qquad \frac{\partial\alpha_{12}}{\partial a_{42}} = \alpha_{12}\left[\alpha_{24} + \frac{a_{14}}{A_1^2}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{43}} = \alpha_{11}\left[\alpha_{34} - \frac{a_{34}}{A_1^1}\right] \qquad \frac{\partial\alpha_{12}}{\partial a_{43}} = \alpha_{12}\alpha_{34}$$

$$\frac{\partial\alpha_{13}}{\partial a_{14}} = \alpha_{13}\left[\alpha_{41} + \frac{a_{43}}{A_1^3}\right] \qquad \frac{\partial\alpha_{14}}{\partial a_{14}} = \alpha_{14}\left[\alpha_{41} + \frac{1}{A_1^4}\right]$$

$$\frac{\partial\alpha_{13}}{\partial a_{24}} = \alpha_{13}\alpha_{42} \qquad \frac{\partial\alpha_{14}}{\partial a_{24}} = \alpha_{14}\alpha_{42}$$

$$\frac{\partial\alpha_{13}}{\partial a_{34}} = \alpha_{13}\alpha_{43} \qquad \frac{\partial\alpha_{14}}{\partial a_{34}} = \alpha_{14}\alpha_{43}$$

$$\frac{\partial\alpha_{13}}{\partial a_{41}} = \alpha_{13}\alpha_{14} \qquad \frac{\partial\alpha_{14}}{\partial a_{41}} = \alpha_{14}\alpha_{14}$$

$$\frac{\partial\alpha_{13}}{\partial a_{42}} = \alpha_{13}\alpha_{24} \qquad \frac{\partial\alpha_{14}}{\partial a_{42}} = \alpha_{14}\alpha_{24}$$

$$\frac{\partial\alpha_{13}}{\partial a_{43}} = \alpha_{13}\left[\alpha_{34} + \frac{a_{14}}{A_1^3}\right] \qquad \frac{\partial\alpha_{14}}{\partial a_{43}} = \alpha_{14}\alpha_{34}$$

$$\frac{\partial\alpha_{21}}{\partial a_{14}} = \alpha_{21}\alpha_{41} \qquad \frac{\partial\alpha_{22}}{\partial a_{14}} = \alpha_{22}\left[\alpha_{41} - \frac{a_{41}}{A_2^2}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{24}} = \alpha_{21}\left[\alpha_{42} + \frac{a_{41}}{A_2^1}\right] \qquad \frac{\partial\alpha_{22}}{\partial a_{24}} = \alpha_{22}\alpha_{42}$$

$$\frac{\partial\alpha_{21}}{\partial a_{34}} = \alpha_{21}\alpha_{43} \qquad \frac{\partial\alpha_{22}}{\partial a_{34}} = \alpha_{22}\left[\alpha_{43} - \frac{a_{43}}{A_2^2}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{41}} = \alpha_{21}\left[\alpha_{14} + \frac{a_{24}}{A_2^1}\right] \qquad \frac{\partial\alpha_{22}}{\partial a_{41}} = \alpha_{22}\left[\alpha_{14} - \frac{a_{14}}{A_2^2}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{42}} = \alpha_{21}\alpha_{24} \qquad \frac{\partial\alpha_{22}}{\partial a_{42}} = \alpha_{22}\alpha_{24}$$

$$\frac{\partial\alpha_{21}}{\partial a_{43}} = \alpha_{21}\alpha_{34} \qquad \frac{\partial\alpha_{22}}{\partial a_{43}} = \alpha_{22}\left[\alpha_{34} - \frac{a_{34}}{A_2^2}\right]$$

$$\frac{\partial\alpha_{23}}{\partial a_{14}} = \alpha_{23}\alpha_{41} \qquad \frac{\partial\alpha_{24}}{\partial a_{14}} = \alpha_{24}\alpha_{41}$$

$$\frac{\partial\alpha_{23}}{\partial a_{24}} = \alpha_{23}\left[\alpha_{42} + \frac{a_{43}}{A_2^3}\right] \qquad \frac{\partial\alpha_{24}}{\partial a_{24}} = \alpha_{24}\left[\alpha_{42} + \frac{1}{A_2^4}\right]$$

$$\frac{\partial\alpha_{23}}{\partial a_{34}} = \alpha_{23}\alpha_{43} \qquad \frac{\partial\alpha_{24}}{\partial a_{34}} = \alpha_{24}\alpha_{43}$$

$$\frac{\partial\alpha_{23}}{\partial a_{41}} = \alpha_{23}\alpha_{14} \qquad \frac{\partial\alpha_{24}}{\partial a_{41}} = \alpha_{24}\alpha_{14}$$

$$\frac{\partial\alpha_{23}}{\partial a_{42}} = \alpha_{23}\alpha_{24} \qquad \frac{\partial\alpha_{24}}{\partial a_{42}} = \alpha_{24}\alpha_{24}$$

$$\frac{\partial\alpha_{23}}{\partial a_{43}} = \alpha_{23}\left[\alpha_{34} + \frac{a_{24}}{A_2^3}\right] \qquad \frac{\partial\alpha_{24}}{\partial a_{43}} = \alpha_{24}\alpha_{34}$$

$$\frac{\partial\alpha_{31}}{\partial a_{14}} = \alpha_{31}\alpha_{41} \qquad \frac{\partial\alpha_{32}}{\partial a_{14}} = \alpha_{32}\alpha_{41}$$

$$\frac{\partial\alpha_{31}}{\partial a_{24}} = \alpha_{31}\alpha_{42} \qquad \frac{\partial\alpha_{32}}{\partial a_{24}} = \alpha_{32}\alpha_{42}$$

$$\frac{\partial\alpha_{31}}{\partial a_{34}} = \alpha_{31}\left[\alpha_{43} + \frac{a_{41}}{A_3^1}\right] \qquad \frac{\partial\alpha_{32}}{\partial a_{34}} = \alpha_{32}\left[\alpha_{43} + \frac{a_{42}}{A_3^2}\right]$$

$$\frac{\partial\alpha_{31}}{\partial a_{41}} = \alpha_{31}\left[\alpha_{14} + \frac{a_{34}}{A_3^1}\right] \qquad \frac{\partial\alpha_{32}}{\partial a_{41}} = \alpha_{32}\alpha_{14}$$

$$\frac{\partial\alpha_{31}}{\partial a_{42}} = \alpha_{31}\alpha_{24} \qquad \frac{\partial\alpha_{32}}{\partial a_{42}} = \alpha_{32}\left[\alpha_{24} + \frac{a_{34}}{A_3^2}\right]$$

$$\frac{\partial\alpha_{31}}{\partial a_{43}} = \alpha_{31}\alpha_{34} \qquad \frac{\partial\alpha_{32}}{\partial a_{43}} = \alpha_{32}\alpha_{34}$$

$$\frac{\partial\alpha_{33}}{\partial a_{14}} = \alpha_{33}\left[\alpha_{41} - \frac{a_{41}}{A_3^3}\right] \qquad \frac{\partial\alpha_{34}}{\partial a_{14}} = \alpha_{34}\alpha_{41}$$

$$\frac{\partial\alpha_{33}}{\partial a_{24}} = \alpha_{33}\left[\alpha_{42} - \frac{a_{42}}{A_3^3}\right] \qquad \frac{\partial\alpha_{34}}{\partial a_{24}} = \alpha_{34}\alpha_{42}$$

$$\frac{\partial\alpha_{33}}{\partial a_{34}} = \alpha_{33}\alpha_{43} \qquad \frac{\partial\alpha_{34}}{\partial a_{34}} = \alpha_{34}\left[\alpha_{43} + \frac{1}{A_3^4}\right]$$

$$\frac{\partial\alpha_{33}}{\partial a_{41}} = \alpha_{33}\left[\alpha_{14} - \frac{a_{14}}{A_3^3}\right] \qquad \frac{\partial\alpha_{34}}{\partial a_{41}} = \alpha_{34}\alpha_{14}$$

$$\frac{\partial\alpha_{33}}{\partial a_{42}} = \alpha_{33}\left[\alpha_{24} - \frac{a_{24}}{A_3^3}\right]. \qquad \frac{\partial\alpha_{34}}{\partial a_{42}} = \alpha_{34}\alpha_{24}$$

$$\frac{\partial\alpha_{33}}{\partial a_{43}} = \alpha_{33}\alpha_{34} \qquad \frac{\partial\alpha_{34}}{\partial a_{43}} = \alpha_{34}^2$$

$$\frac{\partial\alpha_{41}}{\partial a_{14}} = \alpha_{41}^2 \qquad \frac{\partial\alpha_{42}}{\partial a_{14}} = \alpha_{42}\alpha_{41}$$

$$\frac{\partial\alpha_{41}}{\partial a_{24}} = \alpha_{41}\alpha_{42} \qquad \frac{\partial\alpha_{42}}{\partial a_{24}} = \alpha_{42}^2$$

$$\frac{\partial\alpha_{41}}{\partial a_{34}} = \alpha_{41}\alpha_{43} \qquad \frac{\partial\alpha_{42}}{\partial a_{34}} = \alpha_{42}\alpha_{43}$$

$$\frac{\partial\alpha_{41}}{\partial a_{41}} = \alpha_{41}\left[\alpha_{14} + \frac{1}{A_4^1}\right] \qquad \frac{\partial\alpha_{42}}{\partial a_{41}} = \alpha_{42}\alpha_{14}$$

$$\frac{\partial\alpha_{41}}{\partial a_{42}} = \alpha_{41}\alpha_{24} \qquad \frac{\partial\alpha_{42}}{\partial a_{42}} = \alpha_{42}\left[\alpha_{24} + \frac{1}{A_4^2}\right]$$

$$\frac{\partial\alpha_{41}}{\partial a_{43}} = \alpha_{41}\alpha_{34} \qquad \frac{\partial\alpha_{42}}{\partial a_{43}} = \alpha_{42}\alpha_{34}$$

$$\frac{\partial\alpha_{43}}{\partial a_{14}} = \alpha_{43}\alpha_{41} \qquad \frac{\partial\alpha_{44}}{\partial a_{14}} = \alpha_{44}\alpha_{41}$$

$$\frac{\partial\alpha_{43}}{\partial a_{24}} = \alpha_{43}\alpha_{42} \qquad \frac{\partial\alpha_{44}}{\partial a_{24}} = \alpha_{44}\alpha_{42}$$

$$\frac{\partial\alpha_{43}}{\partial a_{34}} = \alpha_{43}^{2} \qquad \frac{\partial\alpha_{44}}{\partial a_{34}} = \alpha_{44}\alpha_{43}$$

$$\frac{\partial\alpha_{43}}{\partial a_{41}} = \alpha_{43}\alpha_{14} \qquad \frac{\partial\alpha_{44}}{\partial a_{41}} = \alpha_{44}\alpha_{14}$$

$$\frac{\partial\alpha_{43}}{\partial a_{42}} = \alpha_{43}\alpha_{24} \qquad \frac{\partial\alpha_{44}}{\partial a_{42}} = \alpha_{44}\alpha_{24}$$

$$\frac{\partial\alpha_{43}}{\partial a_{43}} = \alpha_{43}\left[\alpha_{34} + \frac{1}{A_4^3}\right] \qquad \frac{\partial\alpha_{44}}{\partial a_{43}} = \alpha_{44}\alpha_{34}$$

