AN ABSTRACT OF THE THESIS OF

TSAE-YUN LIN for the Master of Science
(Name of Student) (Degree)

in Agricultural and Resource Economics presented on August 26, 1980

Title: AN APPLICATION OF TRANSFER FUNCTION AND ECONOMETRIC PROCEDURES TO EMPLOYMENT IMPACT ANALYSIS: GRANT COUNTY, OREGON

Abstract approved: Redacted for Privacy
Ronald A. Oliveira

This study centered upon two objectives. First, to demonstrate and apply a statistical method, the transfer function, and to test its usefulness in regional economic modeling and regional impact analysis; and second, to obtain a better understanding of the causal relationships among total employment, government employment, lumber and wood products manufacturing employment, and agriculture employment in Grant County, Oregon. The first objective is because of the deficiencies in using economic base, input-output, and regional econometric models in regional impact analysis. The second objective is because of the need for analyzing the direct and induced dynamic impacts of basic industry employment changes in Grant County.

Transfer function models are based upon an interative approach to the identification of a statistical model that specifies a dynamic relationship between two or more interrelated time series. In this study, one-input and two-input transfer function models, that specify dynamic relationships among total employment, government employment and lumber and wood products manufacturing employment, are developed.

In order to comparatively evaluate the results from the transfer function models, univariate ARIMA and econometric models also are specified and presented for the same employment series.

Six types of goodness-of-fit measures are used to examine the forecasting performances of the ARIMA, transfer function, and econometric models. The empirical results suggest considerable confidence in the accuracy of the ARIMA, one-input transfer function, and econometric models.

Various hypothetical changes of basic employment in Grant County are simulated in order to estimate the employment multiplier effects for the basic industries. The static and dynamic employment multipliers of three basic industries from the econometric models are also presented and interpreted. The agriculture sector is found to have a higher employment multiplier than the government and lumber and wood products manufacturing sectors.

The transfer function models are found to be quite appealing if one is primarily interested in forecasting. However, if the researcher wants to analyze the nature of regional impacts or to explain the "complicated" behavior of an economic system, the input-output model and the regional econometric models appear to be more attractive tools to employ.

An Application of Transfer Function and Econometric Procedures to Employment Impact Analysis: Grant County, Oregon

by

Tsae-Yun Lin

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

June 1981

APPROVED:

Redacted for Privacy

Associate Professor of Agricultural & Resource Economics
in charge of major

Redacted for Privacy

Head of Agricultural & Resource Economics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented ______August 26, 1980______

Typed by ___Lynn M. Hopper___ for ___Tsae-Yun Lin___

ACKNOWLEDGEMENT

I wish to express my most sincere appreciation to several individuals without whose counsel and assistance this thesis would have been infinitely more difficult. In particular, I would like to acknowledge Dr. Ronald A. Oliveira for his advance, support, and service as major professor throughout my term of study in this department.

I also wish to thank the other members of my committee, Dr. William G. Brown, Dr. H. D. Brunk, Dr. Victor T. Neal, and Dr. Frederick W. Obermiller, for their helpful comments on a draft copy of my thesis.

Special thanks is due to my mother, Yue-Wei, and husband, Biing-Hwan for their encouragement and patience, without which graduate study would not have been possible.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF TABLES

AN APPLICATION OF TRANSFER FUNCTION AND ECONOMETRIC PROCEDURES TO EMPLOYMENT IMPACT ANALYSIS: GRANT COUNTY, OREGON

CHAPTER I

INTRODUCTION

Problem

Long term economic stability and prosperity are elusive goals for many people living in Oregon, and indeed, throughout the Western United States. For those now living in the immense rural areas of the region, employment and income are dependent on basic resource-using industries (primarily forestry and agriculture). In contrast to more urbanized areas, levels of income are lower, employment opportunities more scarce, social services few; and the potential of the conventional growth strategies of industrialization and economic diversification are minimal (Harris and Obermiller, 1978).

Grant County, an eastern Oregon county, shares many of the economic and social problems common to rural areas throughout the Western United States. This county is typical of many natural resource based rural communities in that most household income is derived from resource-based industries and government agencies managing natural resource lands. For instance, in 1977 wages paid by resource-based product groups accounted for 47 percent of the total county wage bill (forest products: 44 percent, agriculture: 3 percent). Wages paid by government accounted for 32 percent of the total county wage bill. Within the government sector, the Forest Service accounts for a very high proportion of total wage and salary payments (Obermiller and Miller, 1979, p. 10).

From these statistics, it can be concluded that private firms and public agencies related to forestry provide most of the income (or employment) for local residents in Grant County. Therefore, any decision affecting future growth in these two sectors will have a significant impact on local residents' income (or employment), and will also result in direct and induced economic effects on the county's economy as a whole. The need for research on the extent to which income (or employment) variations in the forest industry and related government agencies (particularly the Forest Service) generate variations in total county income (or employment) is apparent.

In regional economic impact studies, the total impact, or the direct and induced effects of change in economic activity often is measured using "multipliers", which express the total effect as a multiple of the direct effect. Multipliers are very useful tools in impact analysis. For instance, estimates of regional employment multipliers provide a rough but useful means of assessing the total employment impact of gains or losses in a region's export activity. However, one disadvantage of using a static multiplier is its failure to identify the time path of induced impacts. The dynamic multiplier is capable of identifying the impacts in each time period, and in the absence of a technical change, will accumulate to the value of the static multiplier. The additivity and separability over periods are the major advantages of using dynamic multipliers over static multipliers (Liew, 1977, p. 95). For all of the above reasons, analysis of dynamic multipliers in the context of local employment variations is warranted.

Concern for regional economic growth and development has sparked considerable interest in the ability to explain, and thus predict, the regional or local economic impacts of various types of activities. Local policy makers, in order to design better plans for regional development, need accurate predictions of levels of key economic aggregates such as employment, income, and output for different development strategies. Methods such as economic base, input-output, and regional econometric models frequently have been employed to explain the regional economic impacts of various types of activities and to forecast the levels of regional economic indicators. These three approaches have different theoretical implications and, therefore, different advantages and disadvantages. Since the deficiencies in these three approaches are significant, the development of other analytic techniques has been a major concern of regional economic analysts.

The transfer function was introduced to the statistical literature over a decade ago by Box and Jenkins in their article "Some Recent Advances in Forecasting and Control" (Box and Jenkins, 1966). In recent years, this technique has been proven to be useful in economic analysis as it serves to measure the impact from the hypothesized input variable upon the output variable at different time periods. Bennett acquainted regional scientists with the transfer function in his article, "Process Identification for Time Series Modeling in Urban and Regional Planning" (Bennett, 1974). He suggest that the transfer function can form the basis of an inferential methodology for model building which leads naturally to statistical forecasts and policy. The present study is an attempt to employ this technique in regional economic modeling and to test its ability in regional impact analysis. Monthly employment data for Grant County, Oregon will serve as a case study.

In prior research, three static input-output studies of the Grant County economy have been made (Bromley, et. al., 1964; Haroldsen and Youmans, 1972; Obermiller and Miller, 1977). Johnson (1980) employed a simulation method to develop a dynamic input-output model of the Grant County economy which projects aggregate output, income, investment, etc. In these previous input-output studies, the income multipliers of different economic sectors have been derived and evaluated, but little attention has been paid to the dynamic employment multipliers of basic sectors. Therefore, the present study attempts to fill this void by applying transfer function and econometric analysis[1] to analyze the dynamic employment multiplier effects of monthly government employment, lumber and wood products manufacturing employment, and agriculture employment in Grant County, Oregon.[2] Another reason for applying these procedures to employment data is because quarterly or monthly income data are not available at the county level.

---

[1]In this study, an econometric model is mainly used as a comparison to the transfer function model.

[2]These three sectors are considered to be basic industries in Grant County and are discussed in more detail in Chapter IV.

Objectives

The major objectives of this study are twofold: first, to demonstrate and apply a statistical method, the transfer function, and to test its usefulness in regional economic modeling and regional impact analysis; and second, to obtain a better understanding of the causal relationships among total employment, government employment, lumber and wood products manufacturing employment, and agriculture employment in Grant County. The intermediate objectives are to:

(1) analyze the induced dynamic impacts of three basic industries employment changes in Grant County, Oregon;

(2) attempt to determine the accuracy of the ARIMA[3], transfer function, and econometric models;

(3) determine the suitability of time series (transfer function) and econometric procedures for estimating dynamic employment multipliers;

(4) compare the forecasting accuracy of ARIMA, transfer function, and econometric models; and

(5) provide a reliable short-term forecasting tool for projecting local employment consequences of changes in basic industries employment.

----

[3]The Box-Jenkins autoregressive-integrated-moving-average (ARIMA) model is the basic component of the transfer function model.

PROCEDURES

The achievement of the above objectives will depend mainly on the use of the transfer function and more traditional econometric models which are applied to a variation of the economic-base model. Secondary time series data obtained from the Department of Human Resources of the State of Oregon are used.

The regional economic models have progressed from ad hoc methods to more formal models. In Chapter II a review of the characteristics and problems of three types of regional forecasting models is presented. These include economic base, input-output, and econometric models, which are commonly used to estimate regional employment multipliers. At the end of this chapter, the theoretical (economic) rational for using transfer function models to examine employment impacts is discussed.

An outline of the multivariate time series analysis appears in Chapter III. There, the basic differences among ARIMA, transfer function, and econometric models are discussed. The conceptual theory and the modeling procedures of the ARIMA and the transfer function are also presented. Data sources and their limitations, as well as the empirical implementation of the theoretical ARIMA, one-input transfer function, two-input transfer function, and econometric models are discussed in Chapter IV. The forecasting performances of these four models also are presented and compared in this chapter.

The transfer function and econometric models presented in the Chapter IV are employed in Chapter V to estimate the dynamic employment response (or employment multiplier effect) of total employment from changes in government, lumber and wood products manufacturing, and agriculture employment. Results of multipliers from input-output studies are compared with multipliers from time series and econometric models. Finally, conclusions and suggested extensions of this analysis are presented in Chapter VI.

CHAPTER II

OVERVIEW OF REGIONAL EMPLOYMENT MULTIPLIER ANALYSIS

Regional Employment Multiplier, Its Estimation Methods and Problems

Estimates of regional employment multipliers provide a rough but useful means of assessing the total employment impact of gains or losses in a region's export activity. In recent years, considerable attention has been focused on the application of multiplier analysis to various regional economic problems.

There have been three fairly general approaches to the study of regional multipliers: (1) economic base, (2) input-output, and (3) regional econometric. These three approaches have different theoretical implications, causing individual advantages and disadvantages. For instance, economic base models are rooted in a simple theory of urban growth: the expansion of a city or region is determined by the growth of its exports. These models are correspondingly simple in construction, and, as a result, they are beset with many technical and theoretical problems and produce inaccurate forecasts. The input-output approach can introduce a higher degree of sophistication and comprehensiveness than the economic base approach. However, the construction of a complete interindustry model usually entails high costs and substantial data problems. Regional econometric models are in some respects a compromise between economic base and input-output formulations. With respect to data costs, they do not require as much data collection as do input-output models. They are comparable in cost to economic base models, but provide relatively more information about the structure of a region's economy. Unfortunately, some econometric and statistical problems make regional econometric models difficult (if not impossible) to build for small regions.

There are controversies in the literature about the accuracy of multipliers derived from different models. However, Isard and Czamanski (1965) have demonstrated the empirical equivalence of economic base, aggregate input-output, and simple econometric models; Garnick (1970) and Davis (1975) have shown theoretically and empirically that the aggregate input-output multiplier is equal to the multiplier associated with the economic base model; Billings (1969) has proven the mathematical identity of the multipliers derived from the economic base model and the input-output model; Connaughton and McKillop (1979) provide a treatment of the equivalence between Keynesian, input-output, and economic base multipliers.

In the following sections, each of these individual methods of regional analysis and their problems will be examined. An outline of the theoretical (economic) rationale for using transfer function models to examine regional employment impacts appears in the final section.

Economic Base Model

One of the earliest statistical models to be employed in regional research was the economic base model. It was first formulated by Homer Hoyt in the 1930's and has been employed in numerous regional studies.[4]

---

[4]For a comprehensive bibliography, see Isard (1960, pp. 227-231); Bendavid (1974, pp. 179-191); and Glickman (1977, pp. 197-206).

Economic base theory postulates that economic activity in a region expands or contracts primarily because of export activity to other regions. The determinants of supply are not considered; therefore, economic activity is considered to be primarily demand oriented. The economic base model has also been described as essentially a Keynesian type model (Weiss and Gooding, 1968, p. 235). This theory in its simplest formulation provides for the partitioning of a regional economy into two segments (i) export (basic) industries, which are generally held to form the economic base of a region and are the reason for its growth, and (ii) nonexport (residentiary or non-basic) industries, which supply the local requirements of the basic industries and which may also supply their own. Total employment in a region may be broken down according to this basic/nonbasic dichotomy, and it is then possible to derive regional employment multipliers. As export activities expand, nonbasic sectors of the region are stimulated to generate additional economic activity, and total regional employment will grow by some multiple of the initial increase in export oriented jobs.

Regional employment multipliers of the economic base type may be calculated by several different methods. The most commonly used method is the ratio of total employment to basic employment for a given year. The second method used to calculate a multiplier is represented by the change in total employment associated with a unit change in basic employment. These two simple multiplier concepts, which were implicit in the earliest writings on the economic base model, have been widely criticized. An improvement is possible if data are available for more than two periods. It is then possible to employ a stochastic linear estimator, i.e.,

$$EN = \beta_o + \beta_1 EB + \mu, \ \beta_1 > 0$$

where: EN is nonbasic sector employment.
EB is basic sector employment.

Solving for total employment (ET) as a function of basic employment yields:

$$ET = \beta_o + (1 + \beta_1)EB + \mu$$

Where the derivative of ET with respect to EB, $1 + \beta_1$, is regarded as the total employment multiplier associated with a change in basic employment.

The regression analysis procedure, of course, is a more sophisticated formulation of the model than the ratio method. For the past three decades, many individuals have estimated regional employment multipliers by use of the least squares method. Some representative studies are Hildebrand and Mace (1950), Thompson (1959), Sasaki (1963), Weiss and Gooding (1968), Moody and Puffer (1970), Park (1970), and Lewis (1976).

Other measures of regional economic activity such as regional income wages, or personal income, can be substituted for the variables in the above three methods. Given that employment data often are readily available, this has been a popular variable for regional analysts. The advantages and disadvantages of using employment rather than income measures are discussed below.

Problems with the Economic Base Model

The economic base theory is used widely because it is inexpensive and easy to implement at any level of regional aggregation. In earlier years, both short-run and long-run changes were assumed to be explained by the theory. Currently, however, it is generally held that multipliers have predictive value only as long as the industrial structure of the local economy remains unchanged; i.e., the concept that is valid only in the short run. Additional criticism of the theory center around its usefulness. For example, Isard (1960) has discussed in some depth the technical and conceptual difficulties of using two sector economic base multipliers to forecast impacts. The chief technical problems are (1) choosing a unit of measure (e.g., employment of income); (2) distinguishing between basic and residentiary industries; and (3) defining a geographical unit of analysis.

Generally, employment is used as a proxy for general economic activity (e.g., income or production). However, it will carry at least three problems. First, wage differentials among sectors and part-time employment are not accounted for in employment multipliers. Therefore, an employee earning $15,000 annually (or a full-time employee) receives as much weight in the employment multiplier as does the employee earning $7,000 annually (or a part-time employee). This despite the fact that the former employee may provide a significantly greater economic stimulus to the community than the later. This problem becomes more serious when considering the dynamic variation (e.g., technological development might cause an increase in income and output but not in employment).

Second, transfer payments are not reflected in the employment data.

The result, therefore, is an underestimation of the size of the export base, and an overestimation of the multiplier. Gibson and Worden (1979) have indicated that two-thirds of the communities studied showed differences of ten percent or more between unadjusted and fully adjusted multipliers. They also found that the greatest declines in magnitude of the multiplier were for communities with large retirement populations. Similarly, Connaughton and McKillop (1979) have suggested that the importance of recognizing transfer payments in small area analysis cannot be overemphasized.

Third, incommuters, i.e., those employees who live outside the base area, are included in the employment data. Outcommuters, i.e., those individuals living in the base area but drawing earned income from outside the base area, are not reflected in the employment data. Both cases result in a bias of the employment multipliers.

Ideally, the researcher would like to have detailed interregional flow data on which to base an estimate of the volume of export activity. Generally, such data are unavailable and the cost of collecting it through a survey or census of area establishments is prohibitive. Thus, indirect estimation methods such as the location quotients and minimum requirements techniques have been widely used, although there are serious questions about their ability to accurately measure export activity. For instance, Greytak (1969) notes that both the location quotient and minimum requirements techniques yield export estimates which differ significantly from those actually observed. The minimum requirements technique was found to be more accurate than the location quotient, but neither technique was considered adequate. In view of this problem, some other indirect estimation methods have been developed.

And mo~~o~~re equation has been shown to be the only satisfactory alternative to the census survey by Gibson and Worden (1979).

Additional problems exist in economic base theory, for instance, the assumed invariance of the parameters of the system over time. Garnick (1969, 1970) has argued that it is incorrect to treat nonbasic activity as a fixed proportion of basic activity, since interregional industry mixes are becoming more alike. Parameters of the base models are changing, reflecting these interregional shifts as well as changes in income, tastes, and technology. In addition, Moody and Puffer (1969) showed that the estimated constant 89.4 in Sasaki's equation, ET = 89.4 + 1.279 EB, is significantly different than zero. Thus, they conclude that the basic/service ration is not stable. Moreover, the ratio also has the defect of ignoring the feedback effects of economic development.

Time lags also present a problem in economic base analysis. In most studies of regional employment multipliers, the adjustment of local (service) employment to changes in export employment has been hypothesized to occur in the same period, i.e., an unlagged relationship exists between the two. Park (1970) concludes that it is important to measure the lagged relationship in an economic-base model. He demonstrates that if the lagged model is significant, it can produce different results from the static model with its imperfect estimating techniques.

Disaggregation bias also presents problems in economic base analysis. Romanoff (1974) concluded that economic-base models contain a disaggregation bias because they do not measure the interdependent relationship between industries. In addition, Lewis (1976) has demonstrated that estimates from the economic-base model will vary according to alternative definitions of the regional unit, procedures in parameter specification, and the methodology in determination of the export base. Comments such as Romanoff's and Lewis' are constructive and serve as a warning for those who use the economic base approach.

Weiss and Gooding (1968), however, suggested that an economic-base approach can be applicable for small regions. A necessary condition is that the export-base industries be disaggregated into a few homogeneous sectors. They argue that employees in each of these sectors should have similar spending habits and impacts all other interdependent industries in a common manner. Therefore, parameter estimates of the economic-base model may not contain disaggregation bias per Romanoff's critique. Polzin (1980) also suggested that the economic base format was appropriate in the small rural economies of the West. He postulated that most of the past failures of the economic-base approach were due to problems associated with estimating techniques and the quality of the data, rather than a faulty theoretical foundation.

Despite the criticisms raised above, economic base analysis is still widely used. This is because it provides a fast, simple, and inexpensive means of evaluating the existing structure of a region and, by use of the multiplier, of estimating probable change in employment or income resulting from an impact. Moreover, the analyst in a small regional economy most often does not have the luxury of well-defined regional accounts that are needed for the more intricate input-output models. Thus, economic base analysis seems to offer an attractive alternative to input-output methods for analysts or planners in small regions.

Input-Output Model

Regional analysts desirous of a more complete view of the interrelationships obtained in a local economy have often turned to input-output models. It is possible by use of a regional input-output model to derive the interindustry multiplier effects of any change in final demand for all sectors of the economy. Moreover, the multiplier effects induced by increased household consumption expenditures can be estimated by incorporating consumption functions into the model. The major advantage of the input-output approach is that it delineates the interdependence of all defined sectors of the economy.

The input-output analytical method was developed from the theory of general equilibrium. Francois Quesnay's Tableau Economique of 1758 dealt with circular flows between industries and general equilibrium concepts. Walras stressed the interdependence between the production sectors of an economy with his general equilibrium model in the 1870's. The entire concept was not utilized until about 1931 when Wassily Leontief started an input-output system of the U.S. economy. The results were published in 1936. Leontief simplified Walras' generalized model so that equations associated with it could be estimated empirically. He used two simplifying assumptions.[5] These simplifying assumptions provide for important contrasts between input-output and many other economic models.

The Leontief input-output technique is well-known and is summarized

---

[5]First, the large number of commodities in the Walras model was aggregated into relatively few outputs, one for each industrial sector. Second, supply equations for labor and demand equations for final consumption were abandoned. The remaining production equations were expressed in their simplest, linear form (Palmer, et. al., 1978).

briefly here in order to provide a basic notation for the explanation of the income and employment multiplier. The most crucial, and time consuming, step in constructing the model is the collection of primary survey data for specifying the transactions table and which describe in dollar terms the flow of commodities among processing and final demand sectors. Once this step was completed, technical production coefficients were estimated and the Leontief solution was obtained. Equations (1) through (4) summarize the development of the model.

$$(1) \quad X_i = \sum_{j=1}^{n} x_{ij} + Y_i, \qquad (i = 1, 2, \ldots, n)$$

where:

X_i = total output of the ith industry;

x_{ij} = the value of output of industry i purchased by industry j; and

Y_i = the exogenous or "final" demand (consumption, investment, government and exports) for the output of industry i.

Since the data requirements for this type of study are stringent, simplifying assumptions have been made:

(a) Each commodity group is produced by a unique producing industry.

(b) There are no external economies or diseconomies possible.

(c) There is a unique observable production process which does not allow for the substitution of inputs.

Assumption (c) implies:

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

where: a_{ij} is the production coefficient specifying the amount of i needed to produce one unit of j, and

X_j is the output in industry j.

Substituting equation (2) into (1) yields:

$$(3)\quad X_i = \sum_{j=1}^{n} a_{ij}X_j + Y_i$$

Equation (3) can be rewritten in matrix notation for all sectors.

$$X = AX + Y$$

where X is a vector of (X_i), A is a matrix of (a_{ij}), and Y is a vector of (Y_i).

The general solution of the model may now be derived as:

$$(4)\quad X = (I-A)^{-1}Y$$

The $(I-A)^{-1}$ is defined as the direct plus indirect coefficients, or total requirements matrix.

This basic model may be used to generate consistent forecasts where, for each sector, projected industry output must always equal intermediate plus final demands. The procedure is simply to project individual components of final demands and thus calculate a new level of output (X) using the projected final demands (Y) and the Leontief inverse, $(I-A)^{-1}$.

The employment multiplier which is defined as the total change in employment due to a one-unit change in the employed labor force of a particular sector can be computed from the input-output model. The concept of input-output employment multipliers was developed by Moore and Peterson (1955). The basic assumption in computing employment multipliers is that there is a linear relationship between employment and output in a sector.

In general, there are two frequently used types of multipliers. First, there is the so-called Type I multiplier which is based upon the direct and indirect results of an exogenous change in final demand when the household is part of final demand. Second, there is the so-called Type II multiplier which is based upon the direct, indirect and induced results of an exogenous change in final demand when the household is part of the endogenous system of interdependency and final demand consists of government spending, investment expenditures and foreign purchases. Accordingly, both types of multiplier effects can be measured in terms of gross output, household income, and/or employment. For example, for the Type I income multiplier,[6]

$$(5) \quad M_j = \sum_{i=1}^{n} h_{oi} a^*_{ij} / h_{oj} \,, \quad (j = 1, 2, \ldots, n)$$

Here M_j is the Type I income multiplier, representing the ratio of the direct plus indirect income effects to the direct income effect: h_{oi}, h_{oj} are the household coefficients, i.e., the ratio of the household services purchased by the ith (or jth) industry to the total ith (or jth) industry inputs; and a^*_{ij} is the elements of $(I-A)^{-1}$.

For the Type I employment multiplier,

$$(6) \quad M^*_j = \sum_{i=1}^{n} l_{oi} a^*_{ij} / l_{oj} \,, \quad (j = 1, 2, \ldots, n)$$

Here M^*_j is the Type I employment multiplier; l_{oi} and l_{oj} are the direct employment coefficients per dollar of output for the ith (or jth) industry and a^*_{ij} is the elements of $(I-A)^{-1}$.

---

[6]For the theoretical basis of the discussion see Moore and Peterson (1955, pp. 374-378); and Bradley and Gauder (1969, p. 311 and p. 313). For an empirical application see Tweeten and Brinkman (1976, pp. 314-321).

The difference between the Type I and Type II multipliers is a_{ij}^{*} in equation (5), (6) above. For Type II multipliers, the coefficients a_{ij}^{*} are the direct plus indirect plus induced effects on the output of the ith industry for a dollar change in the final demand for the jth industry's output. For Type I multipliers, the a_{ij} represent direct plus indirect, but not induced, effects.

Problems with the Input-Output Model

There has been considerable effort in recent years directed to the application of the input-output model to regions and urban areas to trace the flow of goods among local industries and between regions, and in using it in forecasting and economic impact analysis. For example, Moore and Peterson (1955) employed an input-output model to analyze balance of payments in Utah. Bromley, Blanch and Stoevener (1968), Haroldsen and Youmans (1972) and Obermiller and Miller (1979) employed input-output models to analyze the impacts of natural and community resource changes in Grant County, Oregon, on the local economy.

The disadvantage of an input-output model is that it may require an extensive survey to compile the data for its calculation. As a result, investigators must often undertake the expensive and time-consuming task of primary data gathering. In light of these data problems, recent efforts to reduce the costs of constructing regional input-output models have been directed toward the development of: (1) non-survey techniques of model construction; or (2) truncated or aggregated input-output models with reduced data requirements; and (3) inter-sectoral flows analysis (IFA) as a hybrid model of the economic base multiplier approach and the regional input-output model.[7]

The assumption of constant technical and trade coefficients also presents a problem. In input-output models, no external economies or diseconomies can exist. But, in fact, external economies associated with localization and urbanization should be accounted for in regional analysis. Localization economies arise from the location of many plants in the same industry in close proximity to each other. Urbanization economies occur when firms in different industries locate at one locality and a corresponding urban infrastructure is built to service them. In cases where either localization economies or urbanization economies exist, the assumption of constant coefficients is inaccurate. When technological change happens, the use of constant coefficients may also be questioned (Glickman, 1977).

In view of the constant coefficients problem, two methods have been used to adjust these coefficients to account for technological and trading pattern changes. The first involves the assumption that regional coefficients change at a rate equal to national increments; the work of Almon (1966) is used to make these adjustments. A second method involves the judgment of the researchers as to future technical change (Glickman, 1977, p. 35).

In general, the economic base model and input-output model have several points in common: (1) both have been derived from or are related to the basic Leontief input-output system, (2) both are demand-oriented and view final demand for the output of an area economy as

---

[7] For more detail on these developments see Davis (1976, pp. 18-29); and Barnard and Ballman (1979, pp. 201-215).

the primary exogenous growth force in an area economy, and (3) both are concerned with certain multipliers of the local area or regional economy.

