THE IMPACT OF HIGH INFLATION RATES ON THE DEMAND FOR AIR PASSENGER TRANSPORTATION

Richard L. Vitek

Nawal K. Taneja

R75-6

May 1975

MIT

DEPARTMENT OF AERONAUTICS & ASTRONAUTICS

FLIGHT TRANSPORTATION LABORATORY

Cambridge, Mass. 02139

FLIGHT TRANSPORTATION LABORATORY
DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FTL R-75-6

May 1975

THE IMPACT OF HIGH INFLATION RATES ON THE DEMAND FOR AIR PASSENGER TRANSPORTATION

RICHARD L. VITEK

NAWAL K. TANEJA

THE IMPACT OF HIGH INFLATION RATES ON THE DEMAND FOR AIR PASSENGER TRANSPORTATION

RICHARD L. VITEK

NAWAL K. TANEJA

SUMMARY

The impact of high inflation rates on the demand for domestic air passenger transportation is tested in a demand model using time-series data and linear and non-linear least squares regressions with Revenue Passenger Miles as the dependent variable, and measures of cost, income and inflation as the explanatory variables. The investigation begins with an extensive survey of the past and current air transportation demand models.

The model selected uses linear and non-linear log specifications to account for the secular trend and detrended variables to account for the cyclical variations. These transformations allow determination of the coefficients comparable to delta log models and simultaneously retain the forecasting ability of linear log models. Forecasts are provided to 1990 for both the linear and non-linear secular trends.

Results show that the price is the most stable and significant determinant of demand. Income and the rate of inflation are both significant but are more variable and highly dependent on the type of secular trend and the time period used in the regression. The non-linear secular trend model provided the best overall fit and explained 96% of the variation in demand.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Demand for U.S. Domestic Air Passenger Transportation as measured by revenue passenger miles has shown an average growth rate of 12.1% over the past 34 years. The demand has grown from 1.05 billion RPM in 1940 to 129.5 billion RPM in 1974 (1). While the average growth rate has been high it has not always been consistent. There have been several periods of slower growth which were significant in their impact on the profit and loss columns of many of the major airlines.

An understanding of these important shorter term cyclical effects and of the longer term secular trends is essential to all the principals associated with the air transportation industry. The regulatory agencies, the aircraft manufacturers, the air carriers, and the airport operators may all view the problem from a different aspect but they all have at least one common goal and that is to understand the variation of demand. The use of demand models to gain this knowledge is prevalent throughout the industry, and models do provide valuable information for decisions related to the setting of rates, establishing or modifying routes, and capital investment in aircraft and airports. The airline and airport profits and the quality of service provided to the public is highly dependent on the ability of agencies to produce accurate forecasts. Although considerable effort has been expanded over the past ten years to develop reasonable demand models the results are still not as accurate as desired; the problem is complex and the real world is constantly changing. The development of new and better models will no doubt continue for many years.

The purpose of this study is to determine whether or not high inflation rates are a significant factor in determining demand. While previous studies have taken inflation into account by using various deflators for price and income, they have not treated inflation as a separate variable as is done in the model in this study. To provide a background for this study, chapter 2 includes a discussion of general demand models, the problems associated with least squares regression analysis in the calibration of demand models, and a summary and critique of a number of studies on air passenger demand models which were reviewed in the literature survey. Chapter 3 gives the specifications, defines all the variables and discusses the data sources. Chapter 4 provides an analysis of the results of the various regressions which were run to arrive at the best model; comparison results of the standard log and delta log models; results of a moving time period to test constancy of the coefficients; a test of forecasting ability of the model, and a forecast of demand out to 1990.

CHAPTER 2

SURVEY OF AIR TRANSPORTATION DEMAND MODELS

2.1 The Demand Model

The demand models discussed in this chapter are quantitative rather than qualitative and for the sake of simplicity consist of only one equation which defines the relationship between the endogenous and exogenous variables. While there is no unique or standard classification of demand models, models are generally identified by one or more of the following characteristics: the functional form of the model as defined in the specification; the underlying theory used to establish the relationship between endogenous and exogenous variables; the technique used to calibrate or use the model; and the type of data used in the calibration or operation. The variety of models which are presented in the literature have in general used equations and calibration techniques which were developed in disciplines such as engineering, statistics, and economics. The above model characteristics are discussed in more detail in the following sections.

2.1.1 The Functional Form

The functional form of the model can be additive such as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots\ldots\ldots \beta_n X_n + \varepsilon \qquad (2.1)$$

where Y is the endogenous variable; the β's are the parameters; $X_1, X_2, \ldots\ldots X_n$ are the exogenous variables; and ε is the error term. This equation is a first order linear equation. The functional form can also be additive with higher order terms such as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 + \beta_3 X_1^3 + \beta_4 X_1^4 + \ldots \beta_n X_1^n + \varepsilon \qquad (2.2)$$

or

$$Y = \beta_0 + \beta_1 X_1^2 + \beta_2 X_2^2 + \ldots\ldots \beta_n X_n^2 + \varepsilon \qquad (2.3)$$

These forms are not linear but are in a class which is called intrinsically linear because they can be made linear by a simple transformation of the variable. The next form is multiplicative such as:

$$Y = \beta_0 X_1^{\beta_1} \cdot X_2^{\beta_2} \ldots . X_n^{\beta_n} + \varepsilon \qquad (2.4)$$

This form is also intrinsically linear. The last functional form is a combination of additive and multiplicative such as:

$$Y = \beta_0 + \beta_1 (X_1 X_2 + \beta_2 X_3^2) \ldots\ldots + \varepsilon \qquad (2.5)$$

which is an example of a general class of intrinsically non-linear equations.