4. <u>Variances of $\left[\alpha_{Lk}\right]$:</u>

$$\mathrm{var}(\alpha_{11}) = \alpha_{11}^{2}\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} - \frac{a_{42}}{A_1^1}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} - \frac{a_{43}}{A_1^1}\sqrt{\sigma_{34}}\right.$$

$$\left. + \alpha_{14}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} - \frac{a_{24}}{A_1^1}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\right.$$

$$- \frac{a_{34}}{A_1^1}\sqrt{\sigma_{43}}\Bigg]^2$$

$$\mathrm{var}(\alpha_{12}) = \alpha_{12}^2\Bigg[\alpha_{41}\sqrt{\sigma_{14}} + \frac{a_{42}}{A_1^2}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}$$

$$+ \alpha_{24}\sqrt{\sigma_{42}} + \frac{a_{14}}{A_1^2}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\Bigg]^2$$

$$\mathrm{var}(\alpha_{13}) = \alpha_{13}^2\Bigg[\alpha_{41}\sqrt{\sigma_{14}} + \frac{a_{43}}{A_1^3}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}$$

$$+ \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} + \frac{a_{14}}{A_1^3}\sqrt{\sigma_{43}}\Bigg]^2$$

$$\mathrm{var}(\alpha_{14}) = \alpha_{14}^2\Bigg[\alpha_{41}\sqrt{\sigma_{14}} + \frac{1}{A_1^4}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}$$

$$+ \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\Bigg]^2$$

$$\mathrm{var}(\alpha_{21}) = \alpha_{21}^2\Bigg[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \frac{a_{41}}{A_2^1}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}$$

$$+ \frac{a_{24}}{A_2^1}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\Bigg]^2$$

$$\mathrm{var}(\alpha_{22}) = \alpha_{22}^2\left[\alpha_{41}\sqrt{\sigma_{14}} - \frac{a_{41}}{A_2^2}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} - \frac{a_{43}}{A_2^2}\sqrt{\sigma_{34}}\right.$$

$$+ \alpha_{14}\sqrt{\sigma_{41}} - \frac{a_{14}}{A_2^2}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}$$

$$\left. - \frac{a_{34}}{A_2^2}\sqrt{\sigma_{43}}\right]^2$$

$$\mathrm{var}(\alpha_{23}) = \alpha_{23}^2\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \frac{a_{43}}{A_2^3}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}\right.$$

$$\left. + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} + \frac{a_{24}}{A_2^3}\sqrt{\sigma_{43}}\right]^2$$

$$\mathrm{var}(\alpha_{24}) = \alpha_{24}^2\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \frac{1}{A_2^4}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}\right.$$

$$\left. + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\right]^2$$

$$\mathrm{var}(\alpha_{31}) = \alpha_{31}^2\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \frac{a_{41}}{A_3^1}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}}\right.$$

$$\left. + \frac{a_{34}}{A_3^1}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\right]^2$$

$$\mathrm{var}(\alpha_{32}) = \alpha_{32}^2 \left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \frac{a_{42}}{A_3^2}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} \right.$$

$$\left. + \alpha_{24}\sqrt{\sigma_{42}} + \frac{a_{34}}{A_3^2}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} \right]^2$$

$$\mathrm{var}(\alpha_{33}) = \alpha_{33}^2 \left[\alpha_{41}\sqrt{\sigma_{14}} - \frac{a_{41}}{A_3^3}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} - \frac{a_{42}}{A_3^3}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} \right.$$

$$+ \alpha_{14}\sqrt{\sigma_{41}} - \frac{a_{14}}{A_3^3}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} - \frac{a_{24}}{A_3^3}\sqrt{\sigma_{42}}$$

$$\left. + \alpha_{34}\sqrt{\sigma_{43}} \right]^2$$

$$\mathrm{var}(\alpha_{34}) = \alpha_{34}^2 \left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \frac{1}{A_3^4}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} \right.$$

$$\left. + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} \right]^2$$

$$\mathrm{var}(\alpha_{41}) = \alpha_{41}^2 \left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} + \frac{1}{A_4^1}\sqrt{\sigma_{41}} \right.$$

$$\left. + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} \right]^2$$

$$\mathrm{var}(\alpha_{42}) = \alpha_{42}^2\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} + \frac{1}{A_4^2}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\right]^2$$

$$\mathrm{var}(\alpha_{43}) = \alpha_{43}^2\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}} + \frac{1}{A_4^3}\sqrt{\sigma_{43}}\right]^2$$

$$\mathrm{var}(\alpha_{44}) = \alpha_{44}^2\left[\alpha_{41}\sqrt{\sigma_{14}} + \alpha_{42}\sqrt{\sigma_{24}} + \alpha_{43}\sqrt{\sigma_{34}} + \alpha_{14}\sqrt{\sigma_{41}} + \alpha_{24}\sqrt{\sigma_{42}} + \alpha_{34}\sqrt{\sigma_{43}}\right]^2$$

If we assume that $\frac{\sqrt{\sigma_{ij}}}{a_{ij}} = \frac{\sqrt{\sigma_{Rs}}}{a_{Rs}} = K$ and that all K's are equal, then:

$$\mathrm{var}(\alpha_{11}) = \alpha_{11}^2K^2\left[\frac{2}{|I - A|} - \frac{2}{A_1^1}\right]^2 = 4\alpha_{11}^2K^2\left[\frac{\alpha_{11} - 1}{A_1^1}\right]^2$$

$$\mathrm{var}(\alpha_{12}) = \alpha_{12}^2K^2\left[\frac{2}{|I - A|}\right]^2 = 4\alpha_{12}^2K^2\left[\frac{\alpha_{12}}{A_1^2}\right]^2$$

$$\mathrm{var}(\alpha_{13}) = \alpha_{13}^2K^2\left[\frac{2}{|I - A|}\right]^2 = 4\alpha_{13}^2K^2\left[\frac{\alpha_{13}}{A_1^3}\right]^2$$

$$\mathrm{var}(\alpha_{14}) = \alpha_{14}^2 K^2 \left[\frac{2}{|I - A|} - 1\right]^2 = \alpha_{14}^2 K^2 \left[\frac{2\alpha_{14}}{A_1^4} - 1\right]^2$$

$$\mathrm{var}(\alpha_{21}) = \alpha_{21}^2 K^2 \left[\frac{2}{|I - A|}\right]^2 = 4\alpha_{21}^2 K^2 \left[\frac{\alpha_{21}}{A_2^1}\right]^2$$

$$\mathrm{var}(\alpha_{22}) = \alpha_{22}^2 K^2 \left[\frac{2}{|I - A|} - \frac{2}{A_2^2}\right]^2 = 4\alpha_{22}^2 K^2 \left[\frac{\alpha_{22} - 1}{A_2^2}\right]^2$$

$$\mathrm{var}(\alpha_{23}) = \alpha_{23}^2 K^2 \left[\frac{2}{|I - A|}\right]^2 = 4\alpha_{23}^2 K^2 \left[\frac{\alpha_{23}}{A_2^3}\right]^2$$

$$\mathrm{var}(\alpha_{23}) = \alpha_{24}^2 K^2 \left[\frac{2}{|I - A|} - 1\right]^2 = \alpha_{24}^2 K^2 \left[\frac{2\alpha_{24}}{A_2^4} - 1\right]^2$$

$$\mathrm{var}(\alpha_{31}) = \alpha_{31}^2 K^2 \left[\frac{2}{|I - A|}\right]^2 = 4\alpha_{31}^2 K^2 \left[\frac{\alpha_{31}}{A_3^1}\right]^2$$

$$\mathrm{var}(\alpha_{32}) = \alpha_{32}^2 K^2 \left[\frac{2}{|I - A|}\right]^2 = 4\alpha_{32}^2 K^2 \left[\frac{\alpha_{32}}{A_3^2}\right]^2$$

$$\mathrm{var}(\alpha_{33}) = \alpha_{33}^2 K^2 \left[\frac{2}{|I - A|} - \frac{2}{A_3^3}\right]^2 = 4\alpha_{33}^2 K^2 \left[\frac{\alpha_{33} - 1}{A_3^3}\right]^2$$

$$\mathrm{var}(\alpha_{34}) = \alpha_{34}^2 K^2 \left[\frac{2}{|I - A|} - 1\right]^2 = \alpha_{34}^2 K^2 \left[\frac{2\alpha_{34}}{A_3^4} - 1\right]^2$$

$$\mathrm{var}(\alpha_{41}) = \alpha_{41}^2 K^2 \left[\frac{2}{|I - A|} - 1\right]^2 = \alpha_{41}^2 K^2 \left[\frac{2\alpha_{41}}{A_4^1} - 1\right]^2$$

$$\mathrm{var}(\alpha_{42}) = \alpha_{42}^2 K^2 \left[\frac{2}{|I - A|} - 1\right]^2 = \alpha_{42}^2 K^2 \left[\frac{2\alpha_{42}}{A_4^2} - 1\right]^2$$

$$\mathrm{var}(\alpha_{43}) = \alpha_{43}^2 K^2 \left[\frac{2}{|I - A|} - 1\right]^2 = \alpha_{43}^2 K^2 \left[\frac{2\alpha_{43}}{A_4^3} - 1\right]^2$$

$$\mathrm{var}(\alpha_{44}) = \alpha_{44}^2 K^2 \left[\frac{2}{|I - A|} - 2\right]^2 = 4\alpha_{44}^2 K^2 \left[\frac{\alpha_{44}}{A_4^4} - 1\right]^2$$

<u>Coefficients of Variations of $\left[\alpha_{Lk}\right]$:</u>

$$\sqrt{\frac{\mathrm{var}(\alpha_{11})}{\alpha_{11}^2}} = 2K\left[\frac{\alpha_{11} - 1}{A_1^1}\right] \qquad \sqrt{\frac{\mathrm{var}(\alpha_{12})}{\alpha_{12}^2}} = 2K\,\frac{\alpha_{12}}{A_1^2} = 2K\,\frac{1}{|I - A|}$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{13})}{\alpha_{13}^2}} = 2K\,\frac{\alpha_{13}}{A_1^3} = 2K\,\frac{1}{|I - A|} \qquad \sqrt{\frac{\mathrm{var}(\alpha_{14})}{\alpha_{14}^2}} = K\left[\frac{2\alpha_{14}}{A_1^4} - 1\right] = K\left[\frac{2}{|I - A|} - 1\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{21})}{\alpha_{21}^2}} = 2K\,\frac{\alpha_{21}}{A_2^1} = 2K\,\frac{1}{|I - A|} \qquad \sqrt{\frac{\mathrm{var}(\alpha_{22})}{\alpha_{22}^2}} = 2K\left[\frac{\alpha_{22} - 1}{A_2^2}\right] = 2K\left[\frac{1}{|I - A|} - \frac{1}{A_2^2}\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{23})}{\alpha_{23}^2}} = 2K\,\frac{\alpha_{23}}{A_2^3} = 2K\,\frac{1}{|I - A|} \qquad \sqrt{\frac{\mathrm{var}(\alpha_{24})}{\alpha_{24}^2}} = K\left[\frac{2\alpha_{24}}{A_2^4} - 1\right] = K\left[\frac{2}{|I - A|} - 1\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{31})}{\alpha_{31}^2}} = 2K\,\frac{\alpha_{31}}{A_3^1} = 2K\,\frac{1}{|I - A|} \qquad \sqrt{\frac{\mathrm{var}(\alpha_{32})}{\alpha_{32}^2}} = 2K\,\frac{\alpha_{32}}{A_3^2} = 2K\,\frac{1}{|I - A|}$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{33})}{\alpha_{33}^2}} = 2K\left[\frac{\alpha_{33} - 1}{A_3^3}\right] \qquad \sqrt{\frac{\mathrm{var}(\alpha_{34})}{\alpha_{34}^2}} = K\left[\frac{2\alpha_{34}}{A_3^4} - 1\right] = K\left[\frac{2}{|I - A|} - 1\right]$$

$$= 2K\left[\frac{1}{|I - A|} - \frac{1}{A_3^3}\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{41})}{\alpha_{41}^2}} = K\left[\frac{2\alpha_{41}}{A_4^1} - 1\right] \qquad \sqrt{\frac{\mathrm{var}(\alpha_{42})}{\alpha_{42}^2}} = K\left[\frac{2\alpha_{42}}{A_4^2} - 1\right] = K\left[\frac{2}{|I - A|} - 1\right]$$

$$= K\left[\frac{2}{|I - A|} - 1\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{43})}{\alpha_{43}^2}} = K\left[\frac{2\alpha_{43}}{A_4^3} - 1\right] \qquad \sqrt{\frac{\mathrm{var}(\alpha_{44})}{\alpha_4^4}} = 2K\left[\frac{\alpha_{44}}{A_4^4} - 1\right] = 2K\left[\frac{1}{|I - A|} - 1\right]$$

$$= K\left[\frac{2}{|I - A|} - 1\right]$$

D. Variances and Coefficients of Variation of the Multipliers in (3X3) Matrix where the values of the coefficients below the diagonal are zero:

$$(I - A) = \begin{bmatrix} (1 - a_{11}) & -a_{12} & -a_{13} \\ 0 & (1 - a_{22}) & -a_{23} \\ 0 & 0 & (1 - a_{33}) \end{bmatrix}$$