Romanoff (1974) illustrated that the economic base model is only a very special case of the input-output model. Billings (1969) showed the mathematical identity of aggregated export base multipliers and aggregated input-output multipliers. He pointed out that when the two methods are constructed with corresponding definitions and data then aggregate multipliers determined by the methods are identical.

Input-output analysis and economic base analysis possess a common defect; both present and snapshot of the economy at a given point in time. Although input-output analysis is useful in providing a detailed description of the economy, there is no guarantee that it does not embody some perturbation which will not persist into the future. Liew (1977) indicated that the conventional regional multiplier derived from a static input-output model fails to provide time path of the impact over period. Therefore, he empirically employed dynamic multipliers for the Oklahoma economy. Johnson (1979) concluded that the static input-output model cannot project the time path of local economic adjustment to changes in external conditions or internal structure and technology. He also found that the Leontief-type dynamic input-output model is impractical and incorporates behaviorally inconsistent assumptions. Therefore, he developed a modified Leontief dynamic model and applied it to a rural economy in Grant County, Oregon. Results suggest that the modified dynamic approach is superior to both the static and the Leontief dynamic formulations.[8]

---

[8] For a detailed literature review of the dynamic input-output model, see Johnson (1980, pp. 12-25).

In summary, if one is interested in more detailed information about a local economy, especially those conducting impact studies, then input-output models are a good tool to employ. But an extensive survey to compile the data and stringent assumptions about the nature of production relations are necessary if these models are to be correctly employed.

Regional Econometric Model

Econometric models are the third popular method to calculate employment multipliers. In contrast to economic base and input-output models, econometric models are not necessarily based upon a specific theory of urban strucutre and are, therefore, a more flexible research tool. Conceptual econometric models are constrained only by the broad bounds of economic theory itself (whereas applied models are constrained by data availability). With respect to costs, they do not require primary data collection as do input-output models. They are comparable in cost to economic base models, but superior since they yield far more information about the structure of a region's economy. Thus a multisector econometric approach to multiplier estimation has been generally adopted by some analysts to bridge the gap between the aggregate multiplier of the economic base type and the more disaggregate multiplier of the input-output type.[9]

In general, there are three classes of regional econometric models which can be used to estimate employment multipliers: the simple static model, the simple dynamic model, and the simultaneous equations

---

[9] For instance, Connaughton and McKillop (1979), Glickman (1977), Hall and Licari (1974), and Conopask (1978) have employed multisector econometric approach to estimate multipliers.

system model. The simple static model is a regression equation which describes the relationship between total employment and basic sectors' employment. The simple dynamic model is a regression equation in which lagged endogenous and/or exogenous variables are included in the model. The simultaneous equations system model is a set of regression equations. Each equation has one or more dependent variables which also occur in the other equation.

The general form of the simple static econometric model is

$$(7) \quad Y_t = f(X_{1t}, \ldots, X_{nt}, \mu_t)$$

where: Y_t is the endogenous variable in period t

X_{nt} is the nth exogenous variable in period t, and

μ_t is the error term in period t.

Weiss and Gooding's modified economic base model for estimating differential employment multipliers is a special case of the simple static econometric model.

The general form of the simple dynamic econometric model is:

$$(8) \quad Y_t = f(Y_{t-1}, \ldots, Y_{t-s}, X_{1t}, \ldots, X_{1t-k_1}, \ldots, X_{nt}, \ldots, X_{nt-k_n}, \mu_t)$$

where: Y_t, X_{nt}, and μ_t are defined as in equation (7),

Y_{t-s} is the endogenous variable in period t-s,

X_{nt-k_n} is the nth exogenous variable in period $t-k_n$.

In order to explain clearly the simultaneous equations system model, the structural form of the model is expressed in matrix notation as

(9) $\underset{1xg}{Y_t}\ \underset{gxg}{\Gamma} + \underset{1xg}{Y_{t-j}}\ \underset{gxg}{\beta_1} + \underset{1xk}{X_t}\ \underset{kxg}{\beta_2} = \underset{1xg}{\mu_t}$

where: Y_t is a vector of g current endogenous variables of the model,

Y_{t-j} is a vector of the same g endogenous variables in the previous period,

X_t is a vector of k exogenous variables,

μ_t is a vector of g stochastic disturbance term.

Γ, B_1, and B_2 are unknown parameter matrices of order gxg, gxg, and kxg, respecitvely.

Assuming Γ is non-singular, the reduced form can be written:

(10) $Y_t = Y_{t-j}\Pi_1 + X_t\Pi_2 + \mu_t^*$

where the coefficient matrices and reduced form stochastic disturbance terms are given as

$$\Pi_1 = -\beta_1\Gamma^{-1},\ \Pi_2 = -\beta_2\Gamma^{-1},\ \mu_t^* = \mu_t\Gamma^{-1}$$

From this dynamic simultaneous equation model, impact and total multipliers can be calculated in a generalized manner.[10] For instance, if form (10) is a linear model, and if we take first differences,

$$\Delta Y_t = \Delta Y_{t-j}\Pi_1 + \Delta X_t\Pi_2 + \Delta\mu_t^*$$

then the elements of Π_2 are known as "impact multipliers" (Goldberger, 1959). Goldberger has also shown that the multiplier effects continue to build as time progresses until we reach

---

[10] For more details about impact and total multipliers see Stewart and Venieris (1978) and Glickman (1977, p. 71).

$$Y_t = \sum_{j=0}^{\infty} \Pi_1{}^j \Pi_2 X_{t-j}$$

Theil and Boot (1962) expanded this treatment to develop a "total multiplier", i.e.,

$$(11.1) \quad \sum_{j=0}^{\infty} \Pi_1{}^j \Pi_2 = (I - \Pi_1)^{-1} \Pi_2 \ ,$$

where I is the identity matrix.

If form (10) is a non-linear model, one must use other methods to calculate multipliers (Glickman, 1977). Glickman (1977) suggests employing the Gauss-Seidel method which gives the forecasts from the present to some period S under the assumption that the exogenous variables are to move along some 'reasonable' path. Thus, we compute Y^C_T, Y^C_{T+1}, $Y^C_{T+2}, \ldots,$ Y^C_{T+S}, where the "C" denotes a controlled forecast. Then a 'perturbed solution' in which one or more of the exogenous variables is shocked by the amount δ, to obtain Y^P_T, Y^P_{T+1}, $Y^P_{T+2}, \ldots,$ Y^P_{T+S}, where "P" denotes 'perturbed'. A set of dynamic multipliers can, then be calculated as

$$(11.2) \quad (Y^P_{T+S} - Y^C_{T+S})/\delta$$

In Chapter V of this thesis, equation (11.1) is employed to estimate the dynamic employment multipliers for econometric models.

Several regional econometric models have been constructed for states or combinations of states in the United States and other countries for forecasting purposes. For example, Los Angeles (Hall and Licari, 1974), Philadelphia (Glickman, 1977), Northern California (Connaughton and Mckillop, 1979), Northern Great Plains (Conopask, 1978), and Southern California (Moody and Puffer, 1969). Most of them belong to simultaneous equations system models. In this study, the simple static and dynamic econometric models are applied to a variation of the economic-base model for calculating the dynamic employment multiplier. In turn, these dynamic employment multipliers are compared with those obtained from a transfer function model.

Problems with Regional Econometric Models

Depiste the advantages which regional econometric models may possess over alternative techniques for quantitative regional analysis, there are some shortcomings inherent in the construction of regional econometric models. We can summarize these disadvantages under two headings. The first is the general problems associated with econometric models and the second is the problems associated with applying econometric models to regions.

Within the general problems associated with econometric models, the first problem is misspecification. Often important variables are omitted from equation specifications when they should be included, resulting in biased parameter estimates. The use of time series variables that are individually serially correlated over time may also result in multicollinearity. When multicollinearity is present, the independent variables are intercorrelated, and we are unable to determine the independent effect of these variables on the dependent variable. The result of omitting a collinear variable is thus a biased regression coefficient and a specification error due to the omission.

Autocorrelation is also a serious problem which may arise from omitting a time series variable. Moreover, significant positive autocorrelation is very likely to arise in time series models. It results in downward biases in the standard deviations of regression coefficients. This may lead the model builder to wrongly claim that unimportant variables have regression coefficients which are significantly different than zero.

Another problem in regression models concerns the possibility of changing structures. Additional statistical problems also happen in the econometric models. These include the problem of correctly specifying lagged relationships, heteroscedasticity, and the problems in classifying variables as exogenous and endogenous. By using the techniques treated in standard econometrics texts, it is relatively easy to treat one of the above problems at a time. However, it is indeed difficult to treat more than one problem simultaneously.

Within problems of regional econometric models, the first concerns data. While lack of data is also a problem on the national level, this problem is especially acute for regions. Many of the data series which are available on a national level are non-existent on a regional level. An even bigger problem results when analysts are trying to use simultaneous equation estimating techniques. These techniques require even a greater number of degrees of freedom, since the reduced form must be estimated. This is a significant problem in regional models since the number of exogenous variables is relatively large in comparison to both the sample size and the number of endogenous variables.

Boundaries delineating regions are also a special problem in regional models. Often important regions of economic activities cross state lines, or important regions are only comprised of portions of states.

Although there are many theoretical and statistical problems inherent in the construction of regional econometric models, this does not mean regional econometric models are useless. After we correct the weaknesses of regional econometric models, these models are still useful in providing short-term forecasts, in assessing the probable effects of policy decisions, and in understanding the dynamic structure of a regional economy.

Theoretical Basis for Using the Transfer Function Model to Examine Employment Impacts

The very nature of the employment impacts are characterized by the persistence of dynamic relationships over long periods of time. A variety of models can be used to approximate the lagged nature of these persisntent responses. In attempting to identify a parsimonious model of a given stochastic process, one particularly useful set of processes is the family of transfer function models. The transfer function model has been shown to be useful in economic analysis as it serves to measure the impact from the hypothesized input variable upon the output variable at different time periods.[11] A detailed description of transfer function models is presented in the next chapter.

In the modeling of employment impact systems, the problems often arise in analyzing the interrelationship among different employment time series. For example, when we try to identify the relationship between Y_t and X_t, a regression equation such as

---

[11]For instance, Cramer and Miller (1976), Stokes, Jones and Neuburger (1976), Cook (1979), and Bennett (1974) have shown that the transfer function model is a useful tool for economic analysis and short-term forecasting.

$$Y_t = F(X_t, X_{t-1}, \ldots, X_{t-k})$$

is usually calculated. This approach is often hampered by severe multicollinearity except in the unlikely case where the X variables happen to be prewhitened at the outset (Stokes and Neuburger, 1979). And the diagnostic checkings of the equation will yeild very low Durbin-Watson test statistics, a sign of serial correlation of the residuals and a high degree of multicollinearity between various lags of the X variables.

The transfer function modeling techniques avoid the difficulties of multicollinearity that exist with the regression approach and will result in univariate models of each time series and multivariate models relating the different employment time series, if they are indeed related. The final fitted models will provide an accurate description of each variable and information about the lag structure of the observed relationships.

Recently, Bennett (1974) has applied time series modeling (ARIMA and transfer function) in urban and regional planning. He concluded:

> "Methodologically, the importance of the methods described here is that they provide the possibility of a statistical approach to urban and regional model building. This approach is inferential and leads naturally to the statistical estimation of the parameters of a model which is optimal in the sense of parsimony. Moreover, it also places the urban and regional modeling problem within reach of statistical forecasting and optimal control techniques of which the Box and Jenkins (1970) method is but one example...." (3, p. 172).

Cook (1979) also applied the transfer function to the economic-base model to model employment for Benton and Franklin Counties in Washington. He concluded that the transfer function is a practical tool in analysis of time series data and is potentially suitable for other applications in regional and urban modeling. Therefore, it is important for regional scientists to become familiar with this statistical method and its fitness for measurement of economic impacts and forecasting (Cook, 1979).

Therefore, the theoretical basis for using transfer function models to investigate the long-run employment impacts among government employment, lumber and wood products manufacturing employment, agriculture employment, and total employment is apparent.

CHAPTER III

OVERVIEW OF MULTIVARIATE TIME SERIES ANALYSIS

In applied economic analysis, one is frequently concerned with quantifying the association of one variable with several others. This situation is often described in terms of a time series system which relates several input variables to an output variable. A class of models which have been commonly employed in describing such variable relationships is the multiple regression model. This basic model, however, generally will not be adequate for systems which are dynamic in the sense that previous and present values of the input variables are associated with the present value of the output variable.

This chapter is primarily concerned with multi-input dynamic models (transfer function models) which are formulated via an iterative procedure of identification, estimation, and diagnostic checking. A method of identifying a multiple time series model has been thoroughly elaborated by Box and Jenkins (1970) and Haugh (1972). Nevertheless this procedure has not been widely applied due to the absence of explanations understandable to the non-statistician and reasonably priced computer algorithms.

The intent of this chapter is to 1) examine the basic differences among ARIMA, transfer function, and econometric models, 2) illustrate the transfer function modeling procedure proposed by Haugh (1972) and Haugh and Box (1977), and 3) briefly describe other modeling and estimation approaches.

Comparison with Traditional Econometric Analysis

There are currently two distinct approaches available for the analysis of economic data measured through time: the time series approach and the classifical econometric model building method. The time series approach can be classified into univariate (ARIMA)[12] and multivariate (transfer function) time series analysis. Spivey and Wecker (1971) have classified forecasting models into three groups: intrinsic, extrinsic, and hybrid. According to their classification, the ARIMA model belongs to the intrinsic models which utilize only the past history of a variable in the estimation of a forecasting model.[13] The classical econometric model is an extrinsic model which attempts to quantify relationships that exist between one or more variables and the variable to be forecast. The transfer function is a kind of hybrid model which combines the desirable functional properties of both model categories.

Before discussing the details of transfer function models, however, it will be helpful to examine the basic difference among ARIMA, transfer function, and classical econometric models and the reasons why one may select transfer function models as an alternative.

---

[12]The Box-Jenkins autoregressive-integrated-moving-average (ARIMA) model is based on the present and past observations. This model is discussed in more detail in a later chapter.

[13]Newbold and Granger (1977) have suggested that the ARIMA methodology is, on the average, superior to other intrinsic models.

Comparing ARIMA with Econometric Models

ARIMA models are flexible and usually contain relatively few parameters compared with econometric models, thus making them inexpensive and simpler to construct (Oliveira, et. al., 1977). Econometric models, of either the single equation or multi-equation variety, rely heavily on economic theory to specify causal relationships. Moreover, econometric models require a historical review of both endogenous and exogenous variables, whereas only a past history of the variable being forecast is necessary in an ARIMA model.

In contrast to the ARIMA, forecasting with econometric models requires that assumptions be made about the values of the exogenous variables in the forecast period. For short lead times, this may not be a serious shortcoming but as lead time increases, the uncertainty introduced into the forecasting exercise will grow. However, since the ARIMA model is not based on economic theory, the poor forecasts derived from it may be difficult or impossible to explain. The relative predictive performance of ARIMA and econometric models recently has been evaluated by researchers. For instance, Nelson (1973) compared the predictive performance of an ARIMA model with a large-scale econometric (FRB-MIT-Pen) model of the U.S. economy. He found that the simple ARIMA models were relatively more robust with respect to postsample prediction than the complex FMP model. Narasimham, Castellino, and Singpurwalla (1974), and Schmidt (1979) also found the ARIMA model performed better than the econometric model during the evaluation period used in their analysis. Moreover, Naylor, Seaks, and Wichern (1972) indicated that the Box-Jenkins methods seem to offer an attractive alternative to conventional econometric methods. There is some risk, however, with Box-Jenkins methods. Given that they are void of economic theory, they cannot be used to test hypothesis and to establish confidence intervals for complex economic phenomena.

Comparing ARIMA, Transfer Function, and Econometric Models

The transfer function model has characteristics of both the ARIMA and econometric models. For instance, when studying a variable Y and a leading variable X, we may fit ARIMA models to both variables and combine the models by a transfer function to yield what Box and Jenkins call a dynamic model. Therefore, the transfer function model may not be completely void of economic theory per Naylor's critique.

The transfer function and econometric models are based on quite different philosophies about what is important and how data should be approached. Granger and Newbold (1977) identified two basic differences between these two approaches. First, the transfer function model gives more attention to the lag structure relationships among variables than does the econometric model (unless the econometric models incorporate a distributed lag approach). Second, residual terms are usually treated by econometricians as being a mere nuisance of little real importance. The transfer function approach, rather than just looking at the first or second serial correlation coefficients as the econometric approach does, employs the whole correlogram of the residuals to identify the proper model. Although the econometric approach pays less attention to lags or the modeling of residuals than does the transfer function model, it emphasizes the simultaneity of the relationships among economic variables.

In summary, if economic forecasting models are to be built, it is necessary and desirable that they should incorporate as much economic theory as possible. However, it is equally important that due consideration be paid to problems of time series behavior, if one entertains the hope of producing reasonably accurate forecasts. The transfer function model does not only consider the problems of time series behavior but also incorporates the basic economic theory.[14] Therefore, if one is interested in examining the basic behavior and interaction of economic variables over time then transfer function models may be attractive alternatives or complements to econometric methods.

The Transfer Function Model

The transfer function model, commonly referred to as a rational distributed lag model, was introduced by Box and Jenkins (1966) into the statistical literature over a decade ago. Until recently, the technique has been predominantly used in engineering and the physical sciences; however, it has been successfully applied to economic analysis and control problems by Box and Jenkins (1972), Haugh and Box (1977), Granger and Newbold (1977), and Pierce (1977). Perhaps its success in economic analysis is because it attempts to measure the impact from the hypothesized input variable upon the output variable at different time periods.

The general form of the transfer function model is:

---

[14]However, there is little likelihood of extending the transfer function approach to a large number of series simultaneously due to the complexity of the method.

$$(12)\quad Y_t^* = \mu + \sum_{i=1}^{M} \frac{\omega_i(B)}{\delta_i(B)} X^*_{it-b} + \frac{\theta(B)}{\phi(B)} a_t \quad , \quad \begin{array}{l} t = 1, 2, \ldots, N \\ i = 1, 2, \ldots, M \end{array}$$

where: (1) Y_t^* and X_{it}^* are the output series and the ith input series, respectively, which are assumed to have been appropriately differenced to obtain stationarity.

(2) μ is the output mean and b is the length of pure time delay.

(3) $\delta_i(B)$, $\theta(B)$, and $\phi(B)$ are polynomial lag operators in B the backward shift operator (i.e., $B^k X_t = X_{t-k}$),

$$\delta_i(B) = 1 - (\sum_{j=1}^{r_i} \delta_{ij} B^j)$$

$$\phi(B) = 1 - (\sum_{j=1}^{p} \phi_j B^j)$$

$$\theta(B) = 1 - (\sum_{j=1}^{q} \theta_j B^j)$$

and the roots of these polynomials lie outside the unit circle.

(4) $\omega_i(B) = \omega_{io} - (\sum_{j=1}^{S_i} \omega_{ij} B^j)$

(5) The a_t's are independent, normally distributed with constant variance σ_a^2, and independent of the X_t's.

To fit such a model to a set of data, Box and Jenkins (1976) recommend the four-stage procedure of (1) prewhitening process, (2) model identification, (3) paramter estimation, and (4) diagnostic checking.[15] After these four steps have been completed, the model is

---

[15] These steps are discussed in more detail below, but in terms of the Hangh-Box approach.

assumed to provide an adequate "fit" of the data. At this point, we can consider the economic information conveyed by the coefficients of this transfer function model and employ the model for forecasting purposes.

A modified method of identifying a multivariate time series has been thoroughly elaborated by Haugh (1972) and Haugh and Box (1977). They present a procedure for the prewhitening stage that differs from the Box and Jenkins procedure in that the prewhitened output series β_t is determined by $\phi_y(B)Y_t = \theta_y(B)\ \beta_t$. Thus, rather than applying the ARIMA model developed for X_t to Y_t, one would develop a separate prewhitening model for Y_t. The cross-correlation function of the prewhitened series is then the basis for identifying the basic interrelationships present.

Haugh's work considers the single input model as a special case in a class of general multiple time series models, models for two or more time series which allow mutual input-output relationships between any pair of time series. A procedure was sought which would be applicable in all cases considered. The difference between Haugh's procedure and the Box-Jenkins' procedure seems to arise from this difference in general perspective. One could not speak of getting a prewhitening model for one of several input series and applying it to all other series in the analysis. Therefore, in this study, Haugh's model building procedures were followed in order to obtain the transfer function model.

The Transfer Function Model Building Process

In general, we can classify the Haugh and Box's modeling procedure into five phases: (1) prewhitening phase: Box and Jenkins' iterative procedures are employed to model the univariate series and obtain their residuals; (2) identification phase: the sample cross-correlation functions between residual series are studied in an attempt to identify the relationships between the series, and then postulate a general class of bivariate models; (3) estimation phase: initial and final estimate parameters of tentative model are obtained via a nonlinear least squares regression algorithm such as described by Box and Jenkins (1976); (4) diagnostic check phase: the fitted bivariate model is then checked for adequacy of fit and for significance of the parameter estimate. If these checks show inadequacies in the model, a new model is entertained, estimated, and checked. Once a final bivariate model is obtained, the relationship between the estimated residual from this model and the innovation series for second input will be checked. If these results suggest that the two series are not independent, then phases (2) through (4) are repeated to expand the bivariate model; (5) forecasting phase: after the above steps of the Haugh-Box procedure have been completed, the model is assumed to provide an adequate "fit" of the data. At this point, we can employ this model for forecasting purposes.

The steps of the approach are presented in Figure 1. Detailed descriptions of each phase are given below.

Phase I - Prewhitening

There are five stages included in Phase I. The first stage is to postulate univariate ARIMA models for each individual series; i.e., Y_t, X_{it}, i = 1, 2, $\cdots$, M. For instance, the basic form of the Box-Jenkins univariate model for Y_t is:

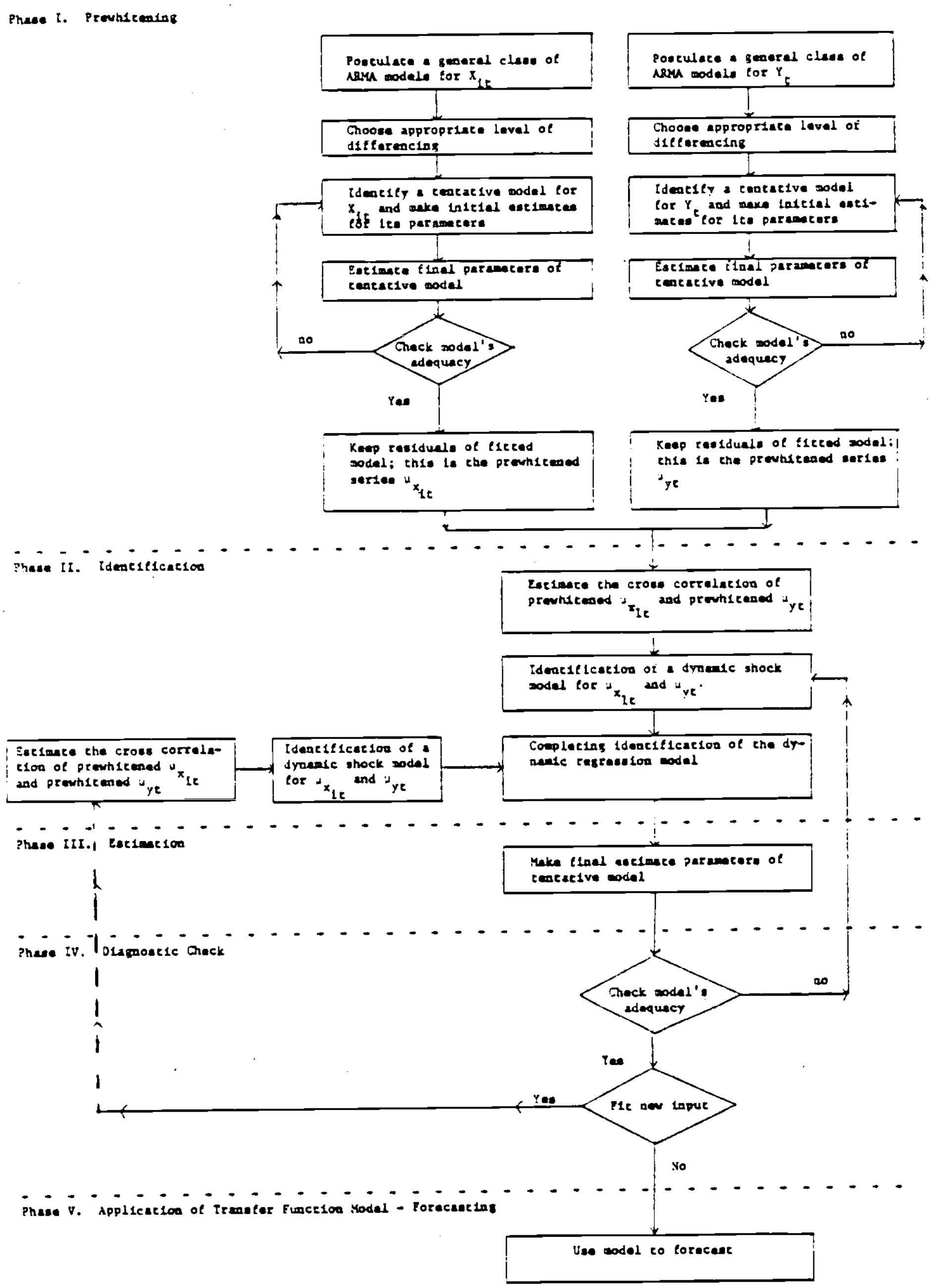

Figure 1. The Basic Steps in Developing a Transfer Function Model

$$(13)\qquad \phi_p(B)\,(Y_t-\mu) = \theta_q(B)\mu_{yt}$$

where $\phi_p(B) = 1 - (\sum_{j=1}^{p} \phi_j B^j)$, $\theta_q(B) = 1 - (\sum_{j=1}^{q} \theta_j B^j)$, μ is the mean of Y_t, and μ_{yt} is random disturbance assumed to be independently distributed as $N(o, \sigma^2)$. (See Appendix I for a more detailed discussion of ARIMA models).

If Y_t is a seasonal series, then the univariate model for Y_t is

$$(14)\qquad \phi_p(B)\Phi_K(B^S)(Y_t-\mu) = \theta_q(B)\Theta_L(B^S)\mu_{yt}$$

where $\phi_p(B)$, $\theta_q(B)$, μ, and μ_{yt} are the same as in equation (13), and

$\Phi_K(B^S) = 1-(\sum_{j=1}^{K} \Phi_j B^{Sj})$, $\Theta_L(B^S) = 1 - (\sum_{j=1}^{L} \Theta_j B^{Sj})$

Many economic time series actually exhibit nonstationary behavior and in particular do not vary about a fixed mean. This kind of nonstationarity may be characterized as being homogeneous in the sense that although the series moves freely without affinity for a particular location, its behavior at different periods in time is essentially the same. Fortunately, homogeneous nonstationarity is displayed by series whose successive changes or differences are stationary. Thus, we can reduce a homogeneous nonstationary series to a stationary series by employing a sufficient degree of differencing, i.e., $(1-B)^d Y_t$, where d denotes the degree of differencing.