1. The Determinant and Cofactors.

$$|I - A| = (1 - a_{11})(1 - a_{22})(1 - a_{33})$$

$$A_1^1 = (1 - a_{22})(1 - a_{33})$$

$$A_1^2 = a_{12}(1 - a_{33})$$

$$A_1^3 = a_{12}a_{23} + a_{13}(1 - a_{22})$$

$$A_2^1 = 0$$

$$A_2^2 = (1 - a_{11})(1 - a_{33})$$

$$A_2^3 = a_{23}(1 - a_{11})$$

$$A_3^1 = 0$$

$$A_3^2 = 0$$

$$A_3^3 = (1 - a_{11})(1 - a_{22})$$

2. The Multipliers $\left[\alpha_{Lk}\right]$:

$$\alpha_{11} = \frac{A_1^1}{I - A} = \frac{(1 - a_{22})(1 - a_{33})}{(1 - a_{11})(1 - a_{22})(1 - a_{33})} = \frac{1}{(1 - a_{11})}$$

$$\alpha_{12} = \frac{A_1^1}{I - A} = \frac{a_{12}(1 - a_{33})}{(1 - a_{11})(1 - a_{22})(1 - a_{33})} = \frac{a_{12}}{(1 - a_{11})(1 - a_{22})}$$

$$\alpha_{12} = \frac{A_1^2}{|I - A|} = \frac{a_{12}(1 - a_{33})}{(1 - a_{11})(1 - a_{22})(1 - a_{33})} = \frac{a_{12}}{(1 - a_{11})(1 - a_{22})}$$

$$+ \frac{a_{13}}{(1 - a_{11})(1 - a_{33})} = \alpha_{12}\frac{a_{23}}{(1 - a_{33})} + \alpha_{11}\frac{a_{13}}{(1 - a_{33})}$$

$$= \alpha_{12}\alpha_{33}a_{23} + \alpha_{11}\alpha_{33}a_{13} = \alpha_{33}(\alpha_{12}a_{23} + \alpha_{11}a_{13})$$

$$\alpha_{21} = \frac{A_2^1}{|I - A|} = 0$$

$$\alpha_{22} = \frac{A_2^2}{|I - A|} = \frac{(1 - a_{11})(1 - a_{33})}{(1 - a_{11})(1 - a_{22})(1 - a_{33})} = \frac{1}{(1 - a_{22})}$$

$$\alpha_{23} = \frac{A_2^3}{|I - A|} = \frac{a_{23}(1 - a_{11})}{(1 - a_{11})(1 - a_{22})(1 - a_{33})} = \frac{a_{23}}{(1 - a_{22})(1 - a_{33})}$$

$$\alpha_{31} = \frac{A_3^1}{|I - A|} = 0$$

$$\alpha_{32} = \frac{A_3^2}{|I - A|} = 0$$

$$\alpha_{33} = \frac{A_3^3}{|I - A|} = \frac{(1 - a_{11})(1 - a_{22})}{(1 - a_{11})(1 - a_{22})(1 - a_{33})} = \frac{1}{(1 - a_{33})}$$

3. The Derivatives of the Multipliers and the Variances and the Coefficients of Variation of $\left[\alpha_{Lk}\right]$:

$$\frac{\partial \alpha_{11}}{\partial a_{11}} = \alpha_{11}^2$$

$$\frac{\partial \alpha_{11}}{\partial a_{12}} = \alpha_{11}\alpha_{21} = 0$$

$$\frac{\partial \alpha_{11}}{\partial a_{13}} = \alpha_{11}\alpha_{31} = 0$$

$$\frac{\partial \alpha_{11}}{\partial a_{21}} = 0$$

$$\frac{\partial \alpha_{11}}{\partial a_{22}} = -\frac{(1 - a_{33})}{A_1^1}\alpha_{11} + \alpha_{11}\alpha_{22} = 0$$

$$\frac{\partial \alpha_{11}}{\partial a_{23}} = \alpha_{11}\alpha_{32} = 0$$

$$\frac{\partial \alpha_{11}}{\partial a_{31}} = 0$$

$$\frac{\partial \alpha_{11}}{\partial a_{32}} = 0$$

$$\frac{\partial \alpha_{11}}{\partial a_{33}} = -\frac{(1 - a_{22})}{A_1^1}\alpha_{11} + \alpha_{11}\alpha_{33} = 0$$

$$\mathrm{var}(\alpha_{11}) = \left[\sum^{ij} \frac{\partial \alpha_{11}}{\partial a_{ij}} (\hat{a}_{ij} - a_{ij})\right]^2 = \left[\alpha_{11}^2(\hat{a}_{11} - a_{11}) + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0\right]^2$$

$$= \left[\alpha_{11}^2(\hat{a}_{11} - a_{11})\right]^2 = \alpha_{11}^4\sigma_{11} = \frac{\sigma_{11}}{(1 - a_{11})^4}, \text{ if } \sigma_{ij} = K^2 a_{ij}^2 \text{ and}$$

K^2 = constant, then

$$\sqrt{\frac{\mathrm{var}(\alpha_{11})}{\alpha_{11}^2}} = \sqrt{\frac{\alpha_{11}^4}{\alpha_{11}^2}\sigma_{11}} = \sqrt{\alpha_{11}^2\sigma_{11}} = \sqrt{\alpha_{11}^2 K^2 a_{11}^2} = \alpha_{11} K a_{11} = K\left[\frac{a_{11}}{(1 - a_{11})}\right]$$

$$\frac{\partial\alpha_{12}}{\partial a_{11}} = \alpha_{12}\alpha_{11} = \frac{a_{12}}{(1 - a_{11})^2(1 - a_{22})} = \alpha_{11}^2\alpha_{22}a_{12}$$

$$\frac{\partial\alpha_{12}}{\partial a_{12}} = \alpha_{12}\alpha_{21} + \alpha_{12}\frac{(1 - a_{33})}{A_1^2} = \frac{(1 - a_{33})}{(1 - a_{11})(1 - a_{22})(1 - a_{33})}$$

$$= \frac{1}{(1 - a_{11})(1 - a_{22})} = \alpha_{11}\alpha_{22} = \frac{\alpha_{12}}{a_{12}}$$

$$\frac{\partial\alpha_{12}}{\partial a_{13}} = \alpha_{12}\alpha_{31} = 0$$

$$\frac{\partial\alpha_{12}}{\partial a_{21}} = \alpha_{12}^2 = 0$$

$$\frac{\partial\alpha_{12}}{\partial a_{22}} = \alpha_{12}\alpha_{22} = \frac{a_{12}}{(1 - a_{11})(1 - a_{22})^2} = \alpha_{11}\alpha_{22}^2 a_{12}$$

$$\frac{\partial\alpha_{12}}{\partial a_{23}} = \alpha_{12}\alpha_{32} = 0$$

$$\frac{\partial\alpha_{12}}{\partial a_{31}} = \alpha_{12}\alpha_{13} = 0$$

$$\frac{\partial \alpha_{12}}{\partial a_{32}} = \alpha_{12}\alpha_{23} = 0$$

$$\frac{\partial \alpha_{12}}{\partial a_{33}} = \alpha_{12}\alpha_{33} - \frac{a_{12}}{|I - A|} = \frac{a_{12}}{(1 - a_{11})(1 - a_{22})(1 - a_{33})}$$

$$- \frac{a_{12}}{(1 - a_{11})(1 - a_{22})(1 - a_{33})} = 0$$

$$\text{var}(\alpha_{12}) = \left[\sum^{ij} \frac{\partial \alpha_{12}}{\partial a_{ij}} (\hat{a}_{ij} - a_{ij})\right]^2 = \left[\alpha_{12}\alpha_{11}(\hat{a}_{11} - a_{11}) + \alpha_{12}\frac{(1 - a_{33})}{A_1^2}\right.$$

$$\left.(\hat{a}_{12} - a_{12}) + \alpha_{12}\alpha_{22}(\hat{a}_{22} - a_{22})\right]^2$$

$$= \alpha_{12}^2\left[\alpha_{11}^2\sigma_{11} + 2\frac{\alpha_{11}}{a_{12}}\sigma_{11,12} + 2\alpha_{11}\alpha_{22}\sigma_{11,22} + \frac{1}{a_{12}^2}\sigma_{12}\right.$$

$$\left. + 2\frac{\alpha_{22}}{a_{12}}\sigma_{12,22} + \alpha_{22}^2\sigma_{22}\right]$$

$$= \alpha_{12}^2\left[\frac{\sqrt{\sigma_{11}}}{(1 - a_{11})} + \frac{\sqrt{\sigma_{12}}}{a_{12}} + \frac{\sqrt{\sigma_{22}}}{(1 - a_{22})}\right]^2$$

If $\sigma_{ij,RS} = 0$, except when $i = R$ and $j = S$, then:

$$= \alpha_{12}^2\left[\alpha_{11}^2\sigma_{11} + \frac{\sigma_{12}}{\alpha_{12}^2} + \alpha_{22}^2\sigma_{22}\right] = \frac{a_{12}^2}{(1 - a_{11})^2(1 - a_{22})^2}$$

$$\left[\frac{\sigma_{11}}{(1 - a_{11})^2} + \frac{\sigma_{12}}{a_{12}^2} + \frac{\sigma_{22}}{(1 - a_{22})^2}\right]$$

If $\sigma_{ij} = K^2 a_{ij}^2$, and K^2 is constant, then

$$= \frac{a_{12}^2}{(1 - a_{11})^2 (1 - a_{22})^2} \left[K^2\right] \left[\frac{a_{11}^2}{(1 - a_{11})^2} + \frac{a_{12}^2}{a_{12}^2} + \frac{a_{22}^2}{(1 - a_{22})^2}\right]$$

$$\sqrt{\frac{\text{var}(\alpha_{12})}{\alpha_{12}^2}} = K \sqrt{\left[\frac{a_{11}^2}{(1 - a_{11})^2} + \frac{a_{12}^2}{a_{12}^2} + \frac{a_{22}^2}{(1 - a_{22})^2}\right]}$$, If $\sigma_{ij,RS} \neq 0$, then

$$= K \sqrt{\frac{a_{11}^2}{(1 - a_{11})^2} + \frac{a_{12}^2}{a_{12}^2} + \frac{a_{22}^2}{(1 - a_{22})^2} + 2\frac{a_{11}a_{12}}{(1 - a_{11})a_{12}} + 2\frac{a_{11}a_{22}}{(1 - a_{11})(1 - a_{22})} + 2\frac{a_{12}a_{22}}{a_{12}(1 - a_{22})}}$$

$$= K \left[\frac{a_{11}}{(1 - a_{11})} + \frac{a_{12}}{a_{12}} + \frac{a_{22}}{(1 - a_{22})}\right]$$

$$\frac{\partial \alpha_{13}}{\partial a_{11}} = \alpha_{13}\alpha_{11}$$

$$\frac{\partial \alpha_{13}}{\partial a_{12}} = \alpha_{13}\alpha_{21} + \alpha_{13}\frac{a_{23}}{A_1^3} = \alpha_{13}\frac{a_{23}}{A_1^3} = \frac{a_{23}}{(1 - a_{11})(1 - a_{22})(1 - a_{33})}$$

$$\frac{\partial\alpha_{13}}{\partial a_{13}} = \alpha_{13}\alpha_{31} + \alpha_{13}\,\frac{(1 - a_{22})}{A_1^3} = \alpha_{13}\,\frac{(1 - a_{22})}{A_1^3} = \frac{1}{(1 - a_{11})(1 - a_{33})}$$

$$\frac{\partial\alpha_{13}}{\partial a_{21}} = 0$$

$$\frac{\partial\alpha_{13}}{\partial a_{22}} = \alpha_{13}\alpha_{22} - \alpha_{13}\,\frac{a_{13}}{A_1^3} = \alpha_{13}\left[\alpha_{22} - \frac{a_{13}}{A_1^3}\right]$$

$$\frac{\partial\alpha_{13}}{\partial a_{23}} = \alpha_{13}\alpha_{32} + \alpha_{13}\,\frac{a_{12}}{A_1^3} = \alpha_{13}\,\frac{a_{12}}{A_1^3} = \frac{a_{12}}{(1 - a_{11})(1 - a_{22})(1 - a_{33})}$$

$$\frac{\partial\alpha_{13}}{\partial a_{31}} = 0$$

$$\frac{\partial\alpha_{13}}{\partial a_{32}} = 0$$

$$\frac{\partial\alpha_{13}}{\partial a_{33}} = \alpha_{13}\alpha_{33}$$

$$\mathrm{var}(\alpha_{13}) = \left[\sum^{ij} \frac{13}{a_{ij}}\,(\hat{a}_{ij} - a_{ij})\right]^2 = \left[\alpha_{13}\alpha_{11}(\hat{a}_{11} - a_{11}) + \alpha_{13}\,\frac{a_{23}}{A_1^3}\,(\hat{a}_{12} - a_{12})\right.$$

$$\left. + \alpha_{13}\,\frac{(1 - a_{22})}{A_1^3}\,(\hat{a}_{13} - a_{13}) + \alpha_{13}(\alpha_{22} - \frac{a_{13}}{A_1^3})(\hat{a}_{22} - a_{22})\right.$$

$$+ \alpha_{13} \frac{a_{12}}{A_1^3}(\hat{a}_{23} - a_{23}) + \alpha_{13}\alpha_{33}(\hat{a}_{33} - a_{33})\Bigg]^2$$

$$= \alpha_{13}^2\left[\frac{\sqrt{\sigma_{11}}}{(1 - a_{11})} + \frac{a_{23}\sqrt{\sigma_{12}} + (1 - a_{22})\sqrt{\sigma_{13}} - a_{13}\sqrt{\sigma_{22}} + a_{12}\sqrt{\sigma_{23}}}{A_1^3}\right.$$

$$\left. + \frac{\sqrt{\sigma_{22}}}{(1 - a_{22})} + \frac{\sqrt{\sigma_{33}}}{(1 - a_{33})}\right]^2$$

$$= \alpha_{13}^2 K^2\left[\frac{a_{11}}{(1 - a_{11})} + \frac{a_{23}a_{12} + a_{13}(1 - a_{22})}{a_{23}a_{12} + a_{13}(1 - a_{22})} + \frac{a_{22}}{(1 - a_{22})}\right.$$

$$\left. - \frac{a_{13}a_{22}}{a_{23}a_{12} + a_{13}(1 - a_{22})} + \frac{a_{12}a_{23}}{a_{23}a_{12} + a_{13}(1 - a_{22})} + \frac{}{(1 - a_{33})}\right]^2$$

$$\frac{var(\alpha_{13})}{\alpha_{13}^2} = K\left[\frac{a_{11}}{(1 - a_{11})} + \frac{a_{22}}{(1 - a_{22})} + \frac{a_{33}}{(1 - a_{33})} + 1 + \frac{a_{12}a_{23} - a_{13}a_{22}}{a_{12}a_{23} + a_{13}(1 - a_{22})}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{ij}} = 0$$

$$var(\alpha_{21}) = 0$$

$$\frac{\partial\alpha_{22}}{\partial a_{11}} = \alpha_{22}\alpha_{11} - \alpha_{22}\frac{(1 - a_{33})}{A_2^2} = \alpha_{22}\left[\alpha_{11} - \frac{(1 - a_{33})}{A_2^2}\right]$$

$$= \alpha_{22}\left[\frac{1}{(1 - a_{11})} - \frac{(1 - a_{33})}{(1 - a_{11})(1 - a_{33})}\right] = 0$$

$$\frac{\partial \alpha_{22}}{\partial a_{12}} = \alpha_{22}\alpha_{21} = 0$$

$$\frac{\partial \alpha_{22}}{\partial a_{13}} = \alpha_{22}\alpha_{31} = 0$$

$$\frac{\partial \alpha_{22}}{\partial a_{21}} = 0$$

$$\frac{\partial \alpha_{22}}{\partial a_{22}} = \alpha_{22}^2$$

$$\frac{\partial \alpha_{22}}{\partial a_{23}} = \alpha_{22}\alpha_{32} = 0$$

$$\frac{\partial \alpha_{22}}{\partial a_{31}} = 0$$

$$\frac{\partial \alpha_{22}}{\partial a_{32}} = 0$$

$$\frac{\partial \alpha_{22}}{\partial a_{33}} = \alpha_{22}\alpha_{33} - \alpha_{22}\,\frac{(1 - a_{11})}{A_2^2} = \alpha_{22}\left[\frac{1}{(1 - a_{33})} - \frac{(1 - a_{11})}{(1 - a_{11})(1 - a_{33})}\right] = 0$$

$$\mathrm{var}(\alpha_{22}) = \left[\alpha_{22}^2(\hat{a}_{22} - a_{22})\right]^2 = \alpha_{22}^2\left[k^2\right]\left[\frac{a_{22}^2}{(1 - a_{22})^2}\right]$$

$$\sqrt{\frac{\mathrm{var}(\alpha_{22})}{\alpha_{22}^2}} = K\left[\frac{a_{22}}{(1 - a_{22})}\right]$$

$$\frac{\partial\alpha_{23}}{\partial a_{11}} = \alpha_{23}\alpha_{11} - \alpha_{23}\frac{a_{23}}{A_2^3} = \alpha_{23}\left[\frac{1}{(1 - a_{11})} - \frac{a_{23}}{a_{23}(1 - a_{11})}\right] = 0$$

$$\frac{\partial\alpha_{23}}{\partial a_{12}} = \alpha_{23}\alpha_{21} = 0$$

$$\frac{\partial\alpha_{23}}{\partial a_{13}} = \alpha_{23}\alpha_{31} = 0$$

$$\frac{\partial\alpha_{23}}{\partial a_{21}} = 0$$

$$\frac{\partial\alpha_{23}}{\partial a_{22}} = \alpha_{23}\alpha_{22}$$

$$\frac{\partial\alpha_{23}}{\partial a_{23}} = \alpha_{23}\alpha_{32} + \alpha_{23}\frac{(1 - a_{11})}{A_2^3} = \alpha_{23}\frac{(1 - a_{11})}{A_2^3} = \frac{1}{(1 - a_{22})(1 - a_{33})}$$

$$\frac{\partial\alpha_{23}}{\partial a_{31}} = 0$$

$$\frac{\partial\alpha_{23}}{\partial a_{32}} = 0$$

$$\frac{\partial\alpha_{23}}{\partial a_{33}} = \alpha_{23}\alpha_{33}$$

$$\mathrm{var}(\alpha_{23}) = \left[\alpha_{23}\alpha_{22}(\hat{a}_{22} - a_{22}) + \alpha_{23}\frac{(1 - a_{11})}{A_2^3}(\hat{a}_{23} - a_{23}) + \alpha_{23}\alpha_{33}(\hat{a}_{33} - a_{33})\right]^2$$

$$= \alpha_{23}^2 K^2 \left[\frac{a_{22}}{(1 - a_{22})} + \frac{a_{23}(1 - a_{11})}{a_{23}(1 - a_{11})} + \frac{a_{33}}{(1 - a_{33})} \right]^2$$

$$\sqrt{\frac{\overline{\mathrm{var}(\alpha_{23})}}{\alpha_{23}^2}} = K \left[\frac{a_{22}}{(1 - a_{22})} + \frac{a_{23}(1 - a_{11})}{a_{23}(1 - a_{11})} + \frac{a_{33}}{(1 - a_{33})} \right]$$

$$\frac{\partial \alpha_{31}}{\partial a_{ij}} = 0$$

$$\mathrm{var}(\alpha_{31}) = 0$$

$$\frac{\partial \alpha_{32}}{\partial a_{ij}} = 0$$

$$\mathrm{var}(\alpha_{32}) = 0$$

$$\frac{\partial \alpha_{33}}{\partial a_{11}} = \alpha_{33}\alpha_{11} - \alpha_{33} \frac{(1 - a_{22})}{A_3^3} = \alpha_{33} \left[\frac{1}{(1 - a_{11})} - \frac{(1 - a_{22})}{(1 - a_{11})(1 - a_{22})} \right] = 0$$

$$\frac{\partial \alpha_{33}}{\partial a_{12}} = \alpha_{33}\alpha_{21} = 0$$

$$\frac{\partial \alpha_{33}}{\partial a_{13}} = \alpha_{33}\alpha_{31} = 0$$

$$\frac{\partial \alpha_{33}}{\partial a_{21}} = 0$$

$$\frac{\partial \alpha_{33}}{\partial a_{22}} = \alpha_{33}\alpha_{22} - \alpha_{33}\frac{(1 - a_{11})}{A_3^3} = \alpha_{33}\left[\frac{1}{(1 - a_{22})} - \frac{(1 - a_{11})}{(1 - a_{11})(1 - a_{22})}\right] = 0$$

$$\frac{\partial \alpha_{23}}{\partial a_{23}} = \alpha_{33}\alpha_{32} = 0$$

$$\frac{\partial \alpha_{33}}{\partial a_{31}} = 0$$

$$\frac{\partial \alpha_{33}}{\partial a_{33}} = \alpha_{33}^2$$

$$\frac{\partial \alpha_{33}}{\partial a_{32}} = 0$$

$$\mathrm{var}(\alpha_{33}) = \alpha_{33}^2\left[\alpha_{33}(\hat{a}_{33} - a_{33})\right]^2 = \alpha_{33}^2 K^2\left[\frac{a_{33}}{(1 - a_{33})}\right]^2$$

$$\frac{\mathrm{var}(\alpha_{33})}{\alpha_{33}^2} = K\frac{a_{33}}{(1 - a_{33})}$$

E. Common Factor with Noncommon Intermediate Goods.

Let us assume the existence of a system whose data are represented by the simultaneous equations:

$$X_1 = a_{15}X_5 + Y_1$$
$$X_2 = a_{21}X_1$$
$$X_3 = a_{35}X_5 + Y_3$$
$$X_4 = a_{43}X_3$$
$$X_5 = a_{51}X_1 + a_{52}X_2 + a_{54}X_3 + a_{54}X_4 + a_{55}X_5 + Y_5$$

Rearranging them gives:

$$X_1 - a_{15}X_5 = Y_1$$
$$X_2 - a_{21}X_1 = 0$$
$$X_3 - a_{35}X_5 = Y_3$$
$$X_4 - a_{43}X_3 = 0$$
$$X_5 - a_{51}X_1 - a_{52}X_2 - a_{53}X_3 - a_{54}X_4 - a_{55}X_5 = Y_5.$$

Putting these equations in matrix form yields:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -a_{15} \\ -a_{21} & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -a_{35} \\ 0 & 0 & -a_{43} & 1 & 0 \\ -a_{51} & -a_{52} & -a_{53} & -a_{54} & (1-a_{55}) \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{bmatrix} = \begin{bmatrix} Y_1 \\ 0 \\ Y_3 \\ 0 \\ Y_5 \end{bmatrix}$$

We want to find the multipliers, their derivatives, their variances, and their coefficients of variations. To derive the multipliers, we should find the determinant of the matrix (I - A) and the cofactors of each element in it because $\alpha_{LK} = \frac{A_L^K}{|I - A|}$.