Seasonality is also one of the most pervasive phenomena of economic life. Seasonal series are characterized by a display of strong serial correlation at the seasonal lag, that is, the lag corresponding to the number of observations per seasonal period, usually at multiples of that lag.

If Y_t in equation (14) is a seasonal homogeneous nonstationary series, then in order to achieve stationarity, Y_t should be replaced by Y^*_t, where $Y^*_t = (1-B)^d(1-B^S)^D Y_t = \nabla^d \nabla_S^D Y_t$, substituting Y^*_t into equation (14), we can obtain the general multiplicative seasonal model

$$(15) \quad \phi_p(B)\Phi_K(B^S)\nabla^d\nabla_S^D Y_t = \theta_q(B)\Theta_L(B^S)\mu_{yt}$$

The resulting multiplicative process will be said to be of order (p,d,q) x (K,D,L). There are two other parameters included in the empirical ARIMA models presented below which are not included in the general form (15). When the degree of differencing d and D are of order zero, Y_t is replaced by $(Y_t-\mu)$ where μ is the mean of Y_t. A deterministic trend parameter θ_o may also be included to indicate the possibility of a trend pattern within the data in the presence of nonstationary noise.

Therefore, the major task at the second stage is to choose appropriate level of differencing, i.e., the order of d and D. The estimated autocorrelation of adequate various differences are needed in this stage to decide the order of difference.

The third stage of the prewhitening phase is to identify the order of p, q, K, and L and make initial estimates for its parameters. The choice of the number and type of parameters is based on the estimated autocorrelation function (ACF) and partial autocorrelation function (PACF), i.e., these two functions are also used to make a preliminary identification of the model (see Appendix II). Characteristic behavior of these tools for three classes of processes is summarized in Table 1 [16] (Nelson, 1973).

Having identified one or more tentative models for a time series, we would like to obtain the best or most efficient estimates of the parameters. This is the goal of Stage 4.

Maximum-likelihood estimates which are closely approximated by the least squares estimates are utilized here to estimate final parameters.

After the model has been identified and the parameters estimated, diagnostic checks are then applied to the fitted model. One useful method of checking a model is to overfit, that is, to estimate the parameters in a model somewhat more general than that which we believe to be true. This method assumes that we can guess the direction in which the model is likely to be inadequate. These checks employ the autocorrelation function of the residuals. Suppose that we have th first k autocorrelations [17] $\hat{r}_k(\mu)$ (k = 1, 2, ...K) from one ARIMA (p,d,q) model, then it is possible to show that, if the fitted model is appropriate,

$$(16) \quad Q_1 = n \sum_{k=1}^{K} \hat{r}_k^2(\mu)$$

is approximately distributed as $\chi^2(k-p-q)$. If the statistic is not significant at some level α, the residual series is assumed to be random and the model accepted. If the residual series is not random, a new

---

[16] Box and Jenkins (1976, pp. 176-177) present a more detailed description.

[17] It is assumed here that k is selected to be sufficiently large so that weights ψ_j in the model, written in the form $Y_t = \phi^{-1}(B)\theta(B)\mu_t = \psi(B)\mu_t$ will be negligibly small after j=k.

TABLE 1. CHARACTERISTIC BEHAVIOR OF AUTOCORRELATIONS AND PARTIAL AUTOCORRELATIONS FOR THREE CLASSES OF PROCESSES

Class of Processes	Autocorrelations	Partial autocorrelations
Moving Average (MA(q))	Spikes at lags 1 through q, then cut off	Tail off
Autoregressive (AR(p))	Tail off according to $\rho_j = \phi_1\rho_{j-1} + \cdots + \phi_p\rho_{j-\rho}$ 1/	Spikes at lags 1 through p, then cut off.
Mixed autoregressive - Moving Average (ARMA(p,q))	Irregular pattern at lags 1 through q, then tail off according to $\rho_j = \phi_1\rho_{j-1} + \cdots + \phi_p\rho_{j-\rho}$	Tail off

1/ These are usually called the Yule-Walker equations (Box and Jenkins, 1976, p. 55).

Source: Nelson, Charles R: "Applied Time Series Analysis for Managerial Forecasting," Holden-Day, Inc., San Francisco, 1973, p. 89.

identification is made and the process of estimation and checking is repeated.

The final step in the prewhitening phase is to keep residuals of fitted model (i.e., $\mu_{x_{it}}, \mu_{yt}$) for example, from equation (15)

$$\mu_{yt} = \frac{\phi_p(B)\Phi_k(B^S)\nabla^d \nabla_S^D Y_t}{\theta_q(B)\Theta_L(B^S)}$$

for checking the relationship between the individual series.

Phase II - Identification

The identification phase is introduced for relating two (or more) time series through a dynamic regression (distributed lag) model. This phase contains four steps. The first step is to check the estimated univariate residuals cross correlation function $r_{\hat{\mu}_x\hat{\mu}_y}(k)$, and then to identify the relationship between X and Y. Depending on the type of pattern in $r_{\hat{\mu}_x\hat{\mu}_y}(k)$, Pierce (1977) and Pierce and Haugh (1977) have shown some possible causality patterns and associated patterns in $r_{\hat{\mu}_x\hat{\mu}_y}(k)$, which are reproduced in Table 2 (Pierce, 1977).

After identifying the relationship between X and Y, we can then build the general dynamic regression model between X and Y. For instance, if relationship 6 exists, then we can build a dynamic regression model for Y on X and thereby improve the forecasting ability of Y.

Haugh (1972) has suggested two alternative chi-square tests[18] to check the independence between the innovation series, i.e.,

---

[18] Haugh suggests the use of S_M^* when M is large relative to N.

TABLE 2. CONDITIONS ON CROSS CORRELATIONS OF THE PREWHITENED SERIES FOR CAUSALITY PATTERNS

Relationship	Restrictions on $r_{\hat{\mu}_x \hat{\mu}_y}(k)$
1. X causes Y	$r_{\hat{\mu}_x \hat{\mu}_y}(k) \neq 0$ for some $k>0$
2. Y causes X	$r_{\hat{\mu}_x \hat{\mu}_y}(k) \neq 0$ for some $k<0$
3. Instantaneous causality	$r_{\hat{\mu}_x \hat{\mu}_y}(0) \neq 0$
4. Feedback	$r_{\hat{\mu}_x \hat{\mu}_y}(k) \neq 0$ for some $k>0$ and for some $k<0$
5. Y does not cause X	$r_{\hat{\mu}_x \hat{\mu}_y}(k) = 0$ for all $k<0$
6. Unidirectional causality from X to Y	$r_{\hat{\mu}_x \hat{\mu}_y}(k) \neq 0$ for some $k>0$ and $r_{\hat{\mu}_x \hat{\mu}_y}(k) = 0$ for all $k<0$
7. X and Y are related instantaneously but in no other way	$r_{\hat{\mu}_x \hat{\mu}_y}(k) = 0$ for all $k \neq 0$ and $r_{\hat{\mu}_x \hat{\mu}_y}(0) \neq 0$
8. X and Y are independent	$r_{\hat{\mu}_x \hat{\mu}_y}(k) = 0$ for all k

Source: Pierce, David A. "Relationships - and the lack thereof - Between Economic Time Series, with Special Reference to Money and Interest Rates." Journal of the American Statistical Association. Vol. 72, No. 357 (March, 1977): 15.

$$(17)\quad S_M = N\sum_{k=-M}^{M} r^2_{\hat{\mu}_x\hat{\mu}_y}(k)$$

and

$$(18)\quad S^*_M = N^2\sum_{k=-M}^{M} (N-|k|)^{-1}\ r^2_{\hat{\mu}_x\hat{\mu}_y}(k)$$

where $r_{\hat{\mu}_x\hat{\mu}_y}(k)$ is the estimated cross-correlation between innovation series μ_x and μ_y at lag k, M is chosen large enough to include expected nonnegligibly nonzero coefficients, and N is the number of observations in each residual series. Both S_M and S^*_M have asymptotic distribution χ^2_{2M+1} under the hypothesis of the series being independent. If the value of S_M (or S^*_M) exceeds the selected table value, one can reject the hypothesis of independence and suspect that X and Y are related.

At the second stage of identification, assuming the chi-square test suggests that a relationship exists between X and Y, the pattern of the cross-correlation function is then used to identify the dynamic shock model for μ_x and μ_y. That is,

$$(19)\quad \mu_{yt} = \frac{\omega'(B)}{\delta'(B)}\mu_{xt} + \frac{\theta'(B)}{\phi'(B)}a_t = \nu'(B)\mu_{xt} + \psi'(B)a_t,$$

where $\omega'(B)$, $\delta'(B)$, $\theta'(B)$, $\phi'(B)$ are all polynomial lag operators. Initial parameter estimates for the prameters appearing in the identified forms of $\nu'(B) = \frac{\omega'(B)}{\delta'(B)}$ may be obtained in Box and Jenkins' procedure (1976, pp. 346-351).[19]

---

[19] Box and Jenkins use one prewhitening filter, and get dynamic regression model

$$Y_t = \frac{\omega(B)}{\delta(B)}x_t + \frac{\theta(B)}{\phi(B)}a_t = \nu(B)\,x_t + \psi(B)a_t$$

The cross correlation function $r_{\hat{\mu}_x \hat{\mu}_y}(k)$ gives a direct indication of $\nu'(B)$; but we may also easily identify the form of $\psi'(B) = \theta'(B)/\phi'(B)$ by making use of the fact that μ_y is white noise. Haugh (1972, pp. 108-110) indicated that $\phi'(B) = \delta'(B)$ and that $\theta'(B)$ is at most of order γ' or S', where γ' and S' are respectively the orders of $\delta'(B)$ and $\omega'(B)$.

The final stage of the identification phase is to substitute the identified univariate models into the identified dynamic shock model giving a preliminary transfer function model, i.e.[20]

$$(20.a.) \qquad y_t^* = \frac{\nu'(B)\phi_x(B)\theta_y(B)}{\theta_x(B)\,\phi_y(B)}\,x_t^* + \frac{\psi'(B)\theta_y(B)}{\phi_y(B)}\,a_t$$

Phase III - Estimation

Final parameter estimates for model (20) are then obtained via a nonlinear least squares regression algorithm such as described by Box and Jenkins (1976, pp. 388-391). This stage is similar to the estimation procedure for the development of an ARIMA model for a single time series.

Phase IV - Diagnostic Check

The fitted bivariate model is then checked for adequacy of fit and for significance of the parameter estimates. Box and Jenkins (1976) suggest employing the residual autocorrelations and cross correlations to check the accuracy of the transfer function model and the noise model.

Suppose after suitable differencing, if necessary, that the model

---

[20]Suppose the identified univariate models for x and Y are $\phi_y(B)y_t^* = \theta_y(B)\mu_{yt}$ and $\phi_x(B)x_t^* = \theta_x(B)\mu_{xt}$, respectively. The identified dynamic shock model is $\mu_{yt} = \nu'(B)\mu_{xt} + \psi'(B)a_t$.

can be written

$$(20.b)\quad y_t^* = \frac{\omega(B)}{\delta(B)} x_{t-b}^* + \frac{\theta(B)}{\phi(B)} a_t = \nu(B)x_t^* + \psi(B)a_t$$

$$\text{(where } \phi_x(B)x_t^* = \theta_x(B)\mu_t\text{).}$$

Now, suppose that we select an incorrect model leading to residuals a_{ot}, where

$$y_t^* = \nu_o(B)x_t^* + \psi_o(B)a_{ot} \ ,$$

then

$$(21)\quad a_{ot} = \psi_o^{-1}(B)\{\nu(B) - \nu_o(B)\}\ x_t^* + \psi_o^{-1}(B)\psi(B)a_t \ ,$$

whence if the transfer function model is correct, but the noise model is incorrect, (i.e., $\nu_o(B) = \nu(B)$ but $\psi_o(B) \neq \psi(B)$), then (21) becomes

$$a_{ot} = \psi_o^{-1}(B)\psi(B)a_t$$

Therefore, the a_{ot} process would be autocorrelated, and the form of the autocorrelation function could indicate the appropriate modification of the noise structure.

Under this situation, we can apply the autocorrelation check

$$(22)\quad Q_2 = M\sum_{K=1}^{k} r_{\hat{a}\hat{a}}^2(K),$$

where Q_2 is approximately distributed as χ^2 with k-p-q degrees of freedom.[21] M = N-u-p, u is the larger of r and S+b. $r_{\hat{a}\hat{a}}^2(K)$ is the estimated residual autocorrelations for K = 1, 2, ... k, and k is chosen

---

[21] Note that the degrees of freedom in χ^2 depends on the number of parameters in the noise model but not on the number of parameters in the transfer function model.

sufficiently large so that the weights ψ_j in (20.6) can be expected to be negligible for $j>K$. If the value of Q_2 exceeds the selected table value, this suggests that the noise model is inadequate.

If the transfer function model were incorrect, from (21) it is apparent that, not only would the a_{ot}'s be cross correlated with the x_t's (and the μ_t's which generate the x_t's), but also the a_{ot}'s would be autocorrelated. This would be true even if the noise model were correct. Under this situation, a cross correlation analysis could indicate the modifications needed in the transfer function model.

Therefore, the cross correlation checks can indicate inadequacy of the transfer function model, i.e.,

$$(23) \quad Q_3 = M \sum_{K=0}^{k} r^2_{\hat{\mu}\hat{a}}(K)$$

where Q_3 is approximately distributed as χ^2 with $k + 1 - (r+S+1)$ degrees of freedom, $(r+S+1)$ is the number of parameters fitted in the transfer function model. $r^2_{\hat{\mu}\hat{a}}(K)$ is the estimated residual cross correlation for $K = 0, 1, 2, \cdots k$. If the value of Q_3 exceeds the selected table value, this suggests inadequacy of the transfer function model.

If the residual autocorrelation and cross correlation checks indicate a good data fit, the next step will be to consider whether it is necessary to expand the one input transfer function to a two input transfer function model. In order to make this decision, we must know whether x_{2t} will be of some additional value in describing Y_t (conditional on x_{1t}) by examining the cross correlations between $\hat{\mu}_{2t}$ and the residuals $\hat{a}_t$ of equation (20.b). If the cross correlations suggest that the two series are not independent, then the inclusion of x_{2t} in the transfer function of Y_t and x_{1t} will be considered. Therefore, Phases II through IV are repeated to expand the transfer function model into two inputs series. If more than two inputs series are proposed, the same steps are repeated for each input series.

Phase V. Application of Transfer Function Model - Forecasting

Forecasting with transfer function models often involves applying the final parameter values to periods for which data are not available. Depending upon the time delay, b, forecasting in such instances is based on past values of X_t and Y_t. It is at this point that the true benefits of a transfer function model can be realized. If X_t is a good leading indicator of Y_t, then Y_t can be accurately forecasted when cyclical changes occur and even X_t values can be forecasted using the univariate model generating X_t.

Other Modeling and Estimation Approaches

In addition to the controversy involving the choice of a prewhitening procedures for Y_t, there are also disagreements about alternative procedures for testing the relationship between economic variables. Feige and Pearce (1979) have examined the conceptual relationship between three popular empirical procedures[22] used for tests of the causal relationships between money and income. They concluded that an essentially arbitrary choice can significantly affect the nature of the economic conclusions derived from the test procedures. In this study, the Haugh-Pierce approach was used to check the relationships between employment variables, and it was also found that the empirical results were not consistent with what one would expect based upon economic theory.

The approach of Priestley is also noted by Haugh. It is philosophically similar to Haugh's approach, but it employs a less general form of the transfer function model and involves an alternative to the cross-correlation function of the prewhitened series, which Priestley labels "covariance contraction."

Recently, attention has turned to the problem of modeling a pair of series related by a two-way causality or feedback mechanism. If this situation exists, then more complicated methods need to be used to model the relationship between the innovations and consequently the original series. These methods and some possible uses for them are described by Granger and Newbold (1977).

---

[22]
(1) The Haugh-Pierce approach.
(2) The Direct Granger Approach.
(3) The Sims Approach.

CHAPTER IV

EMPIRICAL ANALYSIS

Chapter III discussed the conceptual theory and the modeling procedures of the univariate ARIMA model and the transfer function models. In this chapter, we apply these modeling procedures to develop the univariate ARIMA and transfer function models for the Grant County employment data. Simple static and dynamic econometric models are also specified, and their estimated parameters are presented. Finally the forecasting accuracies of the ARIMA, transfer function and econometric models are compared. This chapter is an intermediate step between the theoretical framework of the models and their application to the employment impact analysis.

Data Requirements, Sources, and Limitations

In order to conduct the economic base analysis and, consequently, the transfer function analysis and econometric analysis, the basic industries in Grant County should be identified. In this study, the location quotient technique (LQ) was used to select basic industries, i.e.,

$$LQ = \frac{\frac{X_t}{E_t}}{\frac{X_t^*}{E_t^*}}$$

where X_t = employment in industry X in the Grant County in year t,

X_t^* = employment in industry X in the United States (or Oregon) in year t,

E_t = total employment in the Grant County in year t,

E_t^* = total employment in the United States (or Oregon) in year t.

If LQ > 1, then industry X in Grant County is considered a basic or exporting oriented.

From Table 3 we can see that farm industry, agricultural service, forestry, fish and other industries, mining industry, manufacturing industry, and government and government enterprises are identified as basic industries in Grant County for 1977 (similar results were obtained for other years). Within these basic industries, agriculture service, forestry, fish and other industries, and mining industry include only 46 employees. Therefore, it did not seem worthwhile to include these industries as basic industries for the quantitative analysis in this study. From Oregon Employment Division statistics, we find that most of the manufacturing industry in Grant County produces durable goods (i.e., lumber and wood products). Therefore, lumber and wood products manufacturing industry is selected to represent the manufacturing industry as a basic industry. There are three categories included in government and government enterprises, i.e., federal (civilian), federal (military) and state and local. From the location quotient for federal (military) employment, we find it is not a basic industry in Grant County. However, only total government employment, including military, is identified as a basic industry. The location quotient for total government employment also justifies this assumption.

TABLE 3: Employment and Location Quotients of Grant County by Type and Broad Industrial sources, 1977

Item	Grant County	U.S. (LQ) 1/	State (LQ 2/
Employment by place of work			
Total Employment 3/	3,418		
Number of Proprietors	725		
Farm Proprietors	350		
Nonfarm Proprietors	375		
Total Wage and Salary Employment	2,693		
Farm	206	5.0956	2.5512
Nonfarm	2,487	0.9376	0.9521
Private	1,515	0.7170	0.7333
Ag. Serv., For., Fish, and Other 4/	21	1.7180	1.2105
Mining	25	1.0003	5.1515
Construction	30	0.2551	0.2592
Manufacturing	734	1.2288	1.3139
Non-Durable Goods	(D)	(D)	(D)
Durable Goods	(D)	(D)	(D)
Transportation and Public Utilities	73	0.5115	0.5031
Wholesale Trade	21	0.1461	0.1337
Retail Trade	328	0.7797	0.7223
Finance, Insurance, and Real Estate	48	0.3502	0.3583
Services	235	0.4550	0.4905
Government and Government Enterprises	972	1.8010	1.7794
Federal, civilian	271	3.1200	3.7637
Federal, military	37	0.5126	0.9497
State and local	664	1.7444	1.5254

Source: Regional Economic Information System. Bureau of Economic Analysis. U.S. Department of Commerce

1/ The reference area for these location quotients is United States.

2/ The reference area for these location quotients is Oregon State

3/ Consists of wage and salary jobs plus number of proprietors.

4/ Includes number of jobs held by U.S. residents working for international organizations. Primary source for private non-farm wage and salary employment: Oregon Employment Division.

(D) Not shown to avoid disclosure of confidential information. Data are included in totals.

In summary, government employment, lumber and wood products manufacturing employment, and agriculture employment are selected as basic industries in Grant County for the economic base analysis to follow.

All of these three basic industries employment and total employment data were supplied by the Department of Human Resources of the State of Oregon. A difference exists between the available industries employment and total employment series. The total employment series are adjusted for multiple job-holding and commuting while the industries employment series are not. Comformable total employment and industries employment data are not available. The significance of this difference in Grant County is not assessed in this study. However, the more significant the difference, the lower the estimate of the multiplier.

The sample period for the monthly employment series was from January, 1970 to December, 1977, resulting in 96 observations. The test period was from January, 1978 to December, 1979, resulting in 24 observations which were used to check the forecasting reliability of the fitted model.

TABLE 4

Identification of Data Series[a]

Abbreviation	Series
TEG	Total employment in Grant County
GEG	Government employment in Grant County
LWP	Lumber and Wood products manufacturing employment in Grant County
AGG	Agricultural employment in Grant County

[a]The data series were provided by the Dept. of Human Resources, State of Oregon.

Univariate ARIMA Model

Identification

Four data series were examined and are shown in Table 4. In terms of the notation employed in the previous section, TEG = Y_t, GEG = X_{1t}, LWP = X_{2t}, and AGG = X_{3t}. They are plotted in Figures 2 and 3. All of the plots appear to be nonstationary and each has a very distinct, but similar, seasonal pattern. The yearly peak in these four series occurs mainly in either July or August. The yearly trough of TEG, GEG, and LWP usually occurs in either February or March. However, the yearly trough of AGG occurs in December.

The estimated autocorrelation and partial autocorrelation functions for series TEG, GEG, LWP, and AGG are shown in Appendix III. Plotting of the correlation fucntions greatly assists in their interpretation. For illustration, the autocorrelations and partial autocorrelations for series GEG are plotted in Figures 4 and 5. From these two figures, we find the autocorrelations for GEG are large and fail to die out at higher lags while simple differencing reduces the correlations in general, a very heavy periodic component remains. This is evidenced particularly by very large correlations at lags 12, 24, and 36. Simple differencing with respect to period twelve results in correlations which are first persistently positive and then persistently negative. By contrast, the differencing $\nabla\nabla_{12}$ markedly reduces correlations throughout.

We note that since the autocorrelations cut off, our model will be of moving-average form. Now, in terms of the multiplicative moving-average model, the maximum power of B in the model will be

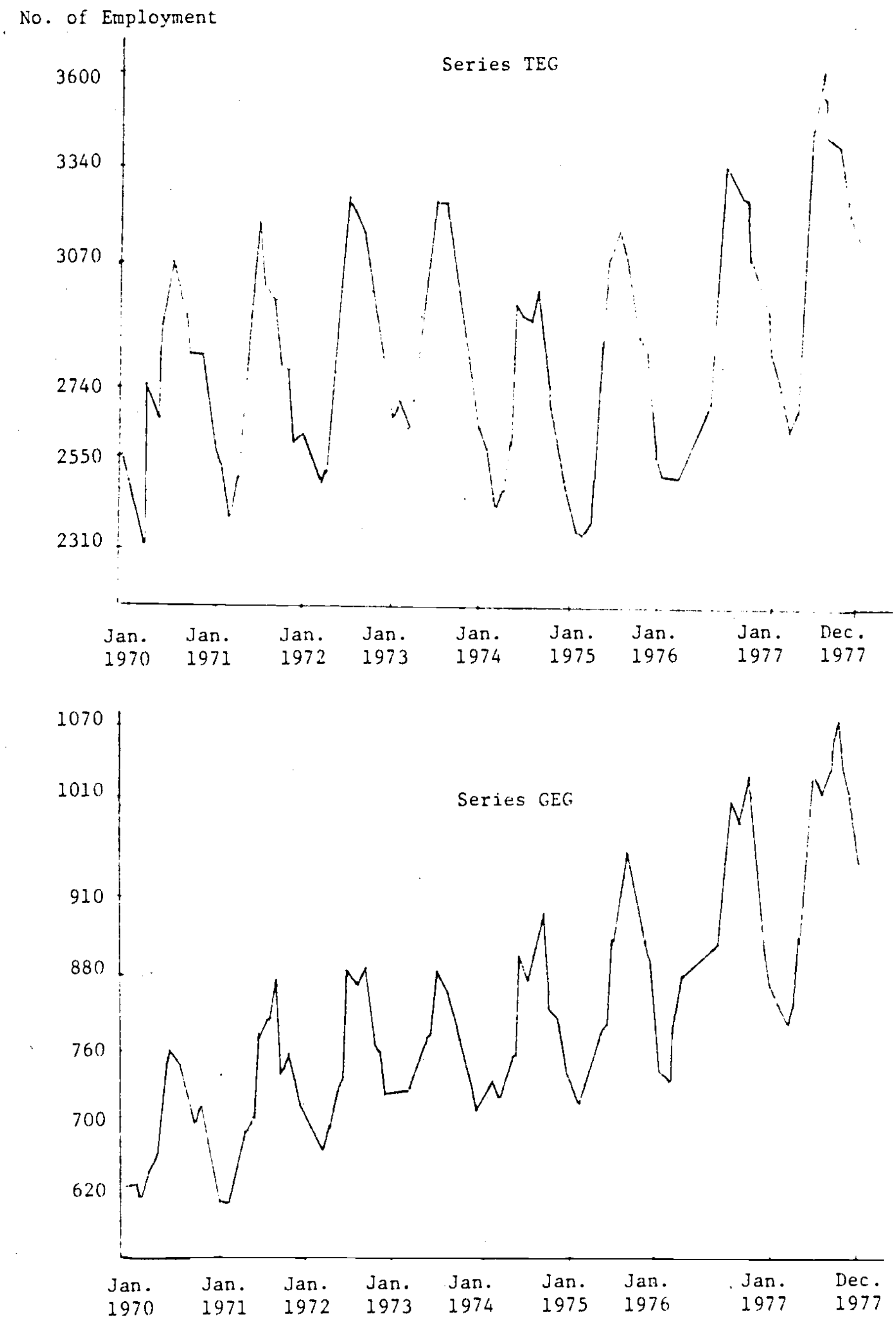

Figure 2. Series TEG and GEG, January, 1970 - December, 1977.