<u>The Determinant</u>.

$$|I - A| = 1 - a_{55} - a_{35}a_{43}a_{54} - a_{35}a_{53} - a_{15}a_{21}a_{52} - a_{15}a_{51}$$

<u>Cofactors</u>:

Since Y_2 and Y_4 are assumed to be zero, only α_{L1}'s, α_{L3}'s and α_{L5}'s have meaning. Therefore, we will find the cofactors of a_{1L}'s, a_{3L}'s, and a_{5L}'s which we call A_L^1's, and A_L^3's, and A_1^5's, respectively.

$A_1^1 = 1 - a_{55} - a_{35}a_{43}a_{54} - a_{35}a_{53}$

$A_2^1 = a_{21} - a_{21}a_{55} - a_{21}a_{43}a_{54}a_{35} - a_{21}a_{35}a_{53}$

$A_3^1 = a_{21}a_{35}a_{52} + a_{35}a_{51}$

$A_4^1 = a_{21}a_{35}a_{43}a_{52} + a_{35}a_{43}a_{51}$

$A_5^1 = a_{21}a_{52} + a_{51}$

$A_1^3 = a_{15}a_{43}a_{54} + a_{53}a_{15}$

$A_2^3 = a_{15}a_{21}a_{43}a_{54} + a_{15}a_{21}a_{53}$

$A_3^3 = 1 - a_{55} - a_{15}a_{51} - a_{15}a_{21}a_{52}$

$A_4^3 = a_{43} - a_{43}a_{55} - a_{15}a_{43}a_{51} - a_{15}a_{43}a_{52}a_{21}$

$A_5^3 = a_{43}a_{54} + a_{53}$

$A_1^5 = a_{15}$

$A_2^5 = a_{15}a_{21}$

$A_3^5 = a_{35}$

$A_4^5 = a_{35}a_{43}$

$A_5^5 = 1$

<u>Multipliers</u>:

$$\alpha_{11} = \frac{1 - a_{55} - a_{35}a_{43}a_{54} - a_{35}a_{53}}{1 - a_{55} - a_{35}a_{43}a_{54} - a_{35}a_{53} - a_{15}a_{21}a_{52} - a_{15}a_{51}} = \frac{A_1^1}{|I - A|}$$

$$\alpha_{21} = \frac{a_{21} - a_{21}a_{55} - a_{21}a_{43}a_{54}a_{35} - a_{21}a_{35}a_{53}}{|I - A|} = a_{21}\alpha_{11}$$

$$\alpha_{31} = \frac{a_{21}a_{35}a_{52} + a_{35}a_{51}}{|I - A|} = \frac{a_{35}A_5^1}{|I-A|} = a_{35}\alpha_{51} = A_3^5\alpha_{51}$$

$$\alpha_{41} = \frac{a_{21}a_{35}a_{52}a_{43} + a_{35}a_{51}a_{43}}{|I - A|} = \frac{a_{43}A_3^1}{|I - A|} = a_{43}\alpha_{31}$$

$$\alpha_{51} = \frac{a_{21}a_{52} + a_{51}}{|I - A|} = \frac{A_5^1}{|I - A|} = \alpha_{51}$$

$$\alpha_{13} = \frac{a_{15}a_{43}a_{54} + a_{53}}{|I - A|} = \frac{A_1^3}{|I - A|}$$

$$\alpha_{23} = \frac{a_{15}a_{21}a_{43}a_{54} + a_{15}a_{21}a_{53}}{|I - A|} = a_{21}a_{15}\frac{A_5^3}{|I - A|} = a_{21}a_{15}\alpha_{53} = A_2^5\alpha_{53}$$

$$\alpha_{33} = \frac{1 - a_{55} - a_{15}a_{51} - a_{15}a_{21}a_{52}}{|I - A|} = \frac{A_3^3}{|I - A|} = \alpha_{33}$$

$$\alpha_{43} = \frac{a_{43} - a_{43}a_{55} - a_{15}a_{51}a_{43} - a_{15}a_{21}a_{52}a_{43}}{|I - A|} = a_{43}\alpha_{33}$$

$$\alpha_{53} = \frac{a_{43}a_{54} + a_{53}}{|I - A|} = \frac{A_5^3}{|I - A|} = \alpha_{53}$$

$$\alpha_{15} = \frac{a_{15}}{|I - A|} = \frac{A_1^5}{|I - A|} = \alpha_{15}$$

$$\alpha_{25} = \frac{a_{15}a_{21}}{|I - A|} = \frac{A_2^5}{|I - A|} = a_{21}\frac{A_1^5}{|I - A|} = a_{21}\alpha_{15}$$

$$\alpha_{35} = \frac{a_{35}}{|I - A|} = \frac{A_3^5}{|I - A|} = \alpha_{35}$$

$$\alpha_{45} = \frac{a_{35}a_{43}}{|I - A|} = \frac{A_4^5}{|I - A|} = a_{43}\,\frac{A_3^5}{|I - A|} = a_{43}\alpha_{35}$$

$$\alpha_{55} = \frac{1}{|I - A|} = \frac{A_5^5}{|I - A|} = \alpha_{55}$$

Derivatives of Multipliers.

We assume that $\frac{\partial\alpha_{LK}}{\partial a_{ij}}$ where $a_{ij} = 0$ is equal to zero, and it will not be shown here. Only $\frac{\partial\alpha_{LK}}{\partial a_{ij}}$'s where $a_{ij} \neq 0$ are shown below:

$$\frac{\partial\alpha_{11}}{\partial a_{15}} = \alpha_{11}\alpha_{51}$$

$$\frac{\partial\alpha_{11}}{\partial a_{21}} = \alpha_{11}\alpha_{12}$$

$$\frac{\partial\alpha_{11}}{\partial a_{35}} = \alpha_{11}\alpha_{53} - \frac{a_{43}a_{54} + a_{53}}{|I - A|}$$

$$= \alpha_{11}\alpha_{53} - \alpha_{53}$$

$$\frac{\partial\alpha_{11}}{\partial a_{43}} = \alpha_{11}\alpha_{34} - \frac{a_{35}a_{54}}{|I - A|}$$

$$= \alpha_{11}\alpha_{34} - \alpha_{34}$$

$$= \alpha_{11}\left[\alpha_{34} - \frac{A_3^4}{A_1^1}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{51}} = \alpha_{11}\alpha_{15}$$

$$\frac{\partial\alpha_{11}}{\partial a_{52}} = \alpha_{11}\alpha_{25}$$

$$\frac{\partial\alpha_{11}}{\partial a_{53}} = \alpha_{11}\alpha_{35} - \frac{a_{35}}{|I - A|}$$

$$= \alpha_{11}\left[\alpha_{35} - \frac{A_3^5}{A_1^1}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{54}} = \alpha_{11}\left[\alpha_{45} - \frac{A_4^5}{A_1^1}\right]$$

$$\frac{\partial\alpha_{11}}{\partial a_{55}} = \alpha_{11}\left[\alpha_{55} - \frac{1}{A_1^1}\right]$$

$$= \alpha_{11}\left[\alpha_{55} - \frac{A_5^5}{A_1^1}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{15}} = \alpha_{21}\alpha_{51} \qquad \frac{\partial\alpha_{21}}{\partial a_{43}} = \alpha_{21}\left[\alpha_{34} - a_{21}\frac{A_3^4}{A_2^1}\right] \qquad \frac{\partial\alpha_{21}}{\partial a_{53}} = \alpha_{21}\left[\alpha_{35} - a_{21}\frac{A_3^5}{A_2^1}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{21}} = \alpha_{21}\left[\alpha_{12} + \frac{A_1^1}{A_2^1}\right] \qquad \frac{\partial\alpha_{21}}{\partial a_{51}} = \alpha_{21}\alpha_{15} \qquad \frac{\partial\alpha_{21}}{\partial a_{54}} = \alpha_{21}\left[\alpha_{45} - a_{21}\frac{A_4^5}{A_2^1}\right]$$

$$\frac{\partial\alpha_{21}}{\partial a_{35}} = \alpha_{21}\left[\alpha_{53} - a_{21}\frac{A_5^3}{A_2^1}\right] \qquad \frac{\partial\alpha_{21}}{\partial a_{52}} = \alpha_{21}\alpha_{25} \qquad \frac{\partial\alpha_{21}}{\partial a_{55}} = \alpha_{21}\left[\alpha_{55} - a_{21}\frac{A_5^5}{A_2^1}\right]$$

$$\frac{\partial\alpha_{31}}{\partial a_{15}} = \alpha_{31}\alpha_{51} \qquad \frac{\partial\alpha_{31}}{\partial a_{43}} = \alpha_{31}\alpha_{34} \qquad \frac{\partial\alpha_{31}}{\partial a_{53}} = \alpha_{31}\alpha_{35}$$

$$\frac{\partial\alpha_{31}}{\partial a_{21}} = \alpha_{31}\left[\alpha_{12} + \frac{A_3^2}{A_3^1}\right] \qquad \frac{\partial\alpha_{31}}{\partial a_{51}} = \alpha_{31}\left[\alpha_{15} + \frac{A_4^5}{A_3^1}\right] \qquad \frac{\partial\alpha_{31}}{\partial a_{54}} = \alpha_{31}\alpha_{45}$$

$$\frac{\partial\alpha_{31}}{\partial a_{35}} = \alpha_{31}\left[\alpha_{53} + \frac{A_5^1}{A_3^1}\right] \qquad \frac{\partial\alpha_{31}}{\partial a_{52}} = \alpha_{31}\left[\alpha_{25} + a_{21}\frac{A_3^5}{A_3^1}\right] \qquad \frac{\partial\alpha_{31}}{\partial a_{55}} = \alpha_{31}\alpha_{55}$$

$$\frac{\partial\alpha_{41}}{\partial a_{15}} = \alpha_{41}\alpha_{51} \qquad \frac{\partial\alpha_{41}}{\partial a_{43}} = \alpha_{41}\left[\alpha_{34} + \frac{A_3^1}{A_4^1}\right] \qquad \frac{\partial\alpha_{41}}{\partial a_{53}} = \alpha_{41}\alpha_{35}$$

$$\frac{\partial\alpha_{41}}{\partial a_{21}} = \alpha_{41}\left[\alpha_{12} + \frac{A_4^2}{A_4^1}\right] \qquad \frac{\partial\alpha_{41}}{\partial a_{51}} = \alpha_{41}\left[\alpha_{15} + \frac{A_4^5}{A_4^1}\right] \qquad \frac{\partial\alpha_{41}}{\partial a_{54}} = \alpha_{41}\alpha_{45}$$