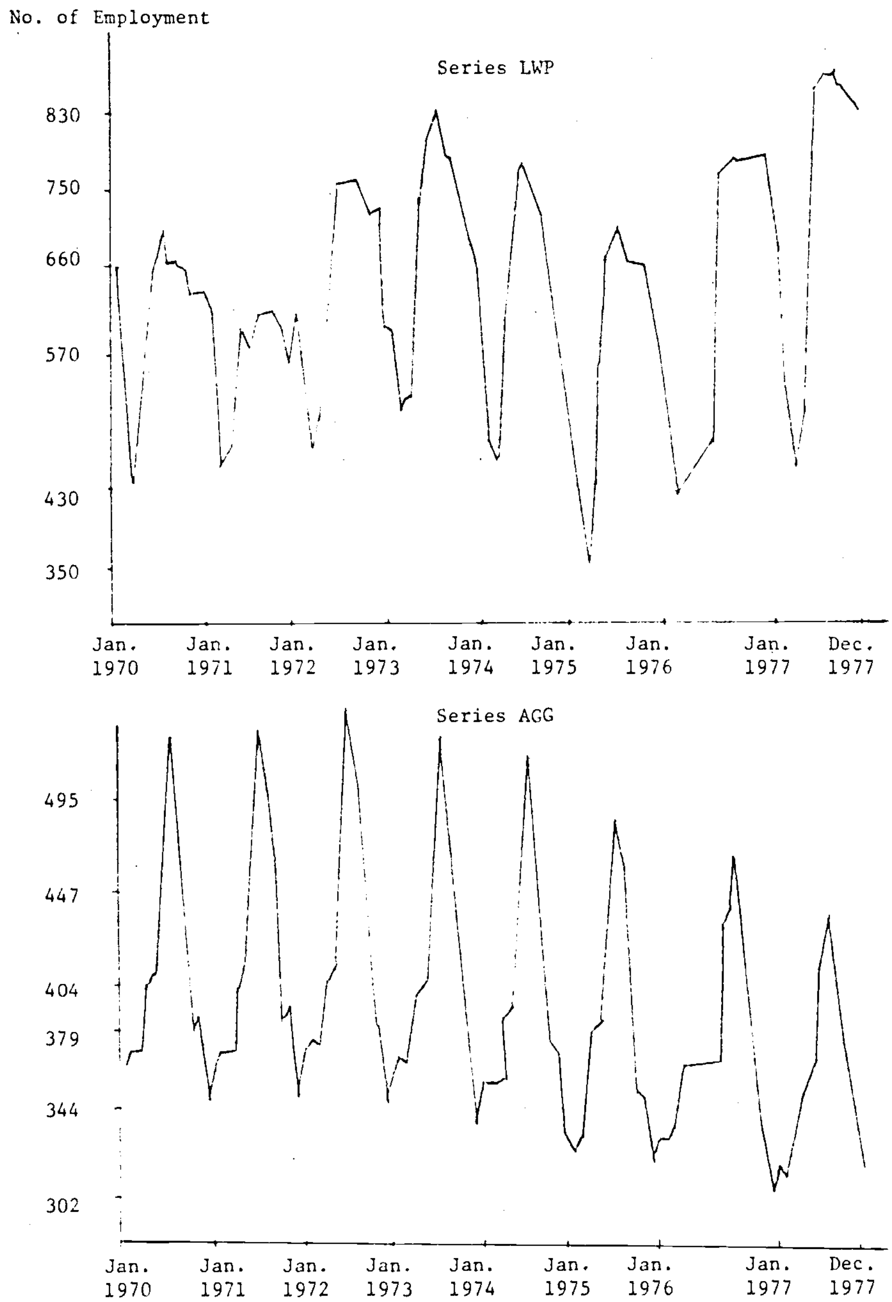

Figure 3. Series LWP and AGG, January, 1970 - December, 1977.

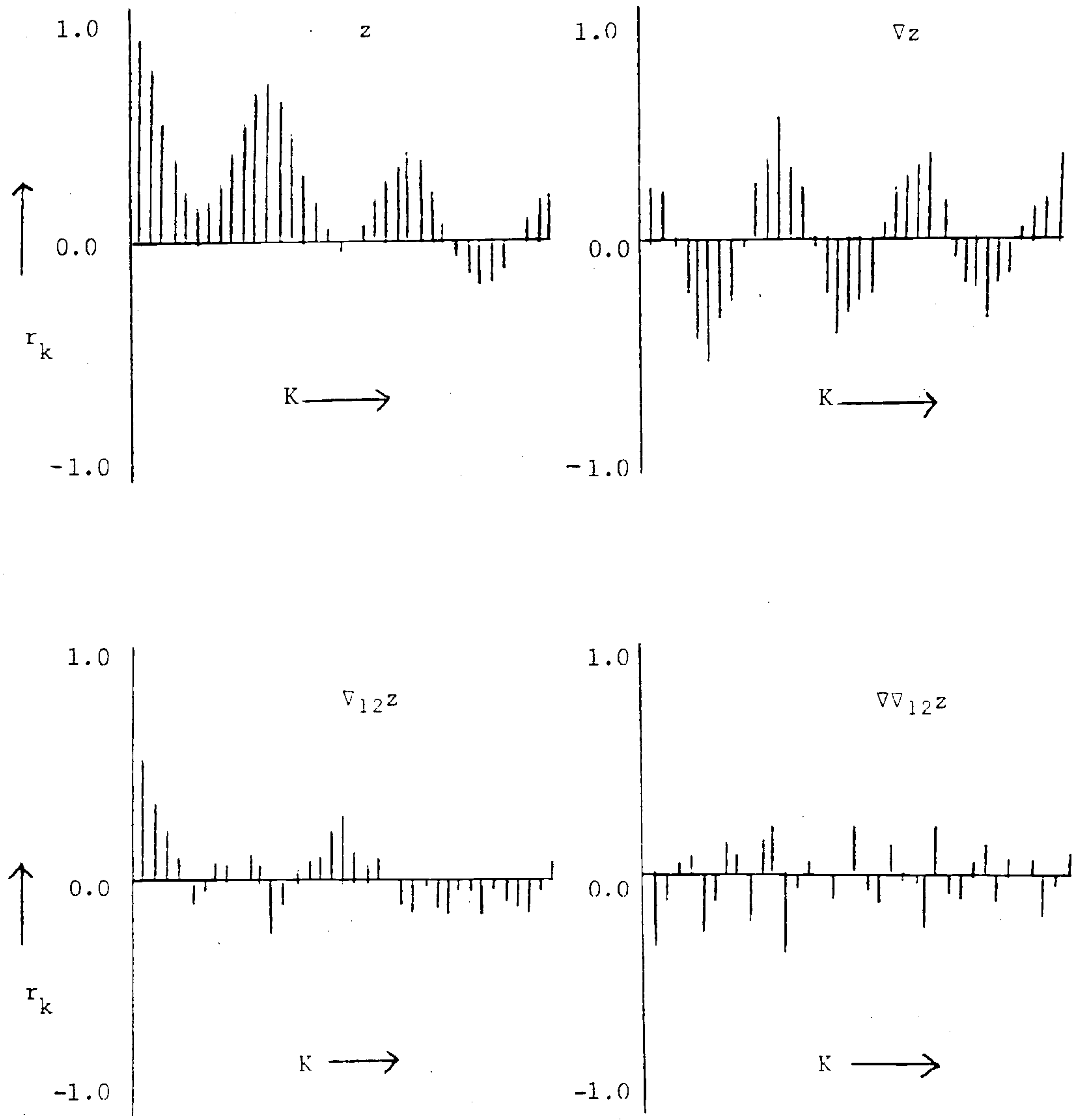

Figure 4. Estimated Autocorrelations of Various Differences of Series GEG.

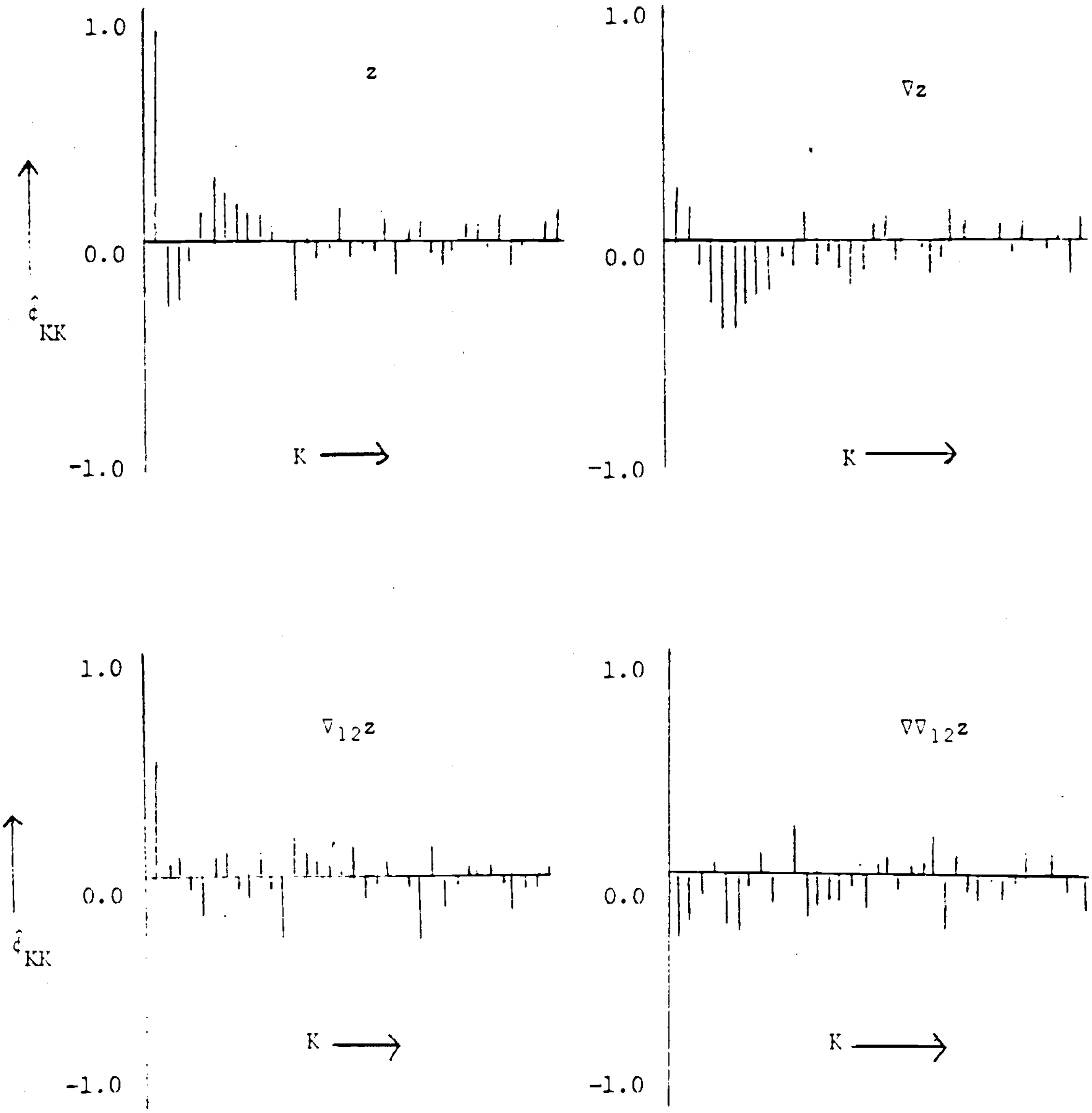

Figure 5. Estimated Partial Autocorrelations of Various Differences of Series GEG.

S=12, and values of Q and q are 1. Our tentative model for the GEG series is then

$$(1-B)\ (1-B^{12})\ GEG_t = \theta_0 + (1-\theta_1 B)\ (1-\Theta_1 B^{12})\mu_{1t}\ .$$

The preliminary estimates of the parameters θ_1 and Θ_1 can be obtained by substituting the sample estimates $r_1 = -0.29$ and $r_{12} = -0.31$ in the expressions

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2}\ ,\ \rho_{12} = \frac{-\Theta_1}{1+\Theta_1^2}$$

Thus, we obtain rough estimates $\hat{\theta}_1 \cong 0.325$ and $\hat{\Theta}_1 \cong 0.345$, and we also can estimate $\hat{\theta}_0 \cong 1.56$ directly from Appendix III.

From the estimated autocorrelation and partial autocorrelation Functions for TEG, LWP and AGG series, their ARIMA models are also tentatively identified to be

$$(1-B)\ (1-B^{12})\ TEG_t = 2.16 + (1 - 0.48B)\ (1 - 0.295B^{12})\mu_{yt}$$

$$(1-0.84B)\ (1-B^{12})\ LWP_t = 15.47 + (1 - 0.148B^{12})\mu_{2t}$$

$$(1-0.65B)\ (1-B^{12})\ AGG_t = \mu_{3t}$$

Estimation and Diagnostic Check

These tentative models were estimated by employing a nonlinear least squares regression algorithm and, then, checked for adequacy of fit as described by Box and Jenkins (1975) and for significance of the parameter estimates. Modifications were found necessary to be made to these models, therefore, they were reestimated. The final ARIMA models for the various employment series are presented in Table 5 along with the usual information concerning the "Goodness of fit" of ARIMA models.

Table 5. Estimated Univariate ARIMA Models

Series	Model[a]	χ^2(D. of F.)[b]	RSE[c]	σ_a[d]
TEG	$\underset{(0.042)}{(1-0.968B)}\ (1-B^{12})\ Y_t = \underset{(0.116)}{(1-0.444B)}\ \underset{(0.115)}{(1-0.368B^{12})}\mu_{yt}$	21.96 (27)	87.68	85.737
GEG	$(1-B)\ (1-B^{12})\ X_{1t} = (1-\underset{(0.101)}{0.409}B - \underset{(0.105)}{0.195}B^5)\ \underset{(0.105)}{(1-0.549B^{12})}\mu_{1t}$	21.11 (27)	32.6	31.68
LWP	$\underset{(0.056)}{(1-0.871B)}\ (1-B^{12})X_{2t} = \underset{(0.110)}{(1-0.421B^{12})}\mu_{2t}$	12.83 (28)	40.54	39.86
AGG	$\underset{(0.065)}{(1-0.805B)}\ (1-B^{12})X_{3t} = \mu_{3t}$	23.91 (29)	8.005	7.795

a/ The numbers in parenthesis below the estimated coefficients are estimated standard errors.

b/ Chi-square statistic with degrees of freedom given in parenthesis.

c/ Residual Standard Error.

d/ σ_a = standard deviation of the estimated residual series.

Forecasting Accuracy Check

The final models for each employment series were employed as forecasting equations, and for each forecast value an upper and lower 95 percent confidence limit was estimated. As an illustration, the forecast values for a 24-month forecasting period along with the actual values for the GEG series are shown in Figure 6. The selected forecast origin is December, 1977. For measuring the forecasting accuracy, six types of goodness-of-fit measures were calculated for each series at forecast periods of lengths 12 and 24 months. These six tests are correlation coefficient (R), regression coefficient of actual on prediction (β_1), mean error (ME), mean absolute error (MAE), root mean square error (RMSE), and Theil's inequality coefficient (U) (see Appendix IV).

The forecasting ability of the various employment series models can be compared from these six types of goodness-of-fit measures presented in Table 6. In general, a disadvantage of the R is that perfect correlation only implies an exact linear relationship between predicted and actual values. For forecasts to be unbiased, and, therefore, perfect, regression parameters of $\beta_1 = 1$ and $\beta_0 = 0$ must also exist. For this situation, we find AGG and GEG models have higher R and β_1 approximates 1 at the same time. The TEG model has higher R, but β_1 is 0.67 and 0.66 for 12-month forecast and 24-month forecast, respectively. ME, MAE, and RMSE will be highly influenced by the average size of the variable. For instance, the TEG series model has the highest ME, MAE and RMSE for 12-month forecasts and 24-month forecasts, partly because this sector has a higher average size of employment.

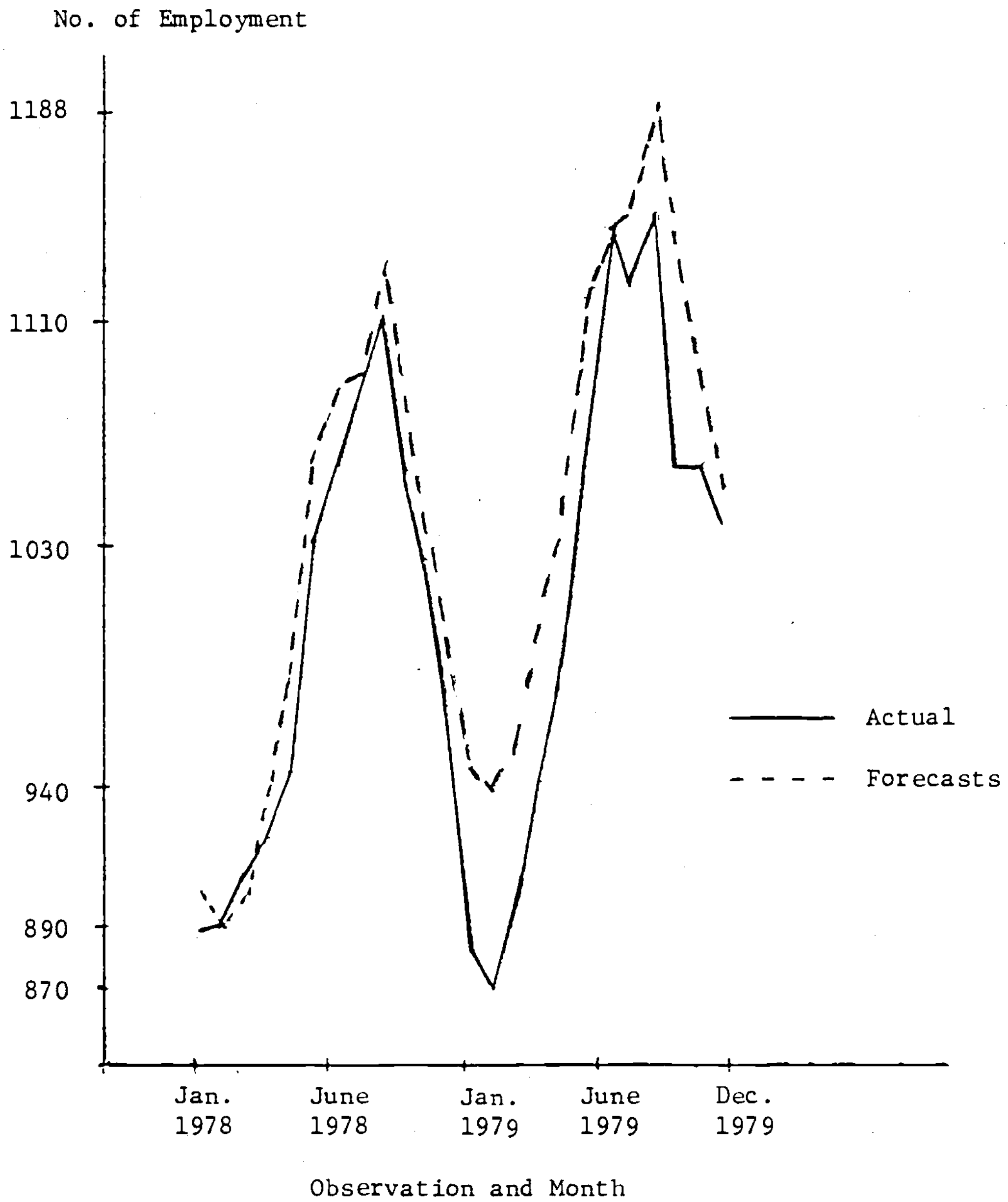

Figure 6. GEG Series: Actual 1978, 1979 Values, and 24-month Forecast Values for Forecast Origin , December 1977.

Table 6. The Forecast Accuracy Test for Four Employment Series. a/

Series	Forecasting Period (months)	R	β_1	ME	MAE	RMSE	U	(I	II	III)
TEG	12	0.962	0.67	-22.16	99.18	113.3	0.0176	(0.038	0.604	0.358)
	24	0.957	0.66	-68.95	112.6	135.0	0.0209	(0.261	0.462	0.277)
GEG	12	0.988	0.93	-15.94	17.55	20.46	0.0102	(0.606	0.056	0.337)
	24	0.961	0.97	-30.55	31.36	38.88	0.019	(0.617	0.92×10^{-2}	0.381)
LWP	12	0.852	0.70	-33.80	68.03	73.39	0.0504	(0.212	0.095	0.693)
	24	0.797	0.59	-39.53	77.73	86.03	0.059	(0.211	0.169	0.619)
AGG	12	0.998	0.99	1.54	2.65	2.92	0.004	(0.278	0.004	0.718)
	24	0.995	0.99	-1.47	3.58	4.46	0.0061	(0.108	0.0021	0.889)

a/ R = Correlation coefficient

β_1 = Regression coefficient of actual on prediction

ME = Mean error

MAE = Mean absolute error

RMSE = Root mean square error

U = Theil inequality coefficient which can decompose into three parts: (I) Fraction of error due to bias; (II) Fraction of error due to different variation; (III) Fraction of error due to different co-variation.

The Theil's inequality coefficient, as proposed by Theil (1966) can be decomposed into three terms: U(Bias), U(variation), and U (covariation). As forecasts will not all be perfect, the goal should be the lowest inequality coefficient possible with a decomposition showing U(bias) and U(variation) approaching zero and U(covariation) approaching one. Observing the 12-week forecasts for the employment series models, the AGG ARIMA model appears to be the most accurate with a Theil's inequality coefficient of 0.004 for 12-month forecast and 0.0061 for 24-month forecast. The U(bias) and U(variation) show approaching zero and U(covariation) approaching one.

The largest Theil's inequality coefficients are those of the LWP series. One plausible explanation of the relatively better predictability of the AGG series is that for this series the variability of the employment was relatively less whereas the LWP series has a relatively high variability of the employment. Although the accuracy of the LWP series model forecasts was not as good as for the other series, in general, the Theil's inequality coefficients for LWP series model are very low. Therefore, we can conclude that these four models appear to have good forecasting ability.

We can also use ' the mean absolute percentage errors for selected forecasting horizongs ' to show the good forecasting ability of these four ARIMA models (see Table 7). The forecasting accuracy of each model was evaluated by computing the percentage error for different forecasting horizons. Then the mean absolute percentage errors for each model and forecast horizon (1, 4, 12 and 24) were computed by averaging across the various origins. Here, the forecasting origins were 96 (December, 1977), 100 (April, 1978), 104

TABLE 7

Mean Absolute Percentage Forecasting Errors
for Selected Forecast Horizons[a]

Series Model	Forecasting Horizon (months) 1	4	12	24
TEG	1.345	2.401	2.496	3.658
GEG	2.447	2.299	2.764	3.165
LWP	7.450	9.278	8.928	11.725
AGG	1.547	1.354	1.441	1.147

[a] Mean absolute percentage forecasting errors

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{|\hat{Y}_t - Y_t|}{Y_t} \cdot 100\%$$

where

T = the number of forecast periods,

$\hat{Y}_t$ = the predicted level of the variable at time period t,

Y_t = the actual level of the variable at time period t.

August, 1978), and 108 (December, 1978).

The results presented in Table 7 also indicate that the best forecasting power within four ARIMA models is the AGG series model, which had only a 1.147 mean absolute percentage error in the postsample 24-month forecast period. The forecasting abilities of the TEG and GEG ARIMA models also were very good. The mean absolute percentage forecasting errors for TEG and GEG series models ranged from 1.345 to 3.658 and 2.299 to 3.165, respectively. The LWP series model had a poorer forecasting ability than the other models. For instance, the LWP 12-month mean absolute percentage error was 8.928, however, the other three series models ranged from 1.441 to 2.764.

In general, the accuracy of the ARIMA models decreased as one moved from 1-month to 12-month or longer forecasts. However, this is not true for AGG series model; the 1-month mean absolute percentage error for AGG model was 1.547, 12-month was 1.441, and 24-month was only 1.147. It appears that the AGG series model can provide both reliable short-term and long-term forecasting values. In conclusion, Tables 6 and 7 demonstrate the accurate forecasting ability of the AGG, TEG and GEG series models.

The One Input Transfer Function Model

Identification

During the first stage of fitting a transfer function model, the two residual series μ_{1t} and μ_{yt}, obtained from the model in Table 5, are cross correlated to obtain the estimated cross correlation function $r_{\hat{\mu}_1 \hat{\mu}_y}(K)$ (see Table 8). Significant individual lagged

Table 8 Estimated Univariate Residual Cross Correlation Function $r_{\hat{u}_1\hat{u}_y}(K)$

Mean of Prewhitened GEG series	= 4.550
ST. DEV. of Prewhitened GEG series	= 31.686
Mean of Prewhitened TEG series	= 7.839
ST. DEV. of Prewhitened TEG series	= 85.731

Lags (K)	$r_{\hat{u}_1\hat{u}_y}(K)$													
-20 - -7	-0.043	-0.011	0.184	-0.105	-0.220	-0.133	0.009	-0.114	0.123	0.178	0.037	-0.147	-0.158	0.135
-6 - 7	0.045	-0.251	0.028	-0.097	0.064	-0.055	0.388	-0.005	0.003	-0.063	0.061	-0.117	-0.038	0.057
8 - 20	0.115	0.025	0.126	0.056	-0.187	0.047	0.053	0.093	-0.036	-0.036	-0.020	0.068	-0.059	

correlations (judged against two standard deviation limits of $\pm 2N^{-1/2}$ or 0.204)[23] occur at lags 0 and -5. According to Pierce's causality relationship, one may conclude that TEG causes GEG and that one can build a dynamic regression model for GEG on TEG and thereby obtain more accurate forecasts of GEG.

Application of the Chi-square tests given in equations (17) and (18) (with M=12 and M=20) yields $S_{12} = 42.22$, $S_{12}^* = 44.739$, $S_{20} = 56.68$, $S_{20}^* = 62.106$. The appropriate critical values for 25 and 41 degrees of freedom are 37.65 and 56.8 ($\alpha = 0.05$). Hence, stochastic dependence is indicated (chiefly at lag 0, or "contemporaneous" correlation). Once the hypothesis of series indepencence is rejected, one may wish to use Table 6 further in identifying possible distributed lag models.

A pattern of the form in which $\rho_{\mu_1\mu_y}(K) \neq 0$ for $K = 0, -5$ leads one to consider the dynamic shock model

$$(24) \quad \mu_{1t} = \nu'(B)\mu_{yt} + \psi'(B)a_t \ ,$$

where $\nu'(B) = \omega_0' - \omega_5'B^5$ and $\psi'(B) = 1-\theta_5'B^5$.

Initial parameter values for $\omega'(B)$ and $\theta'(B)$ are obtained by solving a system of moment equations determined by $r_{\hat{\mu}_y}(K)$ and $r_{\hat{\mu}_1\hat{\mu}_y}(K)$[24] so that we find

$$\omega_0' \simeq 0.143, \ \omega_5' \simeq -0.09, \ \theta_5' \cong 0$$

---

[23] In order to make more exact judgments, Haugh (1976, pp. 379) suggested using $\pm 2(N-|K|)^{-1/2}$. where K is the number of lags.

[24] For more detail about the system of moment equations determined by $r_{\hat{\mu}_y}(K)$ and $r_{\hat{\mu}_1\hat{\mu}_y}(K)$, see Haugh (1972, pp. 106-117) and Haugh and Box (1974)

Using these initial values in (24) gives the identified dynamic shock model

$$\mu_{1t} = (0.143 + 0.09\ B^5)\ \mu_{yt} + a_t\ .$$

In order to complete identification of the dynamic regression model, we combine the identified univariate model from Table 5 and the joint shock model to get

$$(25)\quad (1-B)(1-B^{12})\ X_{1t} = \frac{(0.143-0.196B+0.395B^2-0.117B^5)\ (1-0.549B^{12})(1-B^{12})Y_t}{(1-0.44B)\ (1-0.3689B^{12})\ a_t,}$$

$$+\ (1-0.409B-0.195B^5)\ (1-0.549B^{12})\ a_t,$$

where X_{1t} = GEG, Y_t = TEG, and a_t is a white noise series.

Estimation-Diagnostic Checking-Reidentification

Final parameter estimates for model (25) are then obtained via a nonlinear least squares regression algorithm. The resulting model is

$$(26)\quad (1-B)(1-B^{12})X_{1t} = \frac{\underset{(0.046)\ (0.051)\ (0.029)\ (0.021)\ (0.373)}{(0.128-0.089B-0.057B^2+0.0159B^5)(1+1.48B^{12})(1-B^{12})Y_t}}{\underset{(0.214)\qquad (0.113)}{(1+0.544B)\ (1+1.07B^{12})}}$$

$$+\ \underset{(0.122)\qquad (0.129)\qquad (0.144)}{(1-0.48B-0.236B^5)(1-0.513B^{12})a_t}$$

where the numbers in parentheses associated with each estimated coefficient are the respective estimated standard errors of the coefficients.

The estimated standard deviation of the residuals from model (26) is 26.79, which is a 15.43 percent reduction from that of the univariate model for X_{1t}, i.e., GEG. The usual residual auto- and cross-correlation checks indicate a "good" data fit.

Pierce's causality relationship was used to develop model (26). However, Feige and Pearce (1979) pointed out that an essentially arbitrary choice can significantly affect the nature of the economic conclusions derived from the test procedures. Therefore, we tentatively reidentified the dynamic shock model to be

$$\mu_{yt} = \nu''(B)\mu_{1t} + \psi''(B)a_t$$

where $\nu''(B) = 2.706 + 0.316B^5$ and $\psi(B) = 1$. We then combined the identified univariate model from Table 5 and joint shock model to get the following alternative dynamic regression model:

$$(27) \quad (1-B^{12})Y_t = \frac{(2.706-1.201B+0.316B^5)(1-0.369B^{12})(1-B)(1-B^{12})X_t}{(1-1.377B+0.395B^2)(1-0.549B^{12})}$$

$$+ \frac{(1-0.444B)(1-0.3689B^{12})\ a_t}{(1-0.968B)}$$

Final parameters estimated for model (27) are then again obtained via a nonlinear least squares regression algorithm.

Our final alternative model is then

$$\text{(28)}\quad (1-B^{12})Y_t = \frac{\underset{(0.318)}{(0.876}-\underset{(0.607)}{0.354B}-\underset{(0.186)}{0.394B^5)}(1-\underset{(0.396)}{0.388B^{12}})(1-B)(1-B^{12})X_t}{(1-\underset{(0.525)}{1.227B} + \underset{(0.57)}{0.17B^2})\ (1 - \underset{(0.404)}{0.187B^{12}})}$$

$$+ \frac{(0-\overset{(0.156)}{0.256B})\ (1-\overset{(0.155)}{0.456B^{12}})\ a_t}{(1-\underset{(0.0609)}{0.925B})}$$

The estimated standard deviation of the residuals from model (28) is 68.89, which results in a 19.54 percent reduction in the standard deviation of the residuals from the univariate ARIMA model of TEG. The usual residual auto-and cross correlation checks also indicate a good data fit.

Comparing the results from model (26) and (28), we find the latter has a higher percent reduction in the standard deviation of the residuals. It seems that we should build a dynamic regression model for TEG on GEG instead of building a dynamic regression model for GEG on TEG. However, Pierce's causality relationship is not consistent with the empirical result in this study.

As the parameters $\hat{\omega}_{2,12} = 0.388$ and $\hat{\delta}_{2,12} = 0.187$ and as the model (28) is too complex, we may consider deleting $\omega_{2,12}$ and $\delta_{2,12}$ from our model. This new model fits as follows:

$$(29) \quad (1-B^{12})Y_t = \frac{\underset{(0.985-0.285B-0.165B^5)}{(0.2558)(0.2494)(0.2449)} (1-B)\ (1-B^{12})}{\underset{(0.045)\quad(0.095)}{(1-1.27B - 0.27B^2)}} X_t$$

$$+ \frac{\overset{(0.13)}{(1-0.277B)}\ \overset{(0.122)}{(1-0.601B^{12})}a_t}{\underset{(0.051)}{(1-0.957B)}}$$

The estimated standard deviation of the residual from model (29) is 65.08, which results in a 24.09 percent reduction from that of the univariate model for TEG. Therefore, model (29) is better than model (28), and we are more confident that one should build a dynamic regression model for TEG on GEG and thereby improve the forecasting ability of TEG.

Before accepting model (29) as an adequate representation of these employment data, autocorrelation and cross correlation checks should be applied, as described in Chapter III. The first 30 lags of the residual autocorrelations are tabulated in Table 9, together with their standard errors. There seems to be no evidence of model inadequacy from the behavior of individual autocorrelations. This is confirmed by calculating the Q_2 criterion which is

$$Q_2 = 77 \sum_{K=1}^{30} r^2_{\hat{a}\hat{a}}(K) = 14.63$$

Comparison of Q_2 with the χ^2 table value for $k-p-q = 30-1-2 = 27$ degrees of freedom provides no grounds for questioning model adequacy.

The estimated cross correlations between a_t and the prewhitened input μ_{1t} are given in Table 10. The criterion (equation (23) in Chapter III) yields

$$Q_3 = 77 \sum_{K=0}^{20} r^2_{\hat{\mu}\hat{a}}(K) = 14.982$$

or $$\left(Q_3^* = 77 \sum_{K=0}^{-20} r^2_{\hat{\mu}\hat{a}}(K) = 16.421\right) .$$

Comparison of Q_3 or Q_3* with the χ^2 table for K + 1 -(r+S+1) = 21-5 = 16 degrees of freedom again provides no evidence that the model is inadequate.

TABLE 9

Estimated Autocorrelation Function $r_{\hat{a}\hat{a}}(K)$ of Residuals from Model (29)

Lags K	$r_{\hat{a}\hat{a}}(K)$									
1-10	-0.01	-0.04	0.09	0.09	0.00	0.20	-0.02	-0.06	-0.02	0.00
St. E.	0.11	0.11	0.11	0.12	0.12	0.12	0.12	0.12	0.12	0.12
11-20	0.02	0.05	0.10	-0.12	0.00	-0.02	-0.04	-0.07	0.00	-0.22
St. E.	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
21-30	-0.02	-0.03	0.00	-0.05	-0.16	-0.05	0.09	0.05	-0.05	-0.01
St. E.	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13

Although $\hat{\omega}_{11}$ = 0.285 and $\hat{\omega}_{15}$ = 0.165 in model (29) do not appear to be greatly different from zero, if these two parameters are deleted.

Table 10. Estimated Cross Correlation Function $r_{\hat{\mu}_1\hat{a}}(K)$ Between the Prewhitened Input and Output Residuals.

Lags K	$r_{\hat{\mu}_1\hat{a}}(K)$									
-20 - -10	-0.150	-0.111	0.112	-0.165	-0.191	0.022	-0.001	-0.098	0.135	0.127	0.071
-9 - 0	-0.009	0.040	0.041	0.001	-0.146	0.120	0.008	-0.034	-0.111	0.009	
1 - 10	0.040	0.048	-0.006	0.066	0.008	-0.034	0.118	0.118	0.106	-0.026	
11-20	-0.027	-0.200	0.101	0.122	0.203	-0.027	-0.106	-0.108	0.112	-0.045	

from the model, the resulting residual auto-and cross correlation checks do not indicate a good data fit. Therefore, ω_{11} and ω_{15} are still considered in model (29).

In order to check whether X_{2t}, i.e., LWP, will be of some additional value in describing Y_t (conditional on X_{1t}), estimated cross correlations between a_t and μ_{2t} from Table 5 are presented in Table 11. It appears from Table 11 that the two series are not independent, significant lagged correlations occur at lags 0. Thus the inclusion of X_{2t} in the regression function of Y_t will now be considered.

Two-Input Transfer Function Model

Identification

At the first stage of fitting two-input transfer function model, we must extend our analysis of the various univariate shock series to a set of three series. $r_{\hat{\mu}_1\hat{\mu}_y}(K)$ has already been presented in Table 8. The estimated cross correlations between μ_1 and μ_2 and μ_2 and μ_y, i.e., $r_{\hat{\mu}_1\hat{\mu}_2}(K)$ and $r_{\hat{\mu}_2\hat{\mu}_y}(K)$, are presented in Table 12. Table 13 presents the results of chi-squared tests applied to test the independence of these pairs of residual series. From these tests, we find μ_{1t} and μ_{2t} are independent but μ_{2t} and μ_y are dependent. For this case, Haugh proved that the appropriate dynamic shock model for μ_t is:

$$\mu_t = \omega_{31}(B)\,\mu_{1t} + \omega_{32}(B)\,\mu_{2t} + \frac{\theta(B)}{\phi(B)}\,a_t \quad .$$

$\omega_{31}(B)$ and $\omega_{32}(B)$ can be identified directly from the patterns appearing respectively in $r_{\hat{\mu}_1\hat{\mu}_y}(K)$ and $r_{\hat{\mu}_2\hat{\mu}_y}(K)$. $\theta(B)$ and $\phi(B)$ also can be identified through Haugh's suggested process.

Table 11. Estimated Residual Cross Correlation Function $r_{\hat{\mu}_2\hat{a}}(K)$ (where $\hat{a}_t$ is the Residual Series of Model (29))[a/]

Lags K	$r_{\hat{\mu}_2\hat{a}}(K)$										
-20 - -10	-0.188	-0.112	-0.048	0.181	-0.013	0.141	0.124	0.088	-0.046	0.114	0.083
-9 - 0	0.072	0.010	0.005	0.162	-0.018	0.130	0.109	0.067	-0.048	0.494	
1 - 10	0.036	-0.086	-0.058	-0.045	0.070	0.067	-0.080	-0.087	0.039	-0.028	
11 - 20	-0.044	0.008	0.235	-0.155	-0.000	-0.021	-0.139	-0.037	0.034	-0.169	

a/ Application of the chi-square tests given in equation (17) and (18) with $M = 12$ and $M = 20$ yields $S = 36.022$, $S^*_{12} = 36.883$, $S_{20} = 60.619$, $S^*_{20} = 66.534$. The appropriate critical value for 25 and 41 degrees of freedom is 37.65 and 56.8 ($\alpha = 0.05$).

Table 12. Estimated Residual Cross Correlation Function $r_{\hat{\mu}_1\hat{\mu}_2}(K)$ and $r_{\hat{\mu}_2\hat{\mu}_y}(K)$.

Lags K	$r_{\hat{\mu}_1\hat{\mu}_2}(K)$										
-20 - -10	-0.057	-0.023	0.024	-0.155	-0.070	0.041	-0.081	-0.129	0.107	0.082	-0.004
-9 - -0	-0.140	0.050	0.124	-0.067	-0.100	0.074	-0.020	-0.135	-0.039	0.019	
1 - 10	-0.018	0.062	0.114	0.055	-0.015	0.190	-0.048	-0.036	0.025	0.053	
11 - 20	-0.085	-0.175	0.119	0.104	0.172	-0.018	0.048	0.029	-0.017	-0.182	

Lags K	$r_{\hat{\mu}_2\hat{\mu}_y}(K)$										
-20 - -10	-0.255	-0.095	-0.083	0.159	-0.045	0.105	0.135	0.153	-0.016	0.041	0.123
-9 - 0	0.063	0.096	0.030	0.212	-0.051	0.052	0.163	0.035	-0.059	0.457	
1 - 10	0.032	-0.171	-0.070	0.065	0.074	-0.019	-0.015	-0.100	-0.092	-0.096	
11 - 20	-0.013	0.007	0.057	-0.154	0.021	-0.058	-0.199	0.015	0.024	-0.174	

Table 13. Chi-squared Test Statistics for Lagged Cross-Correlations of Residual Series[a/] [b/]

Residual Pair	M = 12	M = 20
μ_1, μ_2	18.963 (20.550)	33.325 (37.830)
μ_2, μ_y	37.907 (39.032)	63.093 (69.672)

a/ Values are for $S_M = N \sum_{K=-M}^{M} r^2_{\hat{x}\hat{y}}(K)$ with 2M + 1 degrees of freedom

Values in parenthesis are $S^*_M = N^2 \sum_{K=-M}^{M} (N-|K|)^{-1} r^2_{\hat{x}\hat{y}}(K)$

with 2M + 1 degrees of freedom

where $X=\mu_1, \mu_2$, $Y=\mu_2, \mu_y$

b/ The appropriate critical value for 25 and 41 degrees of freedom is 37.65 and 56.8 ($\alpha = 0.05$)

Following Haugh's procedure, the dynamic shock model identified for μ_t is as follows:

(30) $$\mu_t = (1.05 + 0.013\,B)\,\mu_{1t} + (0.982 - 0.0688B + 0.367B^2 - 0.323B^3)\,\mu_{2t} + a_t$$

Substituting the fitted univariate models for each time series into the equation (30), the following dynamic regression model is tentatively obtained:

(31) $$(1-B^{12})\,Y_t = \frac{(1.05-0.453B)\,(1-B)\,(1-B^{12})\,X_{1t}}{(1-1.377B + 0.395B^2)} + \frac{(0.982-1.359B + 0.397B^2)\,(1-B^{12})\,X_{2t}}{(1-0.968B)} + \frac{(1-0.444B)\,(1-0.3689B^{12})}{(1-0.968B)}\,a_t\ ,$$

where Y_t = TEG, X_{1t} = GEG, X_{2t} = LWP, and a_t is a white noise series.

Estimation - Diagnostic Checking

Substituting the initial parameter estimates given in (31) into a nonlinear iterative estimation algorithm results in the following model:

(32) $$(1-B^{12})\,Y_t = \frac{\underset{(0.246)}{(1.127} - \underset{(0.248)}{0.537B)}\,(1-B)\,(1-B^{12})\,X_{1t}}{\underset{(0.0054)}{(1-1.389B} + \underset{(0.069)}{0.366B^2)}}$$

$$+ \frac{(\overset{(0.211)}{0.888}-\overset{(0.339)}{1.095}B + \overset{(0.221)}{0.095}B)(1-B^{12})}{(1-\underset{(0.0707)}{1.023}B)} X_{2t}$$

$$+ \frac{(1-\overset{(0.177)}{0.393}B)\ (1-\overset{(0.122)}{0.451}B^{12})}{(1-\underset{(0.120)}{0.823}B)} a_t \quad .$$

We note that $\hat{\omega}_{22} = -0.095$ does not appear to be greatly different from zero. $\hat{\phi}_1$ and $\hat{\theta}_1$ are highly correlated. The estimated standard deviation of the residuals from this model is disappointing (64.76). Therefore, the model was altered and reestimated to obtain:

(33)
$$(1-B^{12})\ Y_t = \frac{\overset{(0.202)}{0.873}\ (1-B)\ (1-B^{12})}{(1-\underset{(0.195)}{0.903}B + \underset{(0.217)}{0.124}B^2)} X_{1t}$$

$$+ \frac{(\overset{(0.153)}{0.762} - \overset{(0.153)}{0.944}B)(1-B^{12})}{(1-\underset{(0.003)}{1.044}B)} X_{2t}$$

$$+ \frac{(1-\overset{(0.093)}{0.804}B^{12})\ (1-B)}{(1+\underset{(0.087)}{0.543}B + \underset{(0.085)}{0.246}B^2)} a_t$$

with $\hat{\sigma}_a^2 = 56.96$.

Although $\hat{\delta}_{12} = -0.124$ in $\nu_1(B)$ do not appear to be greatly different from zero, we find δ_{11} and δ_{12} in $\nu_1(B)$ are highly correlated. Therefore, δ_{12} is still considered in model (33).

Table 14 gives the estimated residual autocorrelation and cross correlation functions for model (33). From this table, Q_2 and Q_3 are obtained,

$$Q_2 = 80 \sum_{K=1}^{30} r^2_{\hat{a}\hat{a}}(K) = 17.28$$

$$Q_3(\mu_1, a) = 80 \sum_{K=0}^{20} r^2_{\hat{\mu}_1\hat{a}}(K) = 12.526$$

$$(\text{or } Q_3^*(\mu_1, a) = 80 \sum_{K=0}^{-20} r^2_{\hat{\mu}_1\hat{a}}(K) = 14.711\ )$$

$$Q_3(\mu_2, a) = 80 \sum_{K=0}^{20} r^2_{\hat{\mu}_2\hat{a}}(K) = 9.264$$

$$(\text{or } Q_3^*(\mu_2, a) = 80 \sum_{K=0}^{-20} r^2_{\hat{\mu}_2\hat{a}}(K) = 20.171) \quad .$$

Comparison of Q_2, $Q_3(\mu_1,a)$ (or $Q_3^*(\mu_1,a)$), and $Q_3(\mu_2,a)$ (or $Q_3^*(\mu_2,a)$) with the individual χ^2 table values for 25, 18 and 18 degrees of freedom provides no evidence that the model is inadequate.

After these residual checks, it appears that model (33) is an adequate representation of the dynamic relationship between the three employment series. The estimated standard deviation of the residual series for TEG has again been reduced, from $\hat{\sigma}_{\mu_y} = 85.737$ (from Table 5)

Table 14. Estimated Residual Autocorrelation and Cross Correlation Function for Model (33)

Lags K	$r_{\hat{a}\hat{a}}(K)$									
1 - 10	0.12	-0.02	-0.13	-0.08	-0.06	0.03	0.03	-0.04	-0.12	-0.05
St. E.	0.11	0.11	0.11	0.12	0.12	0.12	0.12	0.12	0.12	0.12
11 - 20	-0.04	0.13	0.03	-0.07	-0.02	-0.06	-0.07	-0.11	0.10	-0.04
St. E.	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
21 - 30	0.09	0.11	0.06	-0.09	-0.12	0.09	0.19	0.04	-0.01	-0.07
St. E.	0.12	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13

Lags K	$r_{\hat{\mu}_1\hat{a}}(K)$										
-20 - -10	-0.059	-0.052	0.144	-0.078	-0.192	-0.064	-0.009	-0.029	0.090	0.066	0.024
-9 - 0	0.017	0.088	0.037	0.094	-0.230	-0.037	-0.068	0.086	-0.098	0.043	
1 - 10	-0.042	0.056	-0.038	0.071	-0.043	-0.082	-0.020	0.132	0.089	0.119	
11 - 20	0.017	-0.050	0.002	0.036	0.095	0.011	-0.110	-0.199	0.103	0.139	

Lags K	$r_{\hat{\mu}_2\hat{a}}(K)$										
-20 - -10	-0.181	-0.219	-0.209	0.180	-0.032	0.083	0.094	0.150	0.085	0.015	0.075
-9 - 0	0.036	-0.059	-0.133	0.049	-0.077	0.041	0.064	0.078	0.014	0.013	
1 - 10	-0.016	0.037	0.011	0.035	0.071	0.117	-0.022	-0.133	0.036	0.049	
11 - 20	0.014	-0.009	0.160	-0.046	0.003	0.042	-0.083	0.065	0.040	-0.172	

to $\hat{\sigma}_a$ = 65.08 (from model 29) and now to $\hat{\sigma}_a$ = 56.96 (from model 33).

This result, along with the 33.56 percent reduction in the standard deviation of the residual series from that of the univariate model for TEG, suggests that GEG and LWP are very helpful in explaining the time series behavior of TEG.

The estimated cross correlation function between a_t from model (33) and μ_{3t} from Table 5 are shown in Table 15, for checking whether X_{3t}, ie., AGG, will be of some additional value in describing Y_t (conditional on X_{1t} and X_{2t}). The results suggest that the two series are independent. (A three-input transfer function model was estimated on an experimental basis. The coefficient associated with X_{3t} was not significant and the standard deviation of the residual series was not reduced.) Thus, it is reasonable to conclude that AGG has little to add in explaining the time series behavior of TEG. Therefore, model (33) is the best representation of the Haugh-Box type dynamic relationship between these particular employment series in this study.

Econometric Analysis

In Chapter II, we have discussed three classes of regional econometric models which can be used to estimate employment multipliers: the simple static model, the simple dynamic model, and simultaneous equations model. In this study, only the simple static and dynamic model are employed.

The Simple Static Econometric Model

For this study Weiss and Gooding's modified economic base model for estimating differential employment multiplier is considered as the sim-

Table 15. Estimated Residual Cross Correlation Function $r_{\hat{\mu}_3\hat{a}}(K)$ (where $\hat{a}_t$ is the residual series of model (33).) a/

Lags K	$r_{\hat{\mu}_3\hat{a}}(K)$										
-20 - -10	-0.100	0.003	-0.050	0.103	-0.073	0.177	0.015	-0.012	0.143	0.025	-0.055
-9 - 0	0.045	-0.101	-0.048	0.027	0.272	-0.129	-0.024	0.192	-0.045	0.030	
1 - 10	-0.207	0.110	-0.054	-0.015	0.082	-0.017	0.037	0.063	0.011	0.071	
11 - 20	0.008	0.079	-0.087	0.106	0.078	0.002	0.014	0.040	-0.080	0.043	

a/ Application of the Chi-square tests given in equation (17) and (18) (with M = 12 and M = 20) yields $S_{12} = 24.216$, $S^*_{12} = 25.512$, $S_{20} = 33.344$, $S^*_{20} = 36.477$. The appropriate critical value for 25 and 41 degrees of freedom is 37.65 and 56.8 ($\alpha = 0.05$)

ple static econometric model.

$$Y_t = b_o + b_1 X_{1t} + b_2 X_{2t} + b_3 X_{3t}$$

where Y_t = TEG, X_{1t} = GEG, X_{2t} = LWP, and X_{3t} = AGG were regarded as distinct export sectors in Grant County.

Two principal assumptions underlie this model. These assumptions are (1) that the export sectors are independent of one another, and (2) that workers and firms within the same economic base sector have similar consumption patterns (although consumption may be different across export sectors). The first assumption seems highly plausible for the basic sectors considered in this study; however, the second assumption can only be tested with more detailed survey data.

Based on the same data as employed in fitting the transfer function model, the final equation for the static econometric model is:

$$(34) \quad Y_t = \underset{(107.807)}{569.9} + \underset{(0.146)}{1.29} X_{1t} + \underset{(0.100)}{1.129} X_{2t} + \underset{(0.182)}{1.336} X_{3t},$$

R^2 = 0.9584, D-W = 2.129, F(3,91) = 698.675,

Standard error of regression (SE) = 56.98,

where the numbers in parenthesis below the estimated coefficients are estimated standard errors. From these, we find all the coefficients are significantly different from zero. Looking at the R^2, we see the equation has excellent explanatory power. Furthermore, the Durbin-Watson statistic suggests that serial correlation is not a problem in this case.

The Simple Dynamic Econometric Model

The dynamic econometric models to be employed empirically in the subsequent analysis may be written as:

$$ET_t = b_o + \sum_{i=1}^{r} b_i^* ET_{t-i} + \sum_{j=1}^{m} \sum_{i=0}^{n} b_i^j EB_{t-1}^j ,$$

where ET_{t-i} is the endogenous variable (total employment) at time t-i, and

EB_{t-i}^j is the jth exogenous variable (basic employment in sector j) at time t-i.

For this dynamic employment model, if no lagged endogenous variables are included, the equation is estimated by the Almon distributed lag technique with a fourth-degree polynomial and no end-point restrictions.[25] If the model includes lagged endogenous variables, then the Koyck procedure is used to fit the model.[26] If the model includes both lagged exogenous variables and lagged endogenous variables, the Almon lag technique is used, however, in this case the estimate of the parameters will be biased and inconsistent.

Following four basic criteria,[27] three of ten dynamic models were

---

25 For a detailed explanation about Almon distributed lag technique, see Koutsoyiannis (1972, pp. 299-304) and Johnston (1972, pp. 292-300.)

26 For a detailed explanation about Koyck geometric lag scheme, see Koutsoyiannis (1972, pp. 304-310) and Johnston (1972, pp. 300-303).

27 Four basic criteria for selecting the final model are (1) low standard error of the estimates, (2) higher adjusted R^2 and better F test, (3) significance of most coefficients at the 95 percent confidence level, and (4) better forecasting ability.

chosen for final consideration, the first being:

$$(35)\quad Y_t = \underset{(131.219)}{591.528} - \underset{(0.050)}{0.0148}\, Y_{t-1} + \underset{(0.154)}{1.308}\, X_{1t} + \underset{(0.116)}{1.144}\, X_{2t} + \underset{(0.185)}{1.33}\, X_{3t},$$

$$R^2 = 0.958,\ \text{D-W} = 2.126,\ F(4,90) = 518.74,\ SE = 60.91.$$

For model (35) most of coefficients are significant at 0.01 level, except that for Y_{t-1} which does not appear to be greatly different from zero. Other diagnostic checks show that the equation has excellent explanatory power. An alternative model which omits Y_{t-1} and includes more lags on the exogenous variable is:

$$(36)\quad Y_t = \underset{(142.288)}{512.61} + \underset{(0.176)}{1.286}\, X_{1t} + \underset{(0.178)}{0.144}\, X_{1t-1} - \underset{(0.191)}{0.054}\, X_{1t-2}$$

$$+ \underset{(0.120)}{1.089}\, X_{2t} - \underset{(0.113)}{0.015}\, X_{2t-1} + \underset{(0.224)}{1.226}\, X_{3t} + \underset{(0.241)}{0.117}\, X_{3t-1} + \underset{(0.211)}{0.042}\, X_{3t-2},$$

$$R^2 = 0.964,\ \text{D-W} = 1.9516,\ F(8,83) = 277.826,\ SE = 56.98.$$

For this model the coefficients for lagged variables are not all significantly different from zero; however, other diagnostic checks suggest that the model is very good in terms of fit.

The next model is a combination of models (35) and (36) with the following results:

$$(37)\quad Y_t = \underset{(109.67)}{147.2} + \underset{(0.068)}{0.785}\, Y_{t-1} + \underset{(0.190)}{1.335}\, X_{1t} - \underset{(0.283)}{0.834}\, X_{1t-1} - \underset{(0.214)}{0.212}\, X_{1t-2}$$

$$+ \underset{(0.126)}{1.047}\, X_{2t} - \underset{(0.191)}{0.83}\, X_{2t-1} - \underset{(0.135)}{0.06}\, X_{2t-2} + \underset{(0.301)}{1.16}\, X_{3t} - \underset{(0.374)}{0.99}\, X_{3t-1}$$

$$+ \underset{(0.235)}{0.178}\, X_{3t-2} \;,$$

$$R^2 = 0.966,\ \text{D-W} = 1.798,\ F(10,81) = 232.25,\ SE = 56.03 \;.$$

For this model, most of coefficients are significant at 0.05 level, except those for X_{1t-2}, X_{2t-2}, and X_{3t-2}.