$$\frac{\partial\alpha_{41}}{\partial a_{35}} = \alpha_{41}\left[\alpha_{53} + a_{43}\frac{A_5^1}{A_4^1}\right] \qquad \frac{\partial\alpha_{41}}{\partial a_{52}} = \alpha_{41}\left[\alpha_{25} + a_{21}\frac{A_4^5}{A_4^1}\right] \qquad \frac{\partial\alpha_{41}}{\partial a_{55}} = \alpha_{41}\alpha_{55}$$

$$\frac{\partial\alpha_{51}}{\partial a_{15}} = \alpha_{51}^2 \qquad \frac{\partial\alpha_{51}}{\partial a_{43}} = \alpha_{51}\alpha_{34} \qquad \frac{\partial\alpha_{51}}{\partial a_{53}} = \alpha_{51}\alpha_{35}$$

$$\frac{\partial\alpha_{51}}{\partial a_{21}} = \alpha_{51}\left[\alpha_{12} + \frac{A_5^2}{A_5^1}\right] \qquad \frac{\partial\alpha_{51}}{\partial a_{51}} = \alpha_{51}\left[\alpha_{15} + \frac{A_5^5}{A_5^1}\right] \qquad \frac{\partial\alpha_{51}}{\partial a_{54}} = \alpha_{51}\alpha_{45}$$

$$\frac{\partial\alpha_{51}}{\partial a_{35}} = \alpha_{51}\alpha_{53} \qquad \frac{\partial\alpha_{51}}{\partial a_{52}} = \alpha_{51}\left[\alpha_{25} + a_{21}\frac{A_5^5}{A_5^1}\right] \qquad \frac{\partial\alpha_{51}}{\partial a_{55}} = \alpha_{51}\alpha_{55}$$

$$\frac{\partial\alpha_{13}}{\partial a_{15}} = \alpha_{13}\left[\alpha_{51} + \frac{A_5^3}{A_1^3}\right] \qquad \frac{\partial\alpha_{13}}{\partial a_{43}} = \alpha_{13}\left[\alpha_{34} + \frac{A_1^4}{A_1^3}\right] \qquad \frac{\partial\alpha_{13}}{\partial a_{53}} = \alpha_{13}\left[\alpha_{35} + \frac{A_1^5}{A_1^3}\right]$$

$$\frac{\partial\alpha_{13}}{\partial a_{21}} = \alpha_{13}\alpha_{12} \qquad \frac{\partial\alpha_{13}}{\partial a_{51}} = \alpha_{13}\alpha_{15} \qquad \frac{\partial\alpha_{13}}{\partial a_{54}} = \alpha_{13}\left[\alpha_{45} + a_{43}\frac{A_1^5}{A_1^3}\right]$$

$$\frac{\partial\alpha_{13}}{\partial a_{35}} = \alpha_{13}\alpha_{53} \qquad \frac{\partial\alpha_{13}}{\partial a_{52}} = \alpha_{13}\alpha_{25} \qquad \frac{\partial\alpha_{13}}{\partial a_{55}} = \alpha_{13}\alpha_{55}$$

$$\frac{\partial\alpha_{23}}{\partial a_{15}} = \alpha_{23}\left[\alpha_{51} + a_{21}\frac{A_5^3}{A_2^3}\right] \qquad \frac{\partial\alpha_{23}}{\partial a_{43}} = \alpha_{23}\left[\alpha_{34} + \frac{A_2^4}{A_2^3}\right] \qquad \frac{\partial\alpha_{23}}{\partial a_{53}} = \alpha_{23}\left[\alpha_{35} + \frac{A_2^5}{A_2^3}\right]$$

$$\frac{\partial\alpha_{23}}{\partial a_{21}} = \alpha_{23}\left[\alpha_{12} + \frac{A_1^3}{A_2^3}\right] \qquad \frac{\partial\alpha_{23}}{\partial a_{51}} = \alpha_{23}\alpha_{15} \qquad \frac{\partial\alpha_{23}}{\partial a_{54}} = \alpha_{23}\left[\alpha_{45} + a_{43}\frac{A_2^5}{A_2^3}\right]$$

$$\frac{\partial\alpha_{23}}{\partial a_{35}} = \alpha_{23}\alpha_{53} \qquad \frac{\partial\alpha_{23}}{\partial a_{52}} = \alpha_{23}\alpha_{25} \qquad \frac{\partial\alpha_{23}}{\partial a_{55}} = \alpha_{23}\alpha_{55}$$

$$\frac{\partial\alpha_{33}}{\partial a_{15}} = \alpha_{33}\left[\alpha_{51} - \frac{A_5^1}{A_3^3}\right] \qquad \frac{\partial\alpha_{33}}{\partial a_{43}} = \alpha_{33}\alpha_{34} \qquad \frac{\partial\alpha_{33}}{\partial a_{53}} = \alpha_{33}\alpha_{35}$$

$$\frac{\partial\alpha_{33}}{\partial a_{21}} = \alpha_{33}\left[\alpha_{12} - \frac{A_1^2}{A_3^3}\right] \qquad \frac{\partial\alpha_{33}}{\partial a_{51}} = \alpha_{33}\left[\alpha_{15} - \frac{A_1^5}{A_3^3}\right] \qquad \frac{\partial\alpha_{33}}{\partial a_{54}} = \alpha_{33}\alpha_{45}$$

$$\frac{\partial\alpha_{33}}{\partial a_{35}} = \alpha_{33}\alpha_{53} \qquad \frac{\partial\alpha_{33}}{\partial a_{52}} = \alpha_{33}\left[\alpha_{25} - \frac{A_2^5}{A_3^3}\right] \qquad \frac{\partial\alpha_{33}}{\partial a_{55}} = \alpha_{33}\left[\alpha_{55} - \frac{A_5^5}{A_3^3}\right]$$

$$\frac{\partial\alpha_{43}}{\partial a_{15}} = \alpha_{43}\left[\alpha_{51} - a_{43}\frac{A_5^1}{A_4^3}\right] \qquad \frac{\partial\alpha_{43}}{\partial a_{43}} = \alpha_{43}\left[\alpha_{34} + \frac{A_3^3}{A_4^3}\right] \qquad \frac{\partial\alpha_{43}}{\partial a_{53}} = \alpha_{43}\alpha_{35}$$

$$\frac{\partial\alpha_{43}}{\partial a_{21}} = \alpha_{43}\left[\alpha_{12} - a_{43}\frac{A_1^2}{A_4^3}\right] \qquad \frac{\partial\alpha_{43}}{\partial a_{51}} = \alpha_{43}\left[\alpha_{15} - a_{43}\frac{A_1^5}{A_4^3}\right] \qquad \frac{\partial\alpha_{43}}{\partial a_{54}} = \alpha_{43}\alpha_{45}$$

$$\frac{\partial\alpha_{43}}{\partial a_{35}} = \alpha_{43}\alpha_{53} \qquad \frac{\partial\alpha_{43}}{\partial a_{52}} = \alpha_{43}\left[\alpha_{25} - a_{43}\frac{A_2^5}{A_4^3}\right] \qquad \frac{\partial\alpha_{43}}{\partial a_{55}} = \alpha_{43}\left[\alpha_{55} - a_{43}\frac{A_5^5}{A_4^3}\right]$$

$$\frac{\partial\alpha_{53}}{\partial a_{15}} = \alpha_{53}\alpha_{51} \qquad \frac{\partial\alpha_{53}}{\partial a_{43}} = \alpha_{53}\left[\alpha_{34} + \frac{A_5^4}{A_5^3}\right] \qquad \frac{\partial\alpha_{53}}{\partial a_{53}} = \alpha_{53}\left[\alpha_{35} + \frac{A_5^5}{A_5^3}\right]$$

$$\frac{\partial\alpha_{53}}{\partial a_{21}} = \alpha_{53}\alpha_{12} \qquad \frac{\partial\alpha_{53}}{\partial a_{51}} = \alpha_{53}\alpha_{15} \qquad \frac{\partial\alpha_{53}}{\partial a_{54}} = \alpha_{53}\left[\alpha_{45} + a_{43}\frac{A_5^5}{A_5^3}\right]$$

$$\frac{\partial\alpha_{53}}{\partial a_{35}} = \alpha_{53}^2 \qquad \frac{\partial\alpha_{53}}{\partial a_{52}} = \alpha_{53}\alpha_{25} \qquad \frac{\partial\alpha_{53}}{\partial a_{55}} = \alpha_{53}\alpha_{55}$$

$$\frac{\partial\alpha_{15}}{\partial a_{15}} = \alpha_{15}\left[\alpha_{51} + \frac{A_5^5}{A_1^5}\right] \qquad \frac{\partial\alpha_{15}}{\partial a_{43}} = \alpha_{15}\alpha_{34} \qquad \frac{\partial\alpha_{15}}{\partial a_{53}} = \alpha_{15}\alpha_{35}$$

$$\frac{\partial\alpha_{15}}{\partial a_{21}} = \alpha_{15}\alpha_{12} \qquad \frac{\partial\alpha_{15}}{\partial a_{51}} = \alpha_{15}^2 \qquad \frac{\partial\alpha_{15}}{\partial a_{54}} = \alpha_{15}\alpha_{45}$$

$$\frac{\partial\alpha_{15}}{\partial a_{35}} = \alpha_{15}\alpha_{53} \qquad \frac{\partial\alpha_{15}}{\partial a_{52}} = \alpha_{15}\alpha_{25} \qquad \frac{\partial\alpha_{15}}{\partial a_{55}} = \alpha_{15}\alpha_{55}$$

$$\frac{\partial\alpha_{25}}{\partial a_{15}} = \alpha_{25}\left[\alpha_{51} + a_{21}\frac{A_5^5}{A_2^5}\right] \qquad \frac{\partial\alpha_{25}}{\partial a_{43}} = \alpha_{25}\alpha_{34} \qquad \frac{\partial\alpha_{25}}{\partial a_{53}} = \alpha_{25}\alpha_{35}$$

$$\frac{\partial\alpha_{25}}{\partial a_{21}} = \alpha_{25}\left[\alpha_{12} + \frac{A_1^5}{A_2^5}\right] \qquad \frac{\partial\alpha_{25}}{\partial a_{51}} = \alpha_{25}\alpha_{15} \qquad \frac{\partial\alpha_{25}}{\partial a_{54}} = \alpha_{25}\alpha_{45}$$

$$\frac{\partial\alpha_{25}}{\partial a_{35}} = \alpha_{25}\alpha_{53} \qquad \frac{\partial\alpha_{25}}{\partial a_{52}} = \alpha_{25}^2 \qquad \frac{\partial\alpha_{25}}{\partial a_{55}} = \alpha_{25}\alpha_{55}$$

$$\frac{\partial\alpha_{35}}{\partial a_{15}} = \alpha_{35}\alpha_{51} \qquad \frac{\partial\alpha_{35}}{\partial a_{43}} = \alpha_{35}\alpha_{34} \qquad \frac{\partial\alpha_{35}}{\partial a_{53}} = \alpha_{35}^2$$

$$\frac{\partial\alpha_{35}}{\partial a_{21}} = \alpha_{35}\alpha_{12} \qquad \frac{\partial\alpha_{35}}{\partial a_{51}} = \alpha_{35}\alpha_{15} \qquad \frac{\partial\alpha_{35}}{\partial a_{54}} = \alpha_{35}\alpha_{45}$$

$$\frac{\partial\alpha_{35}}{\partial a_{35}} = \alpha_{35}\left[\alpha_{53} + \frac{A_5^5}{A_3^5}\right] \qquad \frac{\partial\alpha_{35}}{\partial a_{52}} = \alpha_{35}\alpha_{25} \qquad \frac{\partial\alpha_{35}}{\partial a_{55}} = \alpha_{35}\alpha_{55}$$