Forecasting Accuracy Check

Six tests which were used in testing ARIMA forecasting ability are used again as a basis for the evaluation of the forecasting power of these four econometric models.

The results of the measuring of forecasting accuracy for these four models from January 1978 to December 1979 are presented in Table 16.

Generally, the forecasting ability for model 36 (i.e., only using current and lagged explanatory variables) is best. The Theil's inequality coefficient was 0.0102, the root mean square error was 64.62, and R and β_1 approximated 1 at the same time. However, all the lagged coefficients are not significantly different from zero for this model. If we delete all the lagged variables from this model, we get model 34. The forecasting ability for model 34 is also very good, i.e., the Theil's inequality coefficient was 0.0103 and the root mean square error was

Table 16. Comparison of Forecasting Accuracy Among Four Econometric Models[a/]

Forecast Model	R	β_1	ME	MAE	RMSE	U	(I	II	III)
34	0.9908	0.9522	58.87	59.27	65.64	0.0103	(0.8044	0.0154	0.1801)
35	0.9876	0.9525	56.14	58.95	65.17	0.0103	(0.7422	0.013	0.2448)
36	0.9892	0.9101	54.00	56.68	64.62	0.0102	(0.6982	0.0732	0.2286)
37	0.8925	0.7213	7.605	50.11	114.5	0.0179	(0.0044	0.1742	0.8214)

[a/] See Table 6.

65.64. If we compare the forecasting ability and diagnostic checks of the models between model 34 and 36, it then seems that model 34 is better. Model 35 also had the similar situation as model 36. And model 37 was attempted and yielded very good R^2 values, standard error of regression, and F statistics; however, the forecasting ability was not as good as for the other three models.

The Forecasting Accuracy Comparison of ARIMA, Transfer Function, and Econometric Models

In order to evaluate overall relative forecasting ability, one-month to twenty-four month forecasts were compared for the above four models, i.e., the final TEG ARIMA model (from Table 5), the one-input transfer function model (equation 29), the two-input transfer function model (equation 33), and the static econometric model (equation 34). The forecast percentage errors using forecast origin December, 1977, are presented in Table 17. Note that the econometric model had the actual data for the independent variable over the forecast period, giving it an advantage over the ARIMA and transfer function model.[28]

Looking at the one- and two-month ahead forecasts, the econometric model performed appreciably better than any other. However, it had the worst forecasting performance for the three- and four-month ahead forecasts, though the model used the actual data for the independent variable. For the thirteen-, fourteen-, and fifteen-month ahead forecasts, the one-input transfer function and ARIMA model had more accurate forecasts than the other two models.

---

[28]Due to the limitation of the computer package, the forecasted values of the independent variables were used to forecast the output series for the transfer function model.

Table 17. Forecast Percentage Errors Among Four Models[a/]

Periods Ahead	Econometric	ARIMA	One-input transfer function	Two-input transfer function
1	0.56	3.76	3.29	3.76
2	1.66	6.27	5.82	6.37
3	3.37	0.63	0.21	0.63
4	4.03	2.10	1.61	1.97
5	3.64	2.42	1.96	2.82
6	1.65	-3.27	-2.99	-3.30
7	1.89	-6.55	-5.44	-5.08
8	2.08	-2.41	-2.57	-3.34
9	2.18	-2.18	-2.45	-3.70
10	2.39	-2.18	-2.13	-3.55
11	1.92	-2.40	-2.57	-4.90
12	2.75	-2.70	-2.62	-4.73
13	0.77	0.25	-0.20	-3.14
14	1.15	0.27	-0.38	-3.97
15	1.65	1.54	0.67	-3.50
16	1.66	-1.39	-2.65	-7.65
17	2.86	-1.41	-2.59	-7.37
18	1.15	-5.88	-6.33	-12.36
19	0.69	-7.77	-7.42	-13.06
20	-0.14	-6.98	-7.95	-15.51
21	1.29	-4.82	-5.93	-14.44
22	2.26	-4.85	-5.71	-15.27
23	1.64	-5.45	-6.63	-18.10
24	1.75	-5.44	-6.44	-18.59

a/ Forcast Percentage Error = $\frac{Y_t - \hat{Y}_t}{Y_t} \cdot 100\%$.

For the most part, the forecasting errors for the econometric model were for an overestimation situation, except the twenty-month ahead forecast. The forecasting errors for the ARIMA and transfer function model were underestimations, except in the short-term forecasts. For short lead times, the one-input transfer function performed better than the ARIMA and two-input transfer function model. This is probably because the one-input transfer function model has explanatory capabilities and the forecasting performance of the input series (GEG) is very good. This is not true, however, for the two-input transfer function model because the forecasting performance of the LWP series is not very good. Thus, additional variation was introduced into the model due to errors in the forecasts of the second input (LWP). Forecasts from the two-input transfer function model are worse than for any other model. As lead time increases, the drop in accuracy of the forecasts for the GEG series reduced the forecasting ability of one-input transfer function. Therefore, the ARIMA model performed better than the one-input transfer function and appreciately better than the two-input transfer function model for "long-term" forecasts.

In summary, there is no guarantee that the forecasts the econometric model generates will be superior to forecasts given by an ARIMA model which is based only upon the past history of the variable, since additional variation will be introduced into the econometric model due to errors in the forecasts of the independent variables, especially in the future. Our tests were conducted under artificial laboroatory conditions favoring the econometric model because the independent variables were all "perfectly" forecasted on the future. Of course this is not possible in real life applications.

In making short-term or even longer-term forecasts the transfer function model should have the edge over the econometric model because X_t values (independent variable) can be forecasted using the univariate model generating X_t. However, we cannot guarantee that the transfer function model will generate better forecasts than the ARIMA model. This is because the ARIMA model for X_t may not produce accurate forecasts for the X_t, although X_t may be a good leading indicator of Y_t.

Therefore, in order to minimize the forecast errors, predictions from both the transfer function and ARIMA models (or econometric model) should be combined to obtain composite forecasts. For example, by regressing the actual use levels on the two forms of predictions, we obtain estimates of composite weights in terms of the regression coefficients. The estimated composite weights are then employed to obtain linear composite predictions.[29]

---

[29] For more detail on composite forecasts, see Nelson (1973, pp. 212-214), Newbold and Granger (1977, pp. 268-277), Rausser and Oliveira (1976, pp. 282-284), and Oliveira (1978, pp. 524-527).

CHAPTER V

EMPLOYMENT IMPACTS

The transfer function and econometric models presented in the last chapter are used in examining the interrelationships among total, government, lumber and wood products manufacturing, and agriculture employment in this chapter. These two approaches are based on quite different modeling philosophies as has been discussed in Chapter III; thus, the different results coming from these two models are not surprising.

Economists have long been interested in predicting the influence of economic phenomena which may appear in the future. The intent of this chapter is to measure the total employment effect of hypothetical changes in basic employment. Changes in total employment will be measured in terms of dynamic employment response and employment multipliers. Finally, the empirical results are discussed, and various reasons for different employment multipliers are proposed.

Model Behavior - Dynamic Employment Response and Employment Multiplier

The purpose of this section is to utilize the transfer function and econometric models to estimate the impacts from changes in the different export employment categories on changes in total employment.

First, the final transfer function, i.e., equation (33) in Chapter IV, is employed to estimate the dynamic responses, i.e., the responses over time, of total employment from changes in government or lumber and wood products manufacturing employment. Examples of the estimated effects on estimated total employment of hypothetical changes in other employment are shown in Tables 18 and 19. These are the consequences of:

1) an increase of 10 percent in government employment from January 1976 to December 1977,

2) an increase of 74 government employees during the same period,

3) an increase of 10 percent in lumber and wood products manufacturing employment during the same period, and

4) an increase of 60 lumber and wood products manufacturing employees during the same period.

The estimated total employment gains or losses due to these changes were disappointing. The initial increase of 74 persons in government employment generated only 67 persons in the total employment (see Table 18. This result seems quite different from general impact analysis. The results from the second month are even more frustrating; the increase of 74 or 75 persons in government employment generated only 31 or 32 persons in the total employment. That means the induced employment effects from the first month are negative. All of these results are quite the opposite of what one would expect, given the hypothesized relationship. Similar results also were obtained from changes in lumber and wood products manufacturing employment (see Table 19). In the first month, the increase of 60 persons in lumber and wood products manufacturing employment generated only 44 persons in the total employment.

TABLE 18. Impacts of Changes in the Government Employment on the Estimated Total Employment in Grant County from January, 1976 to December, 1977

Month	$\hat{Y}$ a/	$\hat{Y}_1$ b/	$\Delta\hat{Y}_1$ c/	ΔX_1	$\hat{Y}_1$ b/	$\Delta\hat{Y}_1$ c/	ΔX_1
1	2615	2682	67	74	2682	67	74
2	2554	2585	31	73	2586	32	74
3	2549	2549	0	79	2548	-1	74
4	2550	2531	-19	83	2528	-22	74
5	2682	2665	-17	86	2662	-20	74
6	3153	3141	-12	95	3136	-17	74
7	3285	3277	-8	101	3272	-13	74
8	3240	3233	-7	97	3234	-6	74
9	3270	3262	-8	102	3264	-6	74
10	3138	3132	-6	95	3139	1	74
11	3003	3002	-1	86	3009	6	74
12	2902	2906	4	82	2911	9	74
13	2797	2788	-9	80	2790	-7	74
14	2747	2747	0	79	2746	-1	74
15	2654	2660	6	80	2657	3	74
16	2692	2698	6	86	2692	0	74
17	2866	2870	4	92	2864	-2	74
18	3341	3343	2	102	3337	-4	74
19	3454	3459	5	100	3456	2	74
20	3507	3505	-2	102	3504	-3	74
21	3477	3475	-2	107	3474	-3	74
22	3327	3320	-7	103	3323	-4	74
23	3225	3221	-4	99	3224	-1	74
24	3090	3091	1	94	3093	3	74

a/ Estimate total employment based upon the actual government and lumber and wood products manufacturing employment.

b/ Estimate total employment based upon the hypothetical level of government employment and actual lumber and wood products manufacturing employment.

c/ $\Delta\hat{Y}_1 = \hat{Y}_1 - \hat{Y}$

TABLE 19: Impacts of Changes in the Lumber and Wood Products Manufacturing Employment on the Estimated Total Employment in Grant County from January, 1976 to December, 1977.

Month	$\hat{Y}$ a/	$\hat{Y}_2$ b/	$\Delta\hat{Y}_2$ c/	ΔX_2	$\hat{Y}_2$ b/	$\Delta\hat{Y}_2$ c/	ΔX_2
1	2615	2659	44	60	2659	44	60
2	2554	2571	17	57	2577	23	60
3	2549	2549	0	50	2549	0	60
4	2550	2543	-7	41	2546	-4	60
5	2682	2680	-2	48	2675	-7	60
6	3153	3162	9	76	3142	-11	60
7	3285	3290	5	77	3281	-4	60
8	3240	3233	-7	77	3232	-8	60
9	3270	3257	-13	77	3262	-8	60
10	3138	3127	-11	78	3128	-10	60
11	3003	2990	-13	78	2988	-15	60
12	2902	2890	-12	78	2888	-14	60
13	2797	2782	-15	73	2780	-17	60
14	2747	2738	-9	67	2737	-10	60
15	2654	2639	-15	53	2648	-6	60
16	2692	2680	-12	44	2689	-3	60
17	2866	2860	-6	50	2861	-5	60
18	3341	3345	4	75	3334	-7	60
19	3454	3455	1	85	3444	-10	60
20	3507	3499	-8	86	3499	-8	60
21	3477	3469	-8	86	3470	-7	60
22	3327	3318	-9	85	3321	-6	60
23	3225	3215	-10	84	3220	-5	60
24	3090	3078	-12	83	3083	-7	60

a/ Estimate total employment based upon the actual government and lumber and wood products manufacturing employment.

b/ Estimate total employment based upon the actual government employment and the hypothetical level of lumber and wood products manufacturing employment.

c/ $\Delta\hat{Y}_2 = \hat{Y}_2 - \hat{Y}$

In the second month, the increase of 57 or 60 persons in lumber and wood products manufacturing employment generated on 17 or 23 persons in the total employment.

Cook (1979) obtained similar results in his study. He concluded

> "This result indicated that an increase in geographic oriented employment can partially be attributed to a transfer of workers from the other employment sector to the geographic oriented employment sector. To draw this conclusion, it must be postulated that job opportunities in the geographic oriented industries meet the qualifications of some workers in the other industries. In addition, there must be incentives for a worker to transfer from one sector to the other sector...This tends to support the postulate that there may be incentives for workers to transfer to geographic oriented industries because of positive wage and salary differences." (15, p. 90).

In order to explain the incentives for workers to transfer jobs, the average payroll per employee for different sectors was examined (see Table 20). Lumber and wood products manufacturing was found to have the highest average payroll per employee; transportation, communications, electric, gas and sanitary services was second, government was third. Thus, the average payroll data tends to support Cook's postulate. The negative induced employment effect, however, still cannot be explained. What is even more frustrating about the results in Tables 18 and 19 is that it appears as if increases in export oriented employment do not increase the total employment opportunities in Grant County. This seems to imply a negative multiplier effect.

In order to estimate employment causality relationships from econometric models, two approaches were employed. The first was checking the estimated total employment gains or losses due to hypothetical changes in other employment. Several examples illustrating the effects on Grant County's total employment of hypothetical changes in other employment are shown in Table 21. These hypothetical changes are (1)

TABLE 20. Grant County Covered Employment, Payrolls, and Average Payroll per Employee by Industry, First Quarter, 1977.

Industry Description	Covered Employment	Payrolls	Average Payroll per Employee
Contract Construction	69	42,921	622.04
Lumber and Wood Products Manufacturing	1,859	2,237,771	1,203.75
Transportation, Communications, Electric, Gas, and Sanitary Services	207	209,685	1,012.97
Wholesale Trade	59	44,190	748.98
Retail Trade	864	476,745	551.78
Finance, Insurance, and Real Estate	141	87,036	617.27
Services	385	117,539	305.29
Government	2,200	1,777,516	807.96

Source: Oregon Employment Division, Research and Statistics Section. Oregon Covered Employment and Payrolls by Industry and County First Quarter 1977, Salem, pp. 47-48.

an increase of 89 (i.e., 10 percent) government employees in January, 1978, (2) an increase of 80 (i.e., 10 percent) lumber and wood products manufacturing employees in January, 1978 and (3) an increase of 32 (i.e., 10 percent) agriculture employees at January, 1978.

Because the effects of an initial employment change tend to multiply throughout an economy, the total employment impact on county employment may be shown as a multiplier for a particular sector. According to different econometric models, different sets of multipliers were calculated and are also shown in Table 21.

The pattern of employment impacts varies with the type of econometric model. For example, for the simple static econometric model, model 34, the initial increase of 89 government employees generated 115 additional total employees at the same period, but no lag effect happened. Thus, the government employment multiplier was 1.29. The increase of 80 lumber and wood products manufacturing employees and 32 agriculture employees resulted in 90 or 42 additional total employees, respectively. The results indicate that a change in basic employment will immediately and positively impact non-basic employment. A similar result occurred for model 35. However, the impacts on Grant County's total employment from hypothetical changes in basic employment were greater in model 35 than for model 34. Although model 35 is a dynamic model, no lag effect happened, since the coefficient of the lag variable in model 35 was not significant.

For model 36, the initial increase of 89 government employees generated 115 total employees in the first month (zero month lag), generated 13 total employees in the second month (one month lag),

TABLE 21 Impacts and Multipliers of Selected Changes on the Grant County's Total Employment by Econometric Model.

		Alternative and Impact [a/]					
		Government		Lumber and Wood		Agriculture	
Model	Month	Number of Changes	Total Changes	Number of Changes	Total Changes	Number of Changes	Total Changes
34	1	115	115	90	90	42	42
	2	0	115	0	90	0	42
		M_g [b/] = $115/_{89}$ = 1.29		M_m [c/] = $90/_{80}$ = 1.125		M_a [d/] = $42/_{32}$ = 1.312	
35	1	117	117	92	92	43	43
	2	0	117	0	92	0	43
		M_g = $117/_{89}$ = 1.314		M_m = $92/_{80}$ = 1.15		M_a = $43/_{32}$ = 1.34	
36	1	115	115	87	87	40	40
	2	13	128	-1	86	4	44
	3	-5	123	0	86	1	45
	4	0	123	0	86	0	45
		M_g = $123/_{89}$ = 1.382		M_m = $86/_{80}$ = 1.075		M_a = $45/_{32}$ = 1.406	
37	1	118	118	83	83	37	37
	2	-74	44	-64	19	-32	5
	3	-19	25	-4	15	5	10
	4	0	25	0	15	0	10
		M_g = $25/_{89}$ = 0.28028		M_m = $15/_{80}$ = 0.1875		M_a = $10/_{32}$ = 0.3125	

a/ The initial or direct changes under each alternative are (1) an increase of 89 government employees, (2) an increase of 80 lumber and wood products manufacturing employees, and (3) an increase of 32 agriculture employees.

b/ M_g is the multiplier for government employment.

c/ M_m is the multiplier for lumber and wood products manufacturing employment.

d/ M_a is the multiplier for agriculture employment.

decreased 5 total employees in the third month (two months lag). Therefore, the net effect in total employment was 123 employees generated. This can be measured by summing the number of changes in total employment at zero, one, and two month lags. The government employment multiplier from this model was thus 1.382. The negative effect occurring in the third month is quite opposite of what one would expect, given the hypothesized relationship. It is hypothesized that because of the seasonal variation within the employment data, the decrease of five employees in the third month were due to the seasonal variation, not negative induced employment effect.

The results from model 37 were disappointing. The initial increase of 89 government employees generated 118 total employees in the first month. However, the unexpected negative impacts occurred in the second and third months. Therefore, the net effect in total employment was only 25 employees generated. For this result, the seasonal variation within the employment data, and biased and inconsistent estimates for model 37 are the reasons why the negative impact occurs.

The multipliers for each sector also vary with the types of econometric model. In general, most of the multipliers that come from model 36 are higher than others, except the multiplier for lumber and wood products manufacturing employment.

The second method to estimate the employment multipliers from econometric model is to employ Theil and Boot's dynamic impact multiplier approach (this was previously discussed in Chapter II). Suppose the econometric model is

$$ET_t = b_o + \sum_{i=1}^{r} b_i^* \, ET_{t-i} + \sum_{j=1}^{m} \sum_{i=0}^{n} b_i^j \, EB_{t-i}^j ,$$

where:

ET_{t-i} is the total employment at time t-i

EB_{t-i}^j is the jth basic employment at time t-i.

The dynamic differential employment multipliers for this model are defined as K_j, which is the sum of elements in the first row of $(I-A)^{-1} * B^j$,[30]

where:

K_j is the dynamic employment multiplier for jth basic sector,

I is the identity matrix,

$$A = \begin{bmatrix} b_1^* & b_2^* & b_3^* & - - - & b_{r-1}^* & b_r^* \\ 1 & 0 & 0 & - - - & 0 & 0 \\ 0 & 1 & 0 & - - - & 0 & 0 \\ 0 & 0 & 1 & - - - & 0 & 0 \\ - & - & - & - - - & - & - \\ 0 & 0 & 0 & - - - & 0 & 1 \end{bmatrix} , \text{ and}$$

[30] For a mathematic proof and more detailed discussion, see Stewart and Venieris (1978, pp. 459-462) and Theil and Boot (1962, pp. 136-152).

$$
B^j = \begin{bmatrix} b_o^j & b_1^j & b_2^j & ----- & b_n^j \\ 0 & 0 & 0 & ----- & 0 \\ 0 & 0 & 0 & ----- & 0 \\ & - - - - & - - - - & - - - - & \\ & - - - - & - - - - & - - - - & \\ 0 & 0 & 0 & ----- & 0 \end{bmatrix} .
$$

The estimated Theil-Boot employment multipliers for the three basic sectors of four models are shown in Table 22. The range of differential multipliers derived via this method seems quite consistent with estimates obtained from the first method, except for the multipliers of model 37. As we discussed above, the results from model 37 are unreliable due to their biased and inconsistant coefficient estimates.

Interpretation of Empirical Results

The transfer function for Grant County economic-base employment model has specified a dynamic relationship as hypothesized. However, it has been found that the dynamic responses of total employment from changes in government or lumber and wood products manufacturing employment were disappointing. Therefore, in this section, only the results from the econometric models are interpreted.

We have employed two approaches to estimate the employment multipliers for three basic sectors of four econometric models. The results from different approaches were very similar, except for the multipliers of model 37. It has also been found that the employment

TABLE 22. Estimated Employment Multipliers for Basic Sectors of Four Econometric Models

Model	Employment Multiplier		
	Government	Lumber and Wood Products Manufacturing	Agriculture
34	1.29	1.129	1.336
35	1.2889	1.127	1.3106
36	1.3757	1.0739	1.3858
37	1.3418	0.7089	1.5911

multipliers have the same descending order for M_a, M_g, and M_m; that is, $M_a \geq M_g \geq M_m$. These results are different than income multipliers that come from other studies. For example, Miller and Obermiller found that the city/county government had the highest income multiplier of 2.79. They also found the income multipliers for local state/federal agencies were 1.94, lumber and wood products processing were 2.55, dependent ranching (agriculture) was 2.39, other ranching (agriculture) was 2.36, and general agriculture was 1.85. Therefore, the income multiplier of agriculture is smaller than the income multiplier of government and lumber and wood products manufacturing. The main reason is because of the different average annual wages among different sectors.

We do not have good methods to calculate the income multiplier directly from the employment multiplier. However, the employment multiplier analysis could provide some basis for making estimates of changes in income levels which are likely to follow changes in basic employment. Consider, for example, Grant County with an average monthly income per employed person of \$800 experiencing an increase in government employment of 100 persons. Then a government employment multiplier of 1.3 would suggest that the initial increase in government employment would raise total monthly income by \$104,000 [(1.3)(100)(\$800) = \$104,000].

Since the city/county government employees and agriculture employees are mostly civilians, it is reasonable to expect that they would have a greater local impact than other sectors. Moreover, given that state/federal employees receive higher average annual wages than other workers in the county,[31] they might have higher propensities to import and to spend money outside the area or higher propensities to save. Due to these reasons, the government employment multiplier is smaller than the agriculture employment multiplier. Lumber and wood products manufacturing's material inputs and machines are acquired from suppliers outside the county, which causes a smaller multiplier.

Summary

In this chapter, we have explored various employment impacts of different sectors from transfer function and econometric models. The results from the transfer function model were frustrating, although in Chapter IV we found the forecasting ability of this model was good. The results from the econometric models were presented and explained. In general, the employment multipliers of the dynamic econometric models were higher than those of the static econometric model. Finally, we discussed and interpreted the employment impacts of government, lumber and wood products manufacturing, and agriculture sectors.

In the next chapter, the study is summarized and conclusions are drawn. The models are evaluated and the needs for future research outlined.

---

[31] Federal/State government employees have an average monthly income per employed person of $1,060 at first quarter, 1977 (Oregon Employment Division, Research and Statistics Section. Oregon Covered Employment and Payrolls by Industry and County First Quarter, 1977, Salem, p. 48).

CHAPTER VI

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary and Conclusions

The major goals of this study have focused around (a) demonstrating and applying the transfer function and testing its ability in regional economic modeling and regional impact analysis, and (b) obtaining a better understanding of the causal relationships among total employment, government employment, lumber and wood products manufacturing employment, and agriculture employment in Grant County.

The research study started with the investigation of the characteristics and problems of three types of regional forecasting models. These three approaches have different theoretical implications, causing individual advantages and disadvantages. For instance, the input-output approach can introduce a higher degree of sophistication and comprehensiveness than the economic base approach. However, the construction of a complete interindustry model usually entails prohibitive costs and more cumbersome data problems. Regional econometric models are in some respects a compromise between economic base and input-output formulations. With respect to data costs, they do not require as much data collection as for input-output models. They are comparable in cost to economic base models, but provide relatively more information about the structure of a region's economy. Unfortunately, theoretical and data problems make regional econometric models difficult to build for small regions. The significant deficiencies in these three approaches and the implicit theoretical causal relationships between basic and total employment, provided the rationale for employing transfer function models in regional economic analysis.

The conceptual theory and the modeling procedures of the autoregressive-integrated-moving-average (ARIMA) and the transfer function have been proposed in Chapter III. The procedures for the development of a "transfer function model", a model which expresses the interrelationships among time series, involve the iterative repetition of identification, estimation, and diagnostic checking stages. These steps are comparable to the iterative procedures for the development of an ARIMA model for a single time series. A method of identifying a transfer function model has been thoroughly elaborated by Box and Jenkins (1970, 1976) and Haugh (1972). The first step is to model the univariate series and obtain their residuals (prewhitened series). Then the sample cross-correlation functions between the various pairs of residual series are studied in an attempt to identify the relationships among the series (in this study, Pierce's causality approach was then used to test the relationships among employment variables). The second step is to employ the auto- and cross-correlation patterns to identify a dynamic shock model for the prewhitened series. The final step of the identification phase is to substitute the identified univariate ARIMA models into the identified dynamic shock model giving a preliminary transfer function model.

The parameters of the tentatively identified model are estimated and various residual checks are performed. If these checks show inadequacies in the model, a new model is entertained, estimated, and checked. When the residual checks are satisfied, the modeling procedures stop.

Dynamic modeling via a transfer function model has been proven to be a useful tool for economic analysis and short-term forecasting. However, the transfer function modeling procedures are still at a relatively early point in development. Obviously, further efforts in this area are needed.

In Chapter IV the ARIMA and transfer function models were empirically implemented. In order to compare the results from these models, one static and three dynamic econometric models also were specified and their parameters were estimated by using ordinary least squares, Almon's distributed lag, and Koyck's distributed lag estimation procedures. After all the models were used as forecasting equations, the forecasting performances of each model were evaluated using six types of goodness-of-fit measures, i.e., correlation coefficient, root mean square error, mean absolute error, mean error, regression coefficient of actual on prediction, and Theil's inequality coefficient. The findings suggested that the one-input transfer function demonstrated superiority over the ARIMA and two-input transfer function models and similarly for the econometric models. The econometric model forecasts, however, employed the actual data for the independent variable over the forecast period, giving them an "advantage" over the ARIMA and transfer function models.

In summary, there is no guarantee that the forecasts generated by the transfer function model will be superior to forecasts given by an ARIMA model. Additional variation will be introduced into the transfer function model due to errors in the forecasts of independent variables, especially in the future. However, in making short-term or even longer-term forecasts, transfer function models should have the edge over econometric models because X_t values (independent variables) can be forecasted using the univariate model generating each X_t. Of course, this does not mean that econometric models are to be replaced substantially or completely. Transfer function and ARIMA models have some inherent shortcomings compared to structural types of models such as econometric models in analyzing economic time series. For instance, transfer function models do not emphasize the simultaneity of the relationships between economic variables as can econometric models. ARIMA models are generally void of explanatory power, and they are not based on economic theory. If these models yield poor forecasts, one would be at a complete loss to explain the reason (Naylor, Seaks, and Wichern, 1972). Therefore, in order to minimize the forecast error, predictions from both the transfer function and ARIMA models (or econometric model) should be combined to form composite forecasts.