$$\frac{\partial\alpha_{45}}{\partial a_{15}} = \alpha_{45}\alpha_{51} \qquad \frac{\partial\alpha_{45}}{\partial a_{43}} = \alpha_{45}\left[\alpha_{34} + \frac{A_3^5}{A_4^5}\right] \qquad \frac{\partial\alpha_{45}}{\partial a_{53}} = \alpha_{45}\alpha_{35}$$

$$\frac{\partial\alpha_{45}}{\partial a_{21}} = \alpha_{45}\alpha_{12} \qquad \frac{\partial\alpha_{45}}{\partial a_{51}} = \alpha_{45}\alpha_{15} \qquad \frac{\partial\alpha_{45}}{\partial a_{54}} = \alpha_{45}^2$$

$$\frac{\partial\alpha_{45}}{\partial a_{35}} = \alpha_{45}\left[\alpha_{53} + a_{43}\frac{A_5^5}{A_4^5}\right] \qquad \frac{\partial\alpha_{45}}{\partial a_{52}} = \alpha_{45}\alpha_{25} \qquad \frac{\partial\alpha_{45}}{\partial a_{55}} = \alpha_{45}\alpha_{55}$$

$$\frac{\partial\alpha_{55}}{\partial a_{15}} = \alpha_{55}\alpha_{51} \qquad \frac{\partial\alpha_{55}}{\partial a_{43}} = \alpha_{55}\alpha_{34} \qquad \frac{\partial\alpha_{55}}{\partial a_{53}} = \alpha_{55}\alpha_{35}$$

$$\frac{\partial\alpha_{55}}{\partial a_{21}} = \alpha_{55}\alpha_{12} \qquad \frac{\partial\alpha_{55}}{\partial a_{51}} = \alpha_{55}\alpha_{15} \qquad \frac{\partial\alpha_{55}}{\partial a_{54}} = \alpha_{55}\alpha_{45}$$

$$\frac{\partial\alpha_{55}}{\partial a_{35}} = \alpha_{55}\alpha_{53} \qquad \frac{\partial\alpha_{55}}{\partial a_{52}} = \alpha_{55}\alpha_{25} \qquad \frac{\partial\alpha_{55}}{\partial a_{55}} = \alpha_{55}^2$$

The Variances of $\boxed{\alpha_{LK}}$:

$$\mathrm{var}(\alpha_{LK}) = \left[\sum^{i,j} \frac{\partial\alpha_{LK}}{\partial a_{ij}}(\hat{a}_{ij} - a_{ij})\right]^2 \quad \text{where } (\hat{a}_{ij} - a_{ij}) = \sigma_{ij}$$

$$\mathrm{var}(\alpha_{11}) = \alpha_{11}^2\left[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \alpha_{53}\sigma_{35} - \frac{A_5^3}{A_1^1}\sigma_{35} + \alpha_{34}\sigma_{43} - \frac{A_3^4}{A_1^1}\sigma_{43} + \alpha_{15}\sigma_{51}\right.$$

$$+ \alpha_{25}\sigma_{52} + \alpha_{35}\sigma_{53} - \frac{A_3^5}{A_1^1}\sigma_{53} + \alpha_{45}\sigma_{54} - \frac{A_4^5}{A_1^1}\sigma_{54} + \alpha_{55}\sigma_{55} - \frac{A_5^5}{A_1^1}\sigma_{55}\Bigg]^2$$

$$\mathrm{var}(\alpha_{21}) = \alpha_{21}^2\Bigg[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \frac{A_1^1}{A_2^1}\sigma_{21} + \alpha_{53}\sigma_{35} - \frac{a_{21}A_5^3}{A_2^1}\sigma_{35} + \alpha_{34}\sigma_{43}$$

$$- \frac{a_{21}A_3^4}{A_2^1}\sigma_{43} + \alpha_{15}\sigma_{51} + \alpha_{25}\sigma_{52} + \alpha_{35}\sigma_{53} - \frac{a_{21}A_3^5}{A_2^1}\sigma_{53} + \alpha_{55}\sigma_{55}$$

$$- \frac{a_{21}A_5^5}{A_2^1}\sigma_{55} + \alpha_{45}\sigma_{54} - \frac{a_{21}A_4^5}{A_2^1}\sigma_{54}\Bigg]^2$$

$$\mathrm{var}(\alpha_{31}) = \alpha_{31}^2\Bigg[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \frac{A_3^2}{A_3^1}\sigma_{21} + \alpha_{53}\sigma_{35} + \frac{A_5^1}{A_3^1}\sigma_{35} + \alpha_{34}\sigma_{43} + \alpha_{15}\sigma_{51}$$

$$+ \frac{A_3^5}{A_3^1}\sigma_{51} + \alpha_{25}\sigma_{52} + \frac{a_{21}A_3^5}{A_3^1}\sigma_{52} + \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\Bigg]^2$$

$$\mathrm{var}(\alpha_{41}) = \alpha_{41}^2\Bigg[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \frac{A_4^2}{A_4^1}\sigma_{21} + \alpha_{53}\sigma_{35} + \frac{a_{43}A_5^1}{A_4^1}\sigma_{35} + \alpha_{34}\sigma_{43}$$

$$+ \frac{A_3^1}{A_4^1}\sigma_{43} + \alpha_{15}\sigma_{51} + \frac{A_4^5}{A_4^1}\sigma_{51} + \alpha_{25}\sigma_{52} + \frac{a_{21}A_4^5}{A_4^1}\sigma_{52} + \alpha_{35}\sigma_{53}$$

$$+ \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\Bigg]^2$$

$$\mathrm{var}(\alpha_{51}) = \alpha_{51}^2\left[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \frac{A_5^2}{A_5^1}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43} + \alpha_{15}\sigma_{51} + \frac{A_5^5}{A_5^1}\sigma_{51}\right.$$

$$\left. + \alpha_{25}\sigma_{52} + \frac{a_{21}A_5^5}{A_5^1}\sigma_{52} + \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\right]^2$$

$$\mathrm{var}(\alpha_{13}) = \alpha_{13}^2\left[\alpha_{51}\sigma_{15} + \frac{a_{43}A_5^4}{A_1^3}\sigma_{15} + \alpha_{12}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43} + \frac{A_1^4}{A_1^3}\sigma_{43} + \alpha_{15}\sigma_{51}\right.$$

$$\left. + \alpha_{25}\sigma_{52} + \alpha_{35}\sigma_{53} + \frac{A_5^5}{A_1^3}\sigma_{53} + \alpha_{45}\sigma_{54} + \frac{a_{43}A_1^5}{A_1^3}\sigma_{54} + \alpha_{55}\sigma_{55}\right]^2$$

$$\mathrm{var}(\alpha_{23}) = \alpha_{23}^2\left[\alpha_{51}\sigma_{15} + \frac{a_{21}A_5^3}{A_2^3}\sigma_{15} + \alpha_{12}\sigma_{21} + \frac{A_1^3}{A_2^3}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43} + \frac{A_2^4}{A_2^3}\sigma_{43}\right.$$

$$+ \alpha_{15}\sigma_{51} + \alpha_{35}\sigma_{53} + \frac{A_2^5}{A_2^3}\sigma_{53} + \alpha_{45}\sigma_{54} + \frac{a_{43}A_2^5}{A_2^3}\sigma_{54} + \alpha_{25}\sigma_{52}$$

$$\left. + \alpha_{55}\sigma_{55}\right]^2$$

$$\mathrm{var}(\alpha_{33}) = \alpha_{33}^2\left[\alpha_{51}\sigma_{15} - \frac{A_5^1}{A_3^3}\sigma_{15} + \alpha_{12}\sigma_{21} - \frac{A_1^2}{A_3^3}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43} + \alpha_{15}\sigma_{51}\right.$$

$$- \frac{A_1^5}{A_3^3}\sigma_{51} + \alpha_{25}\sigma_{52} - \frac{A_2^5}{A_3^3}\sigma_{52} + \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55} - \frac{A_5^5}{A_3^3}\sigma_{55}\Bigg]^2$$

$$\mathrm{var}(\alpha_{43}) = \alpha_{43}^2\Bigg[\alpha_{51}\sigma_{15} - \frac{a_{43}A_5^1}{A_4^3}\sigma_{15} + \alpha_{12}\sigma_{21} - \frac{a_{43}A_1^2}{A_4^3}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43}$$

$$+ \frac{A_3^3}{A_4^3}\sigma_{43} + \alpha_{15}\sigma_{51} - \frac{a_{43}A_1^5}{A_4^3}\sigma_{51} + \alpha_{25}\sigma_{52} - \frac{a_{43}A_2^5}{A_4^3}\sigma_{52}$$

$$+ \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55} - \frac{a_{43}A_5^5}{A_4^3}\sigma_{55}\Bigg]^2$$

$$\mathrm{var}(\alpha_{53}) = \alpha_{53}^2\Bigg[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43} + \frac{A_5^4}{A_5^3}\sigma_{43} + \alpha_{15}\sigma_{51} + \alpha_{25}\sigma_{52}$$

$$+ \alpha_{35}\sigma_{53} + \frac{A_5^5}{A_5^3}\sigma_{53} + \alpha_{45}\sigma_{54} + \frac{a_{43}A_5^5}{A_5^3}\sigma_{54} + \alpha_{55}\sigma_{55}\Bigg]^2$$

$$\mathrm{var}(\alpha_{15}) = \alpha_{15}^2\Bigg[\alpha_{51}\sigma_{15} + \frac{A_5^5}{A_1^5}\sigma_{15} + \alpha_{12}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43} + \alpha_{15}\sigma_{51} + \alpha_{25}\sigma_{52}$$

$$+ \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\Bigg]^2$$

$$\mathrm{var}(\alpha_{25}) = \alpha_{25}^{2}\left[\alpha_{51}\sigma_{15} + \frac{a_{21}A_5^5}{A_2^5}\,\sigma_{15} + \alpha_{12}\sigma_{21} + \frac{A_1^5}{A_2^5}\,\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43}\right.$$

$$\left. + \alpha_{15}\sigma_{51} + \alpha_{25}\sigma_{52} + \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\right]^{2}$$

$$\mathrm{var}(\alpha_{35}) = \alpha_{35}^{2}\left[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \alpha_{53}\sigma_{35} + \frac{A_5^5}{A_3^5}\,\sigma_{35} + \alpha_{34}\sigma_{43} + \alpha_{15}\sigma_{51} + \alpha_{25}\sigma_{52}\right.$$

$$\left. + \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\right]^{2}$$

$$\mathrm{var}(\alpha_{45}) = \alpha_{45}^{2}\left[\alpha_{51}\,\sigma_{15} + \alpha_{12}\sigma_{21} + \alpha_{53}\sigma_{35} + \frac{a_{43}A_5^5}{A_4^5}\,\sigma_{35} + \alpha_{34}\sigma_{43} + \frac{A_3^5}{A_4^5}\,\sigma_{43}\right.$$

$$\left. + \alpha_{15}\sigma_{51} + \alpha_{25}\sigma_{52} + \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\right]^{2}$$

$$\mathrm{var}(\alpha_{55}) = \alpha_{55}^{2}\left[\alpha_{51}\sigma_{15} + \alpha_{12}\sigma_{21} + \alpha_{53}\sigma_{35} + \alpha_{34}\sigma_{43} + \alpha_{15}\sigma_{51} + \alpha_{25}\sigma_{52}\right.$$

$$\left. + \alpha_{35}\sigma_{53} + \alpha_{45}\sigma_{54} + \alpha_{55}\sigma_{55}\right]^{2}$$