The two-input transfer function and econometric models were subsequently used to measure the employment multipliers for government, lumber and wood products manufacturing and agriculture sectors in Chapter V. The results obtained from the transfer function model were frustrating. The employment multipliers, in terms of the dynamic employment response, were less than one (or negative) for government and lumber and wood products manufacturing sectors. The reasons for these findings are not clear; however, it is suspected that the use of Pierce's causality approach to test the relationship between employment variables, the possibility of feedback between the employment series, and the use of the inconsistent data set may be at the root of the problems.

Feige and Pearce (1979) have indicated that an arbitrary choice of three different test procedures, i.e., the Haugh-Pierce, direct Granger, and Sims approaches, can significantly affect the nature of the economic conclusions. Therefore, it is possible that an implementation of other test approaches could lead to different interrelationship among employment series and, thus, lead to different conclusions.

The employment impact analysis from the transfer function model may not have been successful because of the possibility of feedback between the total employment series and the basic employment series. In Chapter IV, it was noted that dynamic regression models for GEG on TEG and TEG on GEG could be fitted at the same time, and thereby improve the forecast ability of GEG and TEG, individually. However, Pierce's causality test and the Chi-square test do not support this hypothesis.

Using the inconsistent data set is also another possible explanation for these disappointing results. As we have discussed in Chapter IV, the total employment series is adjusted for multiple job-holding and commuting while the industries employment series are not. Thus, it is possible that these differences lead to lower estimates of the employment multipliers.

An employment impact analysis using econometric analysis was also presented and discussed in Chapter V. It was found that the time lag between the initial changes in the causal variable and induced changes in the dependent variable had little effect in the employment multiplier analysis. That is to say, in a given month when a change in the level of employment in the basic sector is observed, a change in the total number employed in all sectors occurred within the same month. The rapidity in adjustment may in part be explained by the small size of the Grant County economy.

The descending order for agriculture and lumber and wood products manufacturing employment multipliers suggests that in a small region such as Grant County a loss of agriculture jobs will have a more severe impact than a loss of an equal number of jobs in government and lumber and wood products manufacturing sectors. Of course, this conclusion is different than the results from income multipliers.

In summary, transfer function models are quite appealing if one is primarily interested in forecasting. They are less desirable if the goal is to analyze regional impacts or to explain the "complicated" behavior of an economic system. If the researcher wants to analyze or predict the total regional impacts, both direct and indirect, of some significant action, such as a new industrial plant or government installation, more accurate and consistent results would be obtained by using regional econometric or input-output models.

Recommendations for Further Research

There are several suggestions for further research. First, the transfer function modeling procedures need to be tested further in practical modeling applications. This is true particularly for systems of three or more variables. Second, it would be interesting to compare results from the Haugh-Pierce, direct Granger, and Sims approaches to test the relationships between employment variables (Feige and Pearce, 1979). It is expected that different conclusions will come from different approaches.

Third, another interesting extension of the present study would be to fit a feedback system among three employment series. That is, a simultaneous structural model consisting of total employment, government employment, and lumber and wood products manufacturing employment should be considered. Granger and Newbold (1976 and 1977) have discussed the modeling procedures for such a feedback system and some possible uses for them.

BIBLIOGRAPHY

1. Barnard, J.R. and R.J. Ballmon. "Tests of Regional intersectoral Flows Analysis multipliers." Journal of Regional Science, Vol. 19, No. 2 (1979): 201-215.

2. Bendavid, Avrom. Regional Economic Analysis for Practitioners. New York: Praeger Publishers, 1974.

3. Bennett, R.J. "Process Identification for Time Series Modeling in Urban and Regional Planning." Regional Studies, Vol. 8 (1974): 157-174.

4. Billings, R. Bruce. "The Mathematical Identity of the Multipliers Derived From The Economic Base Model and the Input-Output Model." Journal of Regional Science, Vol. 9, No. 3 (1969): 471-473.

5. Box, G.E.P. and G.M. Jenkins. Time Series Analysis: Forecasting and Control." Revised Edition. San Francisco: Holden-Day, 1976.

6. ________________. "Some Recent Advances in Forecasting and Control." Applied Statistics, Vol. 17, No. 2 (1966): 91-109.

7. Box, G.E.P. and G.C. Tiao. "Intervention Analysis with Application to Economic and Environmental Problems." Journal of the American Statistical Association, Vol. 70, No. 349 (March 1975): 70-79.

8. Bradley, I.E. and J.P. Gauder. "Input-output Multipliers: Some Theoretical Comments." Journal of Regional Science, Vol. 9, No. 2 (1969): 309-317.

9. Bromley, D.W., G.E. Blanch, and H.H. Stoevener. Effect of Selected Changes in Federal Land Use on a Rural Economy. Station Bulletin 604, Agricultural Experiment Station, Oregon State University, Corvallis, OR, March 1968.

10. Chatfield, C. and D.L. Prothero. "Box-Jenkins Seasonal Forecasting: Problems in a Case-Study." Journal of the Royal Statistical Society - Series A. Vol. 136 (1973): 295-315.

11. Chatfield, C. The Analysis of Time Series: Theory and Practice. London: A Halsted Press, 1975.

12. Chatfield, C. "Some Recent Developments in Time-Series Analysis." Journal of the Royal Statistical Society - Series A. Vol. 140 (1977): 492-510.

13. Collin, Theodore G. An Interindustry Analysis of the Effects of a New Industry on the Public and Private Sectors in Clatsop County, Oregon. M.S. Thesis, Oregon State University, June 1970.

14. Connaughton, Kent P. and William McKillop. "Estimation of 'small area' Multipliers for the Wood Processing Sector--an Econometric Approach." Forest Science, Vol. 25, No. 1 (1979): 7-20.

15. Cook, Thomas. "An Application of the Transfer Function to An Economic-Base Model." The Annals of Regional Science, Vol. XIII, No. 2 (July 1979): 67-80.

16. Conopask, Jeff V. A Data-Pooling Approach to Estimate Employment Multipliers for Small Regional Economies. U.S.D.A., Technical Bulletin No. 1583, May 1978.

17. Cramer, Robert H. and Robert B. Miller. "Dynamic Modeling of Multivariate Time Series for Use in Bank Analysis." Journal of Money, Credit and Banking, Vol. 8 (February 1976): 85-96.

18. Davis, H. Craig. "Economic Base and Input-output Multipliers A Comparison for Vancouver, B.C." The Annals of Regional Science, Vol. IX, No. 3 (November 1975): 1-8.

19. Davis, H. Craig. "Regional Sectoral Multipliers with Reduced Data Requirements." International Regional Science Review, Vol. 1, No. 2 (1976): 18-29.

20. Dent, W.T. and J.A. Swanson. "Forecasting with Limited Information: ARIMA Models of the Trailer on Flatcar Transportation Market." Journal of the American Statistical Association, Vol. 73, No. 362 (June 1978): 293-299.

21. Feige, Edgar L. and Douglas K. Pearce. "The Casual Causal Relationship Between Money and Income: Some Caveats for Time Series Analysis." The Review of Economics and Statistics, Vol. LXI, No. 4 (November 1979): 521-533.

22. Garnick, D.H. "Disaggregated Basic - Service Models in Multiregional Projections." Journal of Regional Science 9 (1969): 87-100.

23. Garnick, D.H. "Differential Regional Multiplier Models." Journal of Regional Science 10 (1970): 35-47.

24. Giarratani, Frank and Paul D. McNelis. "Time Series Evidence Bearing on Crude Theories of Regional Growth." Land Economics Vol. 56, No. 2 (May 1980): 238-248.

25. Glickman, N.J. Econometric Analysis of Regional Systems: Explorations in Model Building and Policy Analysis. New York: Academic Press, 1977.

26. Gibson, L.J. and Marshall A. Worden. Estimating the Economic Base Multiplier: A Test of Alternative Procedures. The University of Arizona, 1979.

27. Granger, Clive W.J. "Investigating Causal Relations by Econometric Models and Cross Spectral Methods." Econometrica, 37 (May 1969): 424-438.

28. Granger, C.W.J. and P. Newbold. "The Time Series Approach to Econometric Model Building." New Methods in Business Cycle Research: Proceedings From a Conference. Federal Reserve Bank of Minneapolis, October 1977.

29. ________________. Forecasting Economic Time Series New York: Academic Press, 1977.

30. Green, R. Jeffery and Gary R. Albrecht. "Testing for Causality in Regional Econometric Models." International Regional Science Review, Vol. 4, No. 2 (1979): 155-164.

31. Greytak, D. "A Statistical Analysis of Regional Export Estimating Techniques." Journal of Regional Science, Vol. 9 (1969): 387-395.

32. Hall, Owen P. and Joseph A Licari. "Building Small Region Econometric Models: Extension of Glickman's Structure to Los Angeles." Journal of Regional Science, Vol. 14, No. 3 (1974): 337-353.

33. Harris, Tiffin D. and Frederick W. Obermiller. Alternatives for Growth in a Based Economy: A Pilot Area Study for Grant County, Oregon. Oregon State University, May 1978.

34. Haugh, L.D. The Identification of Time Series Interrelationships with Special Reference to Dynamic Regression Models. Unpublished Ph.D. Dissertation, Madison: University of Wisconson, Statistics Department, 1972.

35. ________________. "Checking the Independence of Two Covariance-- Stationary Time Series: A Univariate Residual Cross-Correlation Approach." Journal of the American Statistics Association, 71 (354): 378-385, 1976.

36. Haugh, L.D. and G.E.P. Box. "Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series." Journal of the American Statistics Association, 72 (357): 121-130,1977.

37. Isard, Walter. Methods of Regional Analysis: An Introduction to Regional Science. Cambridge, Massachusetts, MIT Press, 1960.

38. Isard, W. and S. Czamanski. "Techniques for Estimating Local and Regional Multiplier Effects of Changes in the Level of Major Governmental Programs." Peace Research Society: Papers, Vol. III (1965): 19-45.

39. Johnson, Thomas G. A Dynamic Input-Output Model for Regional Impact Analysis. Ph.D. Dissertation. Oregon State University, July 1979.

40. Klein, Lawrence R. and Norman J. Glickman." Econometric Model-Building at Regional Level." Regional Science and Urban Economics, 7 (1977): 3-23.

41. Koutsoyiannis, A. Theory of Econometrics. Revised Edition. Ontario: University of Waterloo, 1977.

42. Lewis, W.C. "Export Base Theory and the Multiplier Estimation: A Critique." The Annals of Regional Science, Vol. 10, No. 2 (1976): 58-70.

43. Liew, Chong K. "Dynamic Multipliers for a Regional Input-Output Model." The Annals of Regional Science. Vol. XI, No. 3 (November 1977): 94-106.

44. Makridakis, Spyros and Steven C. Wheelwright. Forecasting: Methods and Applications. New York: John Wiley & Sons, Inc. 1978.

45. Marquardt, D.W. "An Algorithm for Least Squares Estimation of Nonlinear Parameters." Journal of Society for Industrial and Applied Mathematics. 2 (1963): 431-441.

46. McNulty, James E. "A Test of the Time Dimension In Economic Base Analysis." Land Economics, Vol. 53, No. 3 (August 1977): 359-368.

47. Moody, H.L. and F.W. Puffer. "A Gross Regional Product Approach to Regional Model Building." Western Economic Journal 7 (1969): 391-402.

48. Moore, Fredrick and James Peterson. "An Interindustry Model of Utah." Review of Economics and Statistics, 38 (November 1955): 368-383.

49. Narasimham, G.V.L., Victor F. Castellino, and Nozer D. Singpur-Walla. "On the Predictive Performance of the BEA Quarterly Econometric Model and A Box-Jenkins Type ARIMA Model." American Statistical Association 1974 Proceedings of the Business and Economic Statistics Section, American Statistical Association, pp. 501-504.

50. Naylor, Thomas H., Terry G. Seaks, and D.W. Wichern. "Box-Jenkins Methods: An Alternative to Econometric Models." International Statistical Review, Vol. 40, No. 2 (1972): 123-137.

51. Nelson, Charles R. Applied Time Series Analysis: for Managerial Forecasting. San Francisco: Holden-Day, 1973.

52. Newbold, P. and C.W.J. Granger. "Experience with Forecasting Univariate Time Series and the Combination of Forecasts." Journal of the Royal Statistical Society - Series A. Vol. 137 (May 1974): 131-146.

53. Obermiller, Frederick W. and Lester F. Miller. Structure of a Resource-Based Economy: Grant County, Oregon. Corvallis, Oregon, Oregon State University, 1979.

54. Oliveira, R.A., Joseph Buongiorno, and Alice M. Kmiotek. "Time Series Forecasting Models of Lumber Cash, Futures, and Basis Prices." Forest Science. Vol. 23, No. 2 (June 1977): 268-280.

55. Oliveira, Ronald A. "'Combining Forecasts to Predict Property Values for Single-Family Residences' Comment." Land Economics, 54, 4 (November 1978): 524-527.

56. Oliveira, R.A. and G.W. Whittaker. An Examination of Dynamic Relationships--and the Lack Thereof--Among U.S. Lumber Prices, U.S. Housing Starts, U.S. Log Exports to Japan, and Japanese Housing Starts." Special Report 565, Agricultural Experiment Station, Oregon State University, December 1979.

57. Oregon Employment Division. Research and Statistics Section. Oregon Covered Employment and Payrolls by Industry and County First Quarter 1977, Salem.

58. Pack, David J. "Revealing Time Series Interrelationships." Decision Sciences 8 (April 1977): 377-402.

59. Palmer, Charles, Nelson Bills, and Robert Niehaus. "Input-Output Concepts." Regional Development and Plan Evaluation/The Use of Input-Output Analysis, U.S.D.A., Agriculture Handbook No. 530 (May, 1978):22.

60. Park, S.H. "Least Squares Estimates of the Regional Employment Multiplier: An Appraisal." Journal of Regional Science, Vol. 10, No. 3 (1970): 365-374.

61. Pierce, David A. "Residual Correlations and Diagnostic Checking in Dynamic Disturbance Time Series Models." Journal of the American Statistical Association, Vol. 67, No. 339 (September 1972): 636-640.

62. ________________. "Relationships--and the Lack Thereof--Between Economic Time Series, with Special Reference to Money and Interest Rates." Journal of the American Statistical Association. Vol. 72, No. 357 (March, 1977): 11-22.

63. Pierce, David, A. and Larry D. Haugh. "Causality in Temporal Systems: Characterizations and A Survey." Journal of Econometrics, 5 (1977).

64. Polzin, Paul E. "How to Win Friends, Influence People, and Sleep at Night after Using Economic Base Analysis." Paper presented at 1980 WRSA meetings.

65. Rausser, Gordon C. and Ronald A. Oliveira. "An Econometric Analysis of Wilderness Area Use." Journal of The American Statistical Association, Vol. 71, No. 354 (June 1976): 276-285.

66. Romanoff, Eliahu. "The Economic Base Model: A Very Special Case of Input-Output Analysis." Journal of Regional Science, Vol. 14, No. 1 (1974): 121-129.

67. Sasaki, K. "Expenditures and the Employment Multiplier in Hawaii." Review of Economics and Statistics. 45 (1963): 298-304.

68. Schmidt, James R. "Forecasting State Retail Sales: Econometric vs. Time Series Models." The Annals of Regional Science, Vol. 13, No. 1 (November 1979): 91-101.

69. Schmidt, Peter and Roger N. Waud. "The Almon Lag Technique and the Monetary Versus Fiscal Policy Debate." Journal of the American Statistical Association, Vol. 68, No. 341 (March 1973): 11-19.

70. Spivey, W.A. and W.E. Wecker. "Regional Economic Forecasting: Concepts and Methodology." Papers of the Regional Science Association, Vol. 28 (November 1971): 257-276.

71. Stewart, Douglas B. and Yiannis P. Venieris. "Dynamic Multipliers: A note on the Generalization to Many Lags." Journal of Regional Science, Vol. 18, No. 3 (1978): 459-462.

72. Stokes, Houston H., Donald W. Jones and Hugh M. Neuburger." Short-Run Macroeconomic Adjustment in a Small Region." The Wall Street Review of Books, Vol. 4, No. 3 (September 1976): 5-16.

73. Theil, H. and J.C.G. Boot. "The Final Form of Econometric Equation Systems." Review of the International Statistical Institute, 30 (1962): 136-152.

74. Theil, Henri. Applied Economic Forecasting. Chicago: Rand-McNally, 1966.

75. Thompson, Howard E. and George C. Tiao. "Analysis of Telephone Data: A Case Study of Forecasting Seasonal Time Series." Bell Journal of Economic and Management Science, 2 (1971): 515-541

76. Tweeten, Luther and George L. Brinkman. Micropolitan Development - Theory and Practice of Greater Rural Economic Development, The Iowa State University Press, AMES, 1976.

77. Weiss, S.J. and E.C. Gooding. "Estimation of Differential Multipliers in a Small Regional Economy." Land Economics 44 (1968): 235-244.

78. Youmans, Russell, and A. Haroldsen. Grant County: Structure of the County Economy. Special Report 358, Extension Service, Oregon State University, Corvallis, Oregon, May 1972.

APPENDIX I

ARMA MODEL STRUCTURE

The basic components of the ARMA model are the autoregressive and moving-average processes. In the autoregressive model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a shock μ_t. Let us denote the values of a process at equally spaced times t, t-1, t-2, $\cdots$ by Z_t, Z_{t-1}, Z_{t-2}, $\cdots$. Also let $\tilde{Z}_t$, $\tilde{Z}_{t-1}$, $\tilde{Z}_{t-2}$ $\cdots$ be deviations from($\mu = E(Z_t)$) for example, $\tilde{Z}_t = Z_t - \mu$. Then

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \cdots + \phi_p \tilde{Z}_{t-p} + \mu_t$$

is called an autoregressive (AR) process of order p. The reason for this name is that an AR model is essentially a regressive equation in which $\tilde{Z}_t$ is related to its own past values where the number of nonzero term is finite.

If we define an autoregressive operator of order p by

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 \cdots - \phi_p B^p$$

Then the autoregressive model may be written economically as

$$\phi_p(B) \tilde{Z}_t = \mu_t$$

The model contains p + 2 unknown parameters μ, ϕ_1, ϕ_2, $\cdots$ ϕ_p, σ^2_μ, which in practice have to be estimated from the data.

The moving-average (MA) model expresses $\tilde{Z}_t$ as a linear function of the current value and q past values of a random shock series, i.e., μ_t, μ_{t-1}, μ_{t-q}, and may be written as:

$$\tilde{Z}_t = \mu_t - \theta_1\mu_{t-1} - \theta_2\mu_{t-2}\cdots - \theta_q\mu_{t-q}$$

If we define a moving average operator of order q by

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

then the moving average model may be written economically as

$$\tilde{Z}_t = \theta_q(B)\mu_t$$

It contains q + 2 unknown parameters μ, θ_1, θ_2, $\cdots$ θ_q, σ_μ^2, which in practice have to be estimated from the data.

In an attempt to achieve greater flexibility, one may employ both autoregressive and moving average components within the same model. This leads to the mixed autoregressive-moving average (ARMA) model

$$\tilde{Z}_t = \phi_1\tilde{Z}_{t-1} + \cdots + \phi_p\tilde{Z}_{t-p} + \mu_t - \theta_1\mu_{t-1} - \cdots - \theta_q\mu_{t-q}$$

or

$$\phi_p(B)\tilde{Z}_t = \theta_q(B)\mu_t$$

It contains p + q + 2 unknown parameters μ; ϕ_1, $\cdots$ ϕ_p; θ_1, $\cdots$ θ_q; σ_μ^2, which are estimated from the data.

APPENDIX II

ESTIMATED AUTOCORRELATION FUNCTION AND PARTIAL AUTOCORRELATION FUNCTION

The covariance between Z_t and its value Z_{t+K}, separated by K intervals of time, is called the autocovariance of lag K and is defined by

$$\gamma_K = COV(Z_t, Z_{t+K}) = E[(Z_t - \mu)(Z_{t+K} - \mu)]$$

Similarly the autocorrelation at lag K is

$$\rho_K = \frac{E[(Z_t - \mu)(Z_{t+K} - \mu)]}{\sqrt{E[(Z_t - \mu)^2] \; E\,[(Z_{t+K} - \mu)^2]}}$$

$$= \frac{E[(Z_t - \mu)(Z_{t+K} - \mu)]}{\sigma_Z^2}$$

Given that the stationary process, for variance $\sigma_Z^2 = \gamma_o$ is the same at time t+K as at time t, the autocorrelation at lag K is

$$\rho_K = \frac{\gamma_K}{\gamma_o}$$

The partial autocorrelation at lag K is ϕ_{KK}, and the ϕ_{KK} satisfy the set of equations

$$\text{(A.1)} \quad \rho_j = \phi_{K1}\rho_{j-1} + \cdots + \phi_{K(K-1)}\rho_{j-k+1} + \phi_{KK}\rho_{j-K}, \; j = 1, 2, \cdots K .$$

A number of estimates of the autocorrelation function have been suggested by statisticians. It is concluded that the most satisfactory estimate of the Kth lag autocorrelation ρ_K is (Box and Jenkins, 1976)

$$r_K = \frac{C_K}{C_o} \quad ,$$

where

$$C_K = \frac{1}{N} \sum_{t=1}^{N-K} (Z_t - \overline{Z})(Z_{t+K} - \overline{Z}) \qquad K = 0, 1, 2, \cdots \; k ,$$

is the estimate of the autocovariance γ_K, and $\overline{Z}$ is the mean of the time series. The estimated partial autocorrelations then can be obtained by substituting estimates r_j for the theoretical autocorrelations in (A.1), to yield

$$r_j = \hat{\phi}_{K1} r_{j-1} + \hat{\phi}_{K2} r_{j-2} + \cdots + \hat{\phi}_{K(K-1)} r_{j-K+1} + \hat{\phi}_{KK} r_{j-K}$$

$$j = 1, 2, \cdots K ,$$

and solving the resultant equations for K = 1, 2, $\cdots$, we can then get

$$\hat{\phi}_{KK} = \begin{cases} r_1 & K = 1 \\[2ex] \dfrac{r_K - \sum_{j=1}^{K-1} \hat{\phi}_{K-1,j} r_{K-j}}{1 - \sum_{j=1}^{K-1} \hat{\phi}_{K-1,j} r_j} & \begin{array}{l} K = 2, 3, \cdots L \\[2ex] L \le K \end{array} \end{cases}$$

where

$$\hat{\phi}_{Kj} = \hat{\phi}_{K-1,j} - \hat{\phi}_{KK}\,\hat{\phi}_{K-1,K-j} \qquad j = 1, 2, \cdots \; K-1 \, .$$

APPENDIX III

COMPUTER PRINTOUT FOR THE ESTIMATED AUTOCORRELATION AND PARTIAL AUTOCORRELATION FUNCTIONS FOR SERIES TEG, GEG, LWP, AND AGG

AUTOCORRELATION FUNCTION

DATA - TEG

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .28452E+04
ST. DEV. OF SERIES = .29088E+03
NUMBER OF OBSERVATIONS = 96

1- 12	.84	.54	.18	-.14	-.37	-.46	-.40	-.23	.05	.35	.59	.69
ST.E.	.10	.16	.18	.18	.18	.19	.20	.21	.21	.21	.22	.23
13- 24	.58	.32	.01	-.30	-.50	-.58	-.52	-.37	-.11	.15	.36	.45
ST.E.	.25	.27	.27	.27	.27	.28	.30	.30	.31	.31	.31	.32
25- 36	.38	.18	-.07	-.32	-.48	-.55	-.50	-.36	-.15	.09	.29	.39
ST.E.	.32	.33	.33	.33	.33	.34	.35	.35	.36	.36	.36	.36

MEAN DIVIDED BY ST. ERROR = .95838E+02

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .61409E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES = .63158E+01
ST. DEV. OF SERIES = .16084E+03
NUMBER OF OBSERVATIONS = 95

1- 12	.41	.19	-.09	-.33	-.44	-.44	-.35	-.30	-.11	.21	.42	.67
ST.E.	.10	.12	.12	.12	.13	.15	.16	.17	.17	.17	.18	.19
13- 24	.48	.15	-.03	-.34	-.38	-.41	-.27	-.32	-.04	.17	.36	.51
ST.E.	.21	.22	.22	.22	.23	.23	.24	.25	.25	.25	.25	.26
25- 36	.41	.14	-.03	-.24	-.28	-.35	-.26	-.21	-.11	.14	.27	.44
ST.E.	.27	.27	.27	.27	.28	.28	.28	.29	.29	.29	.29	.29

MEAN DIVIDED BY ST. ERROR = .38269E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .38643E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 2
MEAN OF THE SERIES = .63930E+00
ST. DEV. OF SERIES = .17447E+03
NUMBER OF OBSERVATIONS = 94

1- 12	-.31	.06	-.04	-.11	-.09	-.08	.03	-.13	-.11	.09	-.03	.38
ST.E.	.10	.11	.11	.11	.11	.12	.12	.12	.12	.12	.12	.12
13- 24	.12	-.13	.12	-.23	.00	-.16	.15	-.29	.07	.01	.04	.22
ST.E.	.13	.13	.13	.14	.14	.14	.14	.14	.15	.15	.15	.15
25- 36	.14	-.09	.05	-.14	.03	-.16	.04	-.06	-.11	.10	-.04	.23
ST.E.	.15	.15	.15	.16	.16	.16	.16	.16	.16	.16	.16	.16

MEAN DIVIDED BY ST. ERROR = .35470E-01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .27700E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - TEG

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .28452E+04
ST. DEV. OF SERIES = .29088E+03
NUMBER OF OBSERVATIONS = 96

1- 12	.84	-.53	-.32	-.05	-.02	.03	.12	.12	.28	.27	.06	-.00
13- 24	-.35	-.16	.07	-.16	.07	-.06	-.05	-.12	.10	-.17	.03	.01
25- 36	-.09	.06	.07	-.06	.02	-.12	.04	-.05	-.10	.09	.09	-.05

DIFFERENCE 1
MEAN OF THE SERIES = .63158E+01
ST. DEV. OF SERIES = .16086E+03
NUMBER OF OBSERVATIONS = 95

1- 12	.41	.02	-.20	-.29	-.24	-.21	-.21	-.37	-.34	-.11	-.01	.34
13- 24	.11	-.17	.07	-.05	.08	-.03	.09	-.19	.13	-.05	-.06	-.03
25- 36	-.10	-.11	.01	-.01	.09	-.05	.00	.04	-.03	-.10	-.01	.03

DIFFERENCE 2
MEAN OF THE SERIES = .63830E+00
ST. DEV. OF SERIES = .17447E+03
NUMBER OF OBSERVATIONS = 94

1- 12	-.31	-.04	-.03	-.15	-.19	-.20	-.09	-.23	-.40	-.33	-.46	-.07
13- 24	.21	-.07	.08	-.05	.04	-.08	.16	-.18	.04	.03	-.02	.05
25- 36	.03	-.09	-.03	-.10	.08	.02	-.03	.07	.10	-.04	-.05	-.08