Coefficients of Variations of $\boxed{\alpha_{LK}}$:

Assuming that $\frac{\alpha_{ij}}{a_{ij}} = \frac{\sigma_{RS}}{a_{RS}} = C$, and that all C's are equal, yields

$$\frac{\sqrt{var(\alpha_{11})}}{\alpha_{11}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{51} + \alpha_{35}a_{53}\right.$$

$$\left. + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 2 - \frac{2 - A_5^5 a_{55} + a_{54}A_4^5}{A_1^1}\right]$$

$$\frac{\sqrt{var(\alpha_{21})}}{\alpha_{21}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53}\right.$$

$$\left. + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 3 - \frac{a_{21}}{A_2^1}(2 - a_{55}A_5^5 + a_{54}A_4^5)\right]$$

$$\frac{\sqrt{var(\alpha_{31})}}{\alpha_{31}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53}\right.$$

$$\left. + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 3 - \frac{a_{51}A_3^5}{A_3^1}\right]$$

$$\frac{\sqrt{var(\alpha_{41})}}{\alpha_{41}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53}\right.$$

$$\left. + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 4 - \frac{a_{51}A_4^5}{A_4^1}\right]$$

$$\frac{\sqrt{\mathrm{var}(\alpha_{51})}}{\alpha_{51}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 2 - \frac{a_{51}}{A_5^1}\right]$$

$$\frac{\sqrt{\mathrm{var}(\alpha_{13})}}{\alpha_{13}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 3 - \frac{a_{53}A_1^5}{A_1^3}\right]$$

$$\frac{\sqrt{\mathrm{var}(\alpha_{23})}}{\alpha_{23}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 3 + \frac{a_{43}A_2^4}{A_2^3}\right]$$

$$\frac{\sqrt{\mathrm{var}(\alpha_{33})}}{\alpha_{33}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 3 - \frac{3}{A_3^3} + \frac{2a_{55}A_5^5 + a_{51}A_1^5}{A_3^3}\right]$$

$$\frac{\sqrt{var(\alpha_{43})}}{\alpha_{43}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 3 - \frac{a_{43}}{A_4^3}(2 - a_{55}A_5^5 + a_{52}A_2^5)\right]$$

$$\frac{\sqrt{var(\alpha_{53})}}{\alpha_{53}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 2 - \frac{a_{53}}{A_5^3}\right]$$

$$\frac{\sqrt{var(\alpha_{15})}}{\alpha_{15}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + \frac{A_1^5}{A_1^5}\right]$$

$$\frac{\sqrt{var(\alpha_{25})}}{\alpha_{25}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 2\right]$$

$$\frac{\sqrt{var(\alpha_{35})}}{\alpha_{35}} = C\left[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53} + \alpha_{45}a_{54} + \alpha_{55}a_{55} + 1\right]$$

$$\frac{\sqrt{\mathrm{var}(\alpha_{45})}}{\alpha_{45}} = C\Big[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53}$$

$$+ \alpha_{45}a_{54} + \alpha_{55}a_{55} + 2\Big]$$

$$\frac{\sqrt{\mathrm{var}(\alpha_{55})}}{\alpha_{55}} = C\Big[\alpha_{51}a_{15} + \alpha_{12}a_{21} + \alpha_{53}a_{35} + \alpha_{34}a_{43} + \alpha_{15}a_{51} + \alpha_{25}a_{52} + \alpha_{35}a_{53}$$

$$+ \alpha_{45}a_{54} + \alpha_{55}a_{55}\Big]$$

Appendix to Chapter 7

Changes in Income Accompanying Changes in Wages and Salaries

One of the major problems that regional analysts and economists are faced with is the difficulties involved in obtaining accurate and detailed data on industry basis that are necessary for their empirical studies.

In the research work being conducted presently, five counties in the state of Pennsylvania have been selected for empirical study aimed to reveal the implications and the workability of the techniques developed for evaluating and planning the Corps of Engineers projects. One of the objectives involved in this research study has been to construct an improved method for measuring secondary benefits induced by Corps projects where the national income effects and local income effects estimated are consistent with one another.[1] In order to show how the method can be applied empirically, data are required on industry basis on household total income originating in each of the selected counties. Such data however, are available neither by industry nor on aggregate basis. The only data that have been found available for Pennsylvania counties (for the purpose of the research) are the totals and components (by source) of the personal income.

One way to compensate for lack of the necessary data is to investigate whether wages and salaries (which is the largest component in the total income) can be used to estimate the total income originating in the counties under study and, for that matter, in any counties or regions. The method developed for that purpose involves comparison

[1]The reasons why and how local income effects of projects are measured, are well explained in: G. S. Tolley. "National Income and Local Income Effects of Water Projects," June, 1967. (Unpublished.)

between incomes and wages and salaries. Specifically, it involves application of regression analysis to observe the sensitivity of changes as well as the extent of correlations between different incomes and wages and salaries.

The data collected were:

1. Personal income and wages and salaries in each of the 67 Pennsylvania counties for 1963, 1960 and 1962,

2. Personal income and wages and salaries for each of the 48 states in the U.S. for 1963, 1960, and 1962,

3. National income and wages and salaries for the U.S. (48 states) for the period of 37 years from 1929 through 1965, and

4. Another set of data were calculated from the personal incomes of Pennsylvania counties and of 48 states in the U.S. for the years 1963, 1960, and 1962. This set for which the term "income payments to individuals" is used, includes wages and salaries plus other labor income plus proprietors' income plus property income.

Assuming linear functional relationships between the variables, i.e. income variables as functions of wages and salaries, the following regressions were made:

1. The 1963 personal income totals of 48 states in the U.S. were regressed on the corresponding wages and salaries for that year. Another similar regression was made using the data of 1960 and 1962 combined.

2. The 1963, and then separately the combination of 1960 and 1962, personal income totals of the 67 Pennsylvania counties were regressed on their wages and salaries.

3. Income payments to individuals (i.e. personal income minus the difference between transfer payments and personal contribution for social insurance) for the 48 states were regressed on the wages and salaires once using the data of 1963 (i.e. 48 observations), and then using the data of 1960 and 1962 combined (i.e. 96 observations).

The same thing was done concerning the 67 Pennsylvania counties, taking the data of 1963 first (i.e. 67 observations), and then the figures of 1960 and 1962 (i.e. 134 observations).

4. The national income fo the U.S. (48 states) for the period from 1929 to 1965 (i.e. 37 observations) was regressed on the wages and salaries for the nation for the same period.

The same procedures indicated in the items from (1) to (4) above were repeated, but this time using the assumption of non-linear functional relationships between the income variables and wages and salaries.

The results of the regressions are shown in the attached Table 1. On the left-hand side, the columns show the number of observations used for estimating the relations in each category, the constants, the regression coefficients, and the correlation coefficients for each relationship. The same information is given on the right-hand side of the table in terms of percentages, i.e. in logarithmic form. The number 1.466 which is in the top of the slope column, for example, indicates that one unit or $1 change in wages and salaries in the 48 states in 1963 was on the average associated with 1.466 point or dollars change in personal income. The figure .9983 in the top in the correlation coefficient, R^2, column, for example, indicates that the relationship between changes in wages and salaries and in personal income in

1963 was almost perfect. The same interpretation is given by the corresponding figures on the right side of the table in percentage terms where the slope shows the elasticity, i.e. $\frac{dy}{y}/\frac{dw}{w}$ where y is income and w denotes wages and salaries.

It is interesting to observe that the differences between the regression coefficients (slope) in each category is only about (.02) and the difference between the highest and the lowest coefficients is about (.19). Furthermore, the R^2's in all categories are over .99 (the lowest being .9977) indicating very high correlations between the incomes and wages and salaries.

From the results of these observations one can conclude that wages and salaries can well be used to estimate total income originating in the area and study, and in other regions.

The method developed here can be used to investigate the extent of income-wages and salaries relations in any region or area, without being limited for Pennsylvania counties.

able 13. Regressions of personal income and income payments on wages and salaries

Dependent variables	Number of observations	Linear			Logarithm		
		Constant (millions of dollars)	Slope	R^2	Antilog of constant (millions of dollars)	Slope	R^2
Personal Income							
1. United States							
a. 1963	48	135.300	1.466	.9983	2.072	.9632	.9944
		(89.51)	(.008845)		(1.091)	(.01064)	
b. 1960, 1962	96	192.540	1.4570	.9992	2.148	.9593	.9966
		(54.093)	(.005848)		(1.069)	(.008231)	
2. Pennsylvania counties							
a. 1963	67	5.576	1.449	.9997	.0032	.9388	.9960
		(2.183)	(.002883)		(.0011)	(.007349)	
b. 1960, 1962	134	7.9182	1.4270	.9999	.0032	.9390	.9983
		(1.4342)	(.001969)		(.0011)	(.004756)	
Income Payments to Individuals*							
1. United States							
a. 1963	48	115.000	1.403	.9977	1.892	.9679	.9949
		(101.300)	(.01001)		(1.088)	(.01025)	
b. 1960, 1962	96	162.750	1.3835	.9993	1.998	.9595	.9967
		(49.533)	(.005355)		(1.068)	(.008074)	
2. Pennsylvania counties							
a. 1963	67	2.847	1.365	.9998	.0027	.9454	.9967
		(1.764)	(.002330)		(.0011)	(.006787)	
b. 1960, 1962	134	5.1416	1.3442	.9999	.0027	.9455	.9986
		(1.1265)	(.001546)		(.0011)	(.004436)	
National Income							
United States**							
1929-1965	34	3923.2	1.5391	.9994	1.498	1.0040	.9986
		(1591.4)	(.009028)		(1.109)	(.009007)	

*Income payments to individuals include (1) wages and salaries, (2) other labor income, (3) proprietors' income and (4) property income.

**District of Columbia, Alaska, and Hawaii are excluded.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
North Carolina State University and the University of Chicago	UNCLASSIFIED
	2b GROUP: N/A

3 REPORT TITLE

ESTIMATION OF FIRST ROUND AND SELECTED SUBSEQUENT INCOME EFFECTS OF WATER RESOURCES INVESTMENT

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final Report

5 AUTHOR(S) (First name, middle initial, last name)

George S. Tolley and: V.S. Hastings, Adnan Z. Daghestani, Robert Rugg, Shaul Ben-David, G.K. Kripalani

6 REPORT DATE	7a. TOTAL NO OF PAGES	7b NO OF REFS
February 1970		

8a CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
DA-49-129-CIVENG-65-11	IWR Report 70-1
b PROJECT NO: N/A	
c	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report): None
d	

10 DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY
None	U.S. Army Engineer Institute for Water Resources, 206 N. Washington Street, Alexandria, Virginia 22314

13 ABSTRACT

This publication presents research conducted in collaboration with Leven et al, Development Benefits of Water Resources Investments, IWR Report 69-1, November 1969. Ben-David, Hastings and Tolley assert a methodology for estimating those first-round local effects related to the change in the availability of water for supply purposes or recreation. Rugg Develops estimates of demand for reservoir recreation. Kripalani estimates the change in structural unemployment resulting from changes in output. Daghestani and Tolley evaluate the differencees in regional multiplier values and propose a refinement in regional multiplier estimates. A summary paper by Tolley integrates the various research papers and presents a case for the selected strategy for estimating those national income benefits to water resource development in excess of direct output valued in terms of the users "willingness to pay."

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Water Resources Development						
Water Resources Planning						
Evaluation Process						
Economics						
Benefits						
Secondary Benefits						
Regional Analysis						
Investment						
Benefit-Cost Analysis						
Regional Development						
Regional Growth						
Interest Rates						
Allocation Process						
Recreation Demand						
First-Round (Local) Effects						
Regional Multipliers						