AUTOCORRELATION FUNCTION

DATA - TEG 96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1) 1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .50000E+02
ST. DEV. OF SERIES = .15853E+03
NUMBER OF OBSERVATIONS = 84

1- 12	.79	.75	.69	.67	.57	.53	.47	.39	.26	.19	.10	-.04
ST.E.	.11	.16	.20	.23	.25	.26	.28	.29	.29	.29	.30	.30
13- 24	-.06	-.14	-.19	-.29	-.30	-.33	-.36	-.41	-.34	-.36	-.37	-.37
ST.E.	.30	.30	.30	.30	.30	.31	.31	.32	.32	.33	.33	.34
25- 36	-.36	-.34	-.28	-.25	-.24	-.25	-.19	-.16	-.17	-.14	-.09	-.04
ST.E.	.34	.35	.35	.35	.35	.36	.36	.36	.36	.36	.36	.36

MEAN DIVIDED BY ST. ERROR = .28907E+01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .48459E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES = .21687E+01
ST. DEV. OF SERIES = .10090E+03
NUMBER OF OBSERVATIONS = 83

1- 12	-.39	.01	-.08	.19	-.18	.09	.05	.12	-.17	.07	.10	-.27
ST.E.	.11	.13	.13	.13	.13	.13	.13	.13	.13	.14	.14	.14
13- 24	.14	-.05	.10	-.22	.03	-.02	.09	-.31	.25	-.05	-.01	-.07
ST.E.	.14	.15	.15	.15	.15	.15	.15	.15	.16	.16	.16	.16
25- 36	.01	-.12	.12	.01	.06	-.10	.03	.13	-.12	.01	-.03	.09
ST.E.	.16	.16	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17

MEAN DIVIDED BY ST. ERROR = .19581E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .66854E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - TEG

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1) 1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .50000E+02
ST. DEV. OF SERIES = .15853E+03
NUMBER OF OBSERVATIONS = 84

1- 12	.79	.31	.10	.12	-.15	.01	-.03	-.15	-.25	-.07	-.12	-.27
13- 24	.17	-.07	.03	-.06	.06	.10	.02	-.09	.19	-.01	-.14	-.09
25- 36	-.13	-.03	.18	-.16	-.10	-.03	.03	.02	-.02	-.07	.08	.16

DIFFERENCE 1
MEAN OF THE SERIES = .21687E+01
ST. DEV. OF SERIES = .10090E+03
NUMBER OF OBSERVATIONS = 83

1- 12	-.39	-.17	-.16	.12	-.07	.01	.11	.20	.01	.01	.14	-.27
13- 24	-.03	-.13	-.02	-.15	-.22	-.13	.01	-.22	-.03	.09	.09	.10
25- 36	-.00	-.20	.08	.03	-.04	-.06	-.03	.05	.03	-.08	-.17	-.10

AUTOCORRELATION FUNCTION

96 OBSERVATIONS

DATA - GEG

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .79073E+03
ST. DEV. OF SERIES = .10620E+03
NUMBER OF OBSERVATIONS = 96

1- 12	.90	.74	.55	.35	.21	.14	.16	.24	.37	.51	.61	.65
ST.E.	.10	.16	.20	.21	.22	.22	.22	.22	.22	.23	.24	.26
13- 24	.58	.47	.31	.14	.03	-.02	-.01	.06	.17	.27	.34	.36
ST.E.	.27	.29	.29	.30	.30	.30	.30	.30	.30	.30	.30	.31
25- 36	.31	.21	.07	-.05	-.14	-.19	-.15	-.09	.01	.09	.16	.19
ST.E.	.31	.31	.32	.32	.32	.32	.32	.32	.32	.32	.32	.32

MEAN DIVIDED BY ST. ERROR = .72955E+02

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .47426E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES = .32632E+01
ST. DEV. OF SERIES = .43077E+02
NUMBER OF OBSERVATIONS = 95

1- 12	.25	.21	-.01	-.24	-.44	-.50	-.36	-.23	-.01	.24	.34	.52
ST.E.	.10	.11	.11	.11	.12	.13	.15	.16	.16	.16	.17	.18
13- 24	.29	.22	-.04	-.27	-.35	-.31	-.25	-.21	.07	.21	.23	.32
ST.E.	.19	.20	.20	.20	.20	.21	.21	.22	.22	.22	.22	.22
25- 36	.32	.14	-.08	-.18	-.22	-.33	-.18	-.14	.03	.11	.16	.35
ST.E.	.23	.23	.23	.23	.24	.24	.24	.24	.24	.25	.25	.25

MEAN DIVIDED BY ST. ERROR = .73833E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .26392E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 2
MEAN OF THE SERIES =-.53191E+00
ST. DEV. OF SERIES = .52802E+02
NUMBER OF OBSERVATIONS = 94

1- 12	-.48	.12	.02	-.02	-.10	-.11	.00	-.06	-.03	.10	-.06	[illegible]
ST.E.	.10	.12	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13
13- 24	-.12	.13	-.00	-.11	-.06	-.01	.00	-.16	.10	.07	-.06	.07
ST.E.	.14	.14	.14	.14	.14	.14	.14	.14	.14	.14	.14	.14
25- 36	.12	.02	-.06	-.04	.05	-.19	.08	-.08	.05	.02	-.10	.24
ST.E.	.14	.14	.14	.14	.14	.14	.15	.15	.15	.15	.15	.15

MEAN DIVIDED BY ST. ERROR = .97669E-01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .56071E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - GEG 96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .79073E+03
ST. DEV. OF SERIES = .10620E+03
NUMBER OF OBSERVATIONS = 96

1- 12	.90	-.30	-.27	-.06	.15	.28	.21	.16	.12	.10	.01	-.01
13- 24	-.26	-.02	-.09	-.02	.14	-.06	-.01	-.06	.08	-.14	.02	.04
25- 36	-.05	-.10	-.04	.06	.01	-.04	.12	-.11	-.02	-.00	.05	.12

DIFFERENCE 1
MEAN OF THE SERIES = .32632E+01
ST. DEV. OF SERIES = .43077E+02
NUMBER OF OBSERVATIONS = 95

1- 12	.25	.15	-.10	-.27	-.37	-.38	-.24	-.22	-.19	-.09	-.12	.12
13- 24	-.07	-.01	-.09	-.18	-.11	.08	.10	-.08	-.00	-.01	-.10	-.06
25- 36	.14	.09	.00	.01	.09	-.06	.05	-.00	-.03	-.03	-.12	.10

DIFFERENCE 2
MEAN OF THE SERIES =-.53191E+00
ST. DEV. OF SERIES = .52802E+02
NUMBER OF OBSERVATIONS = 94

1- 12	-.48	-.14	.03	.02	-.14	-.31	-.27	-.25	-.27	-.16	-.27	.04
13- 24	-.01	.05	.11	-.03	-.16	-.08	.11	-.00	.02	.07	-.02	-.16
25- 36	-.02	.10	.07	-.01	.10	-.07	-.02	.00	-.00	.07	-.13	.04

AUTOCORRELATION FUNCTION

DATA - GEG 96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1) 1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .35476E+02
ST. DEV. OF SERIES = .41098E+02
NUMBER OF OBSERVATIONS = 84

1- 12	.52	.29	.19	.08	-.09	-.04	.06	.04	-.00	.09	.05	-.21
ST.E.	.11	.14	.14	.15	.15	.15	.15	.15	.15	.15	.15	.15
13- 24	-.11	.01	.06	.10	.21	.27	.12	.06	.07	-.03	-.11	-.14
ST.E.	.15	.15	.15	.15	.15	.16	.16	.16	.16	.16	.16	.16
25- 36	-.03	-.12	-.14	-.07	-.04	-.12	-.04	-.07	-.10	-.15	-.04	.05
ST.E.	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17

MEAN DIVIDED BY ST. ERROR = .79114E+01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .70096E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES = .15663E+01
ST. DEV. OF SERIES = .39162E+02
NUMBER OF OBSERVATIONS = 83

1- 12	-.29	-.10	.03	.06	-.21	-.09	.14	.08	-.17	.13	.22	-.31
ST.E.	.11	.12	.12	.12	.12	.12	.13	.13	.13	.13	.13	.14
13- 24	-.04	.04	-.01	-.08	-.00	.23	-.05	-.10	.14	.00	-.05	-.21
ST.E.	.14	.14	.14	.14	.15	.15	.15	.15	.15	.15	.15	.15
25- 36	.24	-.09	-.08	.05	.13	-.13	.05	.01	.04	-.16	-.01	.09
ST.E.	.16	.16	.16	.16	.16	.16	.16	.16	.16	.16	.17	.17

MEAN DIVIDED BY ST. ERROR = .36436E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .62927E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - GEG

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1) 1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .35476E+02
ST. DEV. OF SERIES = .41098E+02
NUMBER OF OBSERVATIONS = 84

1- 12	.52	.03	.03	-.06	-.16	.09	.12	-.02	-.06	.11	-.06	-.30
13- 24	.17	.11	.09	.05	.03	.13	-.08	-.03	.01	.00	-.04	-.26
25- 36	.16	-.12	-.04	.04	.01	.04	-.04	-.15	-.03	-.05	.06	-.02

DIFFERENCE 1
MEAN OF THE SERIES = .15663E+01
ST. DEV. OF SERIES = .39162E+02
NUMBER OF OBSERVATIONS = 83

1- 12	-.29	-.20	-.08	.03	-.21	-.26	-.06	.07	-.12	.01	.22	-.17
13- 24	-.13	-.12	-.11	-.04	-.13	.02	.02	-.03	.02	.04	.18	-.21
25- 36	.09	-.07	-.10	-.03	-.08	-.03	.10	-.00	.04	-.04	-.01	-.17

AUTOCORRELATION FUNCTION

DATA - LWP   96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .64229E+03
ST. DEV. OF SERIES = .12030E+03
NUMBER OF OBSERVATIONS = 96

1- 12	.80	.45	.11	-.13	-.28	-.35	-.34	-.22	-.02	.24	.50	.64
ST.E.	.10	.15	.17	.17	.17	.17	.18	.19	.19	.19	.19	.21
13- 24	.52	.24	-.05	-.27	-.42	-.49	-.50	-.40	-.21	.03	.27	.40
ST.E.	.23	.24	.24	.24	.24	.25	.26	.27	.28	.28	.28	.28
25- 36	.33	.10	-.13	-.31	-.42	-.48	-.46	-.36	-.19	.04	.26	.41
ST.E.	.29	.29	.29	.29	.30	.30	.31	.32	.32	.32	.32	.32

MEAN DIVIDED BY ST. ERROR = .52313E+02

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .47653E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES = .17895E+01
ST. DEV. OF SERIES = .73240E+02
NUMBER OF OBSERVATIONS = 95

1- 12	.43	-.06	-.25	-.25	-.23	-.18	-.20	-.19	-.21	-.06	.33	.67
ST.E.	.10	.12	.12	.13	.13	.14	.14	.14	.14	.15	.15	.15
13- 24	.45	.01	-.19	-.19	-.21	-.18	-.18	-.22	-.18	-.00	.26	.55
ST.E.	.18	.19	.19	.20	.20	.20	.20	.20	.21	.21	.21	.21
25- 36	.41	.03	-.18	-.17	-.16	-.16	-.18	-.17	-.18	-.00	.19	.44
ST.E.	.23	.23	.23	.23	.24	.24	.24	.24	.24	.24	.24	.24

MEAN DIVIDED BY ST. ERROR = .23814E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .25482E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 2
MEAN OF THE SERIES = .85106E+00
ST. DEV. OF SERIES = .77852E+02
NUMBER OF OBSERVATIONS = 94

1- 12	-.08	-.23	-.17	-.02	-.03	.06	-.03	.02	-.15	-.20	.03	.51
ST.E.	.10	.10	.11	.11	.11	.11	.11	.11	.11	.11	.12	.12
13- 24	.17	-.19	-.16	.01	-.06	.05	.01	-.07	-.13	-.09	-.00	.36
ST.E.	.14	.14	.14	.15	.15	.15	.15	.15	.15	.15	.15	.15
25- 36	.21	-.14	-.18	-.01	.01	.01	-.03	.02	-.17	-.01	-.08	.31
ST.E.	.16	.16	.16	.16	.16	.16	.16	.16	.16	.17	.17	.17

MEAN DIVIDED BY ST. ERROR = .10599E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .88781E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - LWP 96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .64229E+03
ST. DEV. OF SERIES = .12030E+03
NUMBER OF OBSERVATIONS = 96

1- 12	.80	-.56	.01	-.10	-.11	-.06	-.02	.18	.09	.33	.25	-.02
13- 24	-.37	-.10	.02	-.12	-.11	-.01	-.14	-.01	-.04	.01	-.04	.01
25- 36	-.16	-.05	.03	-.03	-.03	-.00	.07	-.10	-.05	.12	-.02	.07

DIFFERENCE 1
MEAN OF THE SERIES = .17895E+01
ST. DEV. OF SERIES = .73240E+02
NUMBER OF OBSERVATIONS = 95

1- 12	.43	-.30	-.12	-.11	-.17	-.13	-.25	-.23	-.39	-.27	.06	.36
13- 24	.04	-.08	.07	.09	-.03	.03	-.03	-.11	-.02	.05	-.05	.09
25- 36	-.01	-.07	-.05	.00	.00	-.05	-.03	.06	-.11	.03	-.12	.06

DIFFERENCE 2
MEAN OF THE SERIES = .85106E+00
ST. DEV. OF SERIES = .77852E+02
NUMBER OF OBSERVATIONS = 94

1- 12	-.08	-.24	-.23	-.14	-.17	-.07	-.13	-.07	-.26	-.43	-.45	.04
13- 24	.12	-.02	-.02	.10	-.01	.06	.07	-.07	-.10	.01	-.08	.01
25- 36	.09	.04	-.02	.01	.07	.01	-.07	.10	-.05	.08	-.12	.04

AUTOCORRELATION FUNCTION

DATA - LWP

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1) 1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .15476E+02
ST. DEV. OF SERIES = .81422E+02
NUMBER OF OBSERVATIONS = 84

1- 12	.84	.74	.67	.57	.49	.40	.31	.24	.15	.06	-.02	-.14
ST.E.	.11	.17	.20	.23	.25	.26	.26	.27	.27	.27	.27	.27
13- 24	-.14	-.18	-.25	-.31	-.35	-.36	-.37	-.41	-.41	-.42	-.44	-.42
ST.E.	.27	.27	.28	.28	.28	.29	.29	.30	.30	.31	.32	.33
25- 36	-.42	-.42	-.37	-.32	-.30	-.32	-.30	-.24	-.17	-.10	-.03	.03
ST.E.	.33	.34	.34	.35	.35	.36	.36	.36	.36	.36	.36	.36

MEAN DIVIDED BY ST. ERROR = .17421E+01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .50284E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES = .84337E+00
ST. DEV. OF SERIES = .45581E+02
NUMBER OF OBSERVATIONS = 83

1- 12	-.17	-.08	.04	-.04	.02	.01	-.05	.07	-.01	-.06	.15	-.41
ST.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.12	.12
13- 24	.15	.08	.03	-.08	-.11	.01	.10	-.08	-.01	.02	-.11	.06
ST.E.	.13	.14	.14	.14	.14	.14	.14	.14	.14	.14	.14	.14
25- 36	-.01	-.13	-.00	.09	.15	-.15	-.13	.00	-.05	.01	.05	.10
ST.E.	.14	.14	.14	.14	.14	.14	.15	.15	.15	.15	.15	.15

MEAN DIVIDED BY ST. ERROR = .16857E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .41265E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - LWP

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1).1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .15476E+02
ST. DEV. OF SERIES = .81422E+02
NUMBER OF OBSERVATIONS = 84

1- 12	.84	.11	.06	-.08	-.02	-.07	-.06	-.01	-.11	-.08	-.05	-.25
13- 24	.31	-.13	-.08	-.20	.03	.07	-.05	-.10	-.07	-.08	-.03	-.05
25- 36	.07	-.16	.11	-.09	-.03	-.23	.09	.02	.10	.05	.03	-.05

DIFFERENCE 1
MEAN OF THE SERIES = .84337E+00
ST. DEV. OF SERIES = .45581E+02
NUMBER OF OBSERVATIONS = 83

1- 12	-.17	-.12	.00	-.05	.01	.00	-.05	.05	.00	-.05	.13	-.39
13- 24	.07	.02	.11	-.10	-.15	-.02	.05	-.03	-.02	-.09	-.03	-.18
25- 36	.09	-.17	.02	-.03	.15	-.18	-.05	-.18	-.11	-.10	.04	.04

AUTOCORRELATION FUNCTION

DATA - AGG 96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .39670E+03
ST. DEV. OF SERIES = .60386E+02
NUMBER OF OBSERVATIONS = 96

1- 12	.80	.45	.04	-.28	-.45	-.53	-.42	-.24	.06	.41	.70	.86
ST.E.	.10	.15	.17	.17	.17	.18	.20	.21	.21	.21	.22	.24
13- 24	.68	.37	-.00	-.29	-.44	-.50	-.40	-.25	.01	.31	.56	.69
ST.E.	.27	.29	.29	.29	.30	.30	.31	.32	.32	.32	.32	.33
25- 36	.54	.27	-.06	-.30	-.43	-.48	-.38	-.25	-.03	.21	.42	.51
ST.E.	.35	.36	.36	.36	.36	.37	.37	.38	.38	.38	.38	.38

MEAN DIVIDED BY ST. ERROR = .64366E+02

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .71397E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES =-.55789E+00
ST. DEV. OF SERIES = .37288E+02
NUMBER OF OBSERVATIONS = 95

1- 12	.40	.17	-.19	-.37	-.22	-.51	-.21	-.35	-.17	.14	.36	.88
ST.E.	.10	.12	.12	.12	.14	.14	.16	.16	.17	.17	.17	.18
13- 24	.38	.16	-.17	-.33	-.20	-.45	-.19	-.30	-.13	.11	.32	.74
ST.E.	.22	.23	.23	.23	.23	.24	.24	.25	.25	.25	.25	.26
25- 36	.33	.14	-.16	-.27	-.18	-.37	-.16	-.25	-.10	.09	.27	.61
ST.E.	.28	.28	.28	.28	.29	.29	.29	.29	.30	.30	.30	.30

MEAN DIVIDED BY ST. ERROR = .14583E+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .41350E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 2
MEAN OF THE SERIES =-.22340E+00
ST. DEV. OF SERIES = .40877E+02
NUMBER OF OBSERVATIONS = 94

1- 12	-.30	.10	-.15	-.28	.38	-.50	.37	-.26	-.11	.07	-.25	.85
ST.E.	.10	.11	.11	.12	.12	.13	.15	.16	.17	.17	.17	.17
13- 24	-.23	.08	-.14	-.24	.32	-.42	.31	-.23	-.07	.02	-.17	.70
ST.E.	.21	.21	.21	.22	.22	.22	.23	.24	.24	.24	.24	.24
25- 36	-.18	.08	-.15	-.18	.26	-.34	.25	-.20	-.04	.02	-.15	.57
ST.E.	.26	.26	.26	.26	.26	.27	.27	.27	.28	.28	.28	.28

MEAN DIVIDED BY ST. ERROR = .52987E-01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .34227E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - AGG

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA.

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES = .39670E+03
ST. DEV. OF SERIES = .60386E+02
NUMBER OF OBSERVATIONS = 96

1- 12	.80	-.52	-.34	-.01	-.01	-.30	.31	-.12	.39	.43	.31	.18
13- 24	-.44	.17	.20	-.09	.00	-.16	-.07	-.04	-.04	-.09	-.03	-.10
25- 36	-.08	.01	.07	-.05	-.03	.00	-.00	.05	-.04	-.07	-.12	-.01

DIFFERENCE 1
MEAN OF THE SERIES =-.55789E+00
ST. DEV. OF SERIES = .37288E+02
NUMBER OF OBSERVATIONS = 95

1- 12	.40	.01	-.32	-.25	.10	-.58	.04	-.53	-.41	-.25	.09	.54
13- 24	-.26	-.23	.15	.05	-.06	.11	-.02	.02	.15	-.09	.06	-.01
25- 36	-.06	.05	.04	-.03	.05	.03	-.02	.07	.02	.07	-.12	-.05

DIFFERENCE 2
MEAN OF THE SERIES =-.22340E+00
ST. DEV. OF SERIES = .40877E+02
NUMBER OF OBSERVATIONS = 94

1- 12	-.30	.01	-.13	-.40	.24	-.47	.09	-.28	-.41	-.52	-.64	.20
13- 24	.11	-.23	-.11	.02	-.14	.00	-.04	-.13	.10	-.06	.01	.05
25- 36	-.05	-.05	.03	-.04	-.04	.03	-.05	-.01	-.04	.14	.01	-.05

AUTOCORRELATION FUNCTION

DATA - AGG   96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1) 1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES =-.84167E+01
ST. DEV. OF SERIES = .10265E+02
NUMBER OF OBSERVATIONS = 84

1- 12	.65	.47	.35	.34	.40	.36	.40	.40	.41	.38	.29	.21
ST.E.	.11	.15	.17	.17	.18	.19	.20	.21	.22	.23	.23	.24
13- 24	.24	.24	.23	.16	.09	.05	.08	.12	.14	.09	.12	.06
ST.E.	.24	.24	.25	.25	.25	.25	.25	.25	.25	.25	.25	.25
25- 36	.01	-.03	-.12	-.04	.01	-.04	-.07	-.13	-.14	-.15	-.18	-.15
ST.E.	.25	.25	.25	.25	.25	.25	.25	.26	.26	.26	.26	.26

MEAN DIVIDED BY ST. ERROR = .75151E+01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .20923E+03
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

DIFFERENCE 1
MEAN OF THE SERIES = .84337E-01
ST. DEV. OF SERIES = .83712E+01
NUMBER OF OBSERVATIONS = 83

1- 12	-.22	-.14	-.13	-.04	.15	-.10	-.04	.02	.05	.12	-.03	-.16
ST.E.	.11	.11	.12	.12	.12	.12	.12	.12	.12	.12	.12	.12
13- 24	.05	.01	.10	.03	-.06	-.12	-.02	.06	.08	-.14	.10	.05
ST.E.	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13
25- 36	.02	.07	-.26	.04	.19	-.06	.01	-.07	-.01	.11	-.10	-.01
ST.E.	.13	.13	.13	.14	.14	.14	.14	.14	.14	.14	.14	.15

MEAN DIVIDED BY ST. ERROR = .91785E-01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .37389E+02
SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 30 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - AGG

96 OBSERVATIONS

DIFFERENCING - ORIGINAL SERIES IS YOUR DATA DIFFERENCED BY

1) 1 OF ORDER 12

DIFFERENCES BELOW ARE OF ORDER 1

ORIGINAL SERIES
MEAN OF THE SERIES =-.84167E+01
ST. DEV. OF SERIES = .10265E+02
NUMBER OF OBSERVATIONS = 84

1- 12	.65	.08	.04	.14	.20	-.01	.18	.09	.08	.03	-.09	-.10
13- 24	.10	-.05	-.04	-.09	-.12	-.10	.07	.04	.07	-.05	.11	-.08
25- 36	-.01	-.03	-.14	.08	.05	-.20	-.00	-.02	-.10	-.03	-.03	-.00

DIFFERENCE 1
MEAN OF THE SERIES = .84337E-01
ST. DEV. OF SERIES = .83712E+01
NUMBER OF OBSERVATIONS = 83

1- 12	-.22	-.20	-.23	-.19	.02	-.14	-.12	-.06	-.01	.10	.08	-.10
13- 24	.03	-.00	.07	.12	.08	-.11	-.09	-.05	.03	-.12	.05	.03
25- 36	.02	.13	-.10	-.02	.22	-.04	.01	.07	-.05	.04	-.00	-.11

APPENDIX IV

SIX TYPES OF GOODNESS-OF-FIT MEASURES

The definitions of the six types of goodness-of-fit measures employed are:

1) Correlation coefficient (R) $= \dfrac{\sum_{t=1}^{T}(Y_t - \bar{Y}_t)(\hat{Y}_t - \bar{\hat{Y}}_t)}{\left[\sum_{t=1}^{T}(Y_t - \bar{Y}_t)^2 \sum_{t=1}^{T}(\hat{Y}_t - \bar{\hat{Y}}_t)^2\right]^{\frac{1}{2}}}$

2) β_1 = regression coefficient of actual on predicted value

$Y_t = \beta_o + \beta_1 \hat{Y}_t + \varepsilon_t$

if $\beta_o = 0$ and $\beta_1 = 1$, then $\hat{Y}_t$ is equal to Y_t for all t.

3) Mean Error (ME) $= \frac{1}{T} \sum_{t=1}^{T} (\hat{Y}_t - Y_t)$

The mean error can be misleading. Large positive and negative errors offset each other and bias the mean error downward.

4) Mean Absolute Error (MAE) $= \frac{1}{T} \sum_{t=1}^{T} |\hat{Y}_t - Y_t|$

The mean absolute error is not subject to the bias associated with the mean error.

5) Root-Mean-Square Error (RMSE) $= \sqrt{\frac{1}{T} \sum_{t=1}^{T} (\hat{Y}_t - Y_t)^2}$

This measure is more frequently used in the literature. It weights large errors more than the mean absolute error.

6) Theil's Inequality Coefficient (U)

$$= \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}((\hat{Y}_t - Y_{t-1}) - (Y_t - Y_{t-1}))^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t - Y_{t-1})^2}}$$

$$0 < U < \infty$$

For perfect forecasting, i.e., when $\hat{Y}_t = Y_t$ for all periods, $U = 0$.

The Theil inequality coefficient can be decomposed into three parts, each reflecting a different type of error.

$$U(\text{Bias}) = \frac{(\bar{\hat{Y}}-\bar{Y})^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2}$$

$$U(\text{Variation}) = \frac{(S_{\hat{Y}} - S_Y)^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \qquad S_{\hat{Y}} = \sum_{t=1}^{T}(\hat{Y}_t - \bar{\hat{Y}})^2$$

$$S_Y = \sum_{t=1}^{T}(Y_t - \bar{Y})^2$$

$$U(\text{Covariation}) = \frac{2(1 - r)S_{\hat{Y}}S_Y}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2}$$

And

$$U(\text{Bias}) + U(\text{Variation}) + U(\text{Covariation}) = 1$$

where

T = the number of forecast periods,

$\hat{Y}_t$ = the predicted level of the variable at time period t,

Y_t = the actual level of the variable at time period t.

The first part is zero only when the means of actual and predicted variables are equal. Errors that lead to a positive value for this part can be interpreted as a bias or central tendency error. The second part is zero only when standard deviations of actual and predicted variables are equal. A positive value for this part can be interpreted as error due to different variation. The third part is zero only when the correlation coefficient between predicted and actual values is one. Therefore, a positive value for this part can be interpreted as an error due to different covariation.