The functional forms of the equations can therefore be classified as additive, multiplicative or combinations of additive and multiplicative which gives rise to equations which are linear, intrinsically linear or intrinsically non-linear. All the above model formulations are represented by single equations which are the most prevalent form; however, two or more simultaneous equations are required to define the relationship between variables if two way causality exists. Equations 2.1 through 2.5 are also defined as the structural form of the model since the β's represent the true parameters and ε represents the disturbance or error term which is the sum of all other factors which may influence Y but which are not included in the equation explicitly.

2.1.2 The Theoretical Basis of the Model

Richard E. Quandt (2) points out that demand for travel is commonly viewed as the result of an individuals' rational decision-making which is subject to economic, social and demographic constraints; and that various modes or destinations of travel are regarded as commodities, each with its own price and among which the consumer chooses so as to maximize some index of satisfaction. This viewpoint is very broad and depends on consumer theory, economic theory, utility theory and other related but generally accepted concepts to be used as the basis for the selection of variables in the model. It is important that the theoretical basis of the model be understood so that an evaluation can be made of the causal or accidental relationship of the variables.

2.1.3 Types of Data

The data used to calibrate or run the models is classified in two ways. First, as time series data, which is a sample set of data over a period of time with fixed time intervals; as cross-section data, which is data collected for one specific time period but a selection that is representative of a set of cells or catagories that make up the total population of data; or data which is a combination of both time series and cross-section data. The second classification is even more general and is aggregated or disaggregated data and simply refers to the level to which the data has been summarized.

Government agencies are the primary source of both airline and socio-economic data. The Civil Aeronautics Board (1) provides financial and traffic statistics on all the major airlines in both aggregate form and on major city-pairs. The Department of Commerce and the Department of Labor provide statistics on income, income distribution, population and various other demographic and economic variables which may be desired in a model. The specific data and sources for this study are given in Chapter 3.

2.1.4 Techniques Used in the Solution

The structural form of the model very nicely isolates the real parameters from the error term, however, the modified form of the equation used in the solution with real data cannot make this distinction. The problem is then to use a technique which will separate the error from the true information and provide the most consistent and unbiased estimate of the real parameters. The most common method for both time-series and cross-section data is regression analysis using ordinary, two stage, or non-linear least squares. However, time-series data is also amenable to smoothing and fitting techniques such as ordinary averaging, exponentially weighted averaging, adaptive filtering and spectral analysis. These techniques are used to separate secular, cyclical, and seasonable variations contained within the time-series data. Computer simulation which has been used in other fields to model dynamic systems has had only limited application to demand models; but it is a very flexible tool and will undoubtedly be used more in the future.

2.1.5 Types of Models

A cursory survey of the literature would give the impression that there are as many models as there are individuals to perform studies; however, a more detailed analysis shows this is not true. For example, the following is a partial list of typical models' names: time-series, cross-section, gravity, inter-zonal, land-use, abstract mode, non-linear, econometric, simulation, aggregate and distributed lag. The names do not represent unique models, but rather that characteristic that the author felt was most relevant. As mentioned in the preceding sections the characteristic named may be type of data, the functional form, the technique applied in the solution, or the underlying theory. Most of them are variations of a few basic formulations which are dictated by the type and amount of data that is available.

The majority of the work with demand models has used the multiplicative functional form with log or delta log transformations and regression analysis with ordinary least squares. When time-series data is used it is usually aggregated to a high degree and the major differences in the models relate to the selection of variables to be tested rather than structural form or technique. However, if quarterly or monthly time-series data is used, the models will frequently employ a lagged variable. The lag can be employed in the dependent or independent variable or both and can be a simple step of one or more periods, or according to a specific distribution function.

When using cross-section data various levels of disaggregated data, usually related to city-pairs, are combined with the gravity model or one of its many variations. The gravity model is a special case of the multiplicative functional form, and is structured to resemble the equation for the gravitational attraction between two masses. The assumption is that the demand for travel between two cities is analogous to the gravitational attraction between masses. The land-use, interzonal, and abstract mode models are all variations of the gravity model. The one notably different model is the N.Y. Port Authority model which uses a market research approach to forecast demand. Examples of all of these models will be discussed in Section 2.4, Review of the Literature.

2.2 Model Development

While the models and the results they produce may vary considerably, the procedure to develop the model follows some standardized and straight-forward steps. Step one is to select the explanatory variables based on stated assumptions, the predictability of the variables and the availability of data. The second step is to determine the functional form of the model. Steps one and two together determine the specifications of the model. The third step is to calibrate the model through the use of regression techniques and determine with statistical tests the significance and reliability of the individual variables and the overall goodness-of-fit. Steps 1 through 3 may have to be repeated in an iterative process until the results of step 3 are satisfactory. If the model is to be used for forecasting, step 4 is to test its forecasting ability. This is normally accomplished by calibrating with only part of the historical data and then forecasting with past and known values of the explanatory variables. This forecast can then be compared with actual historical values that were not used in the calibration. Step 5 is to forecast the future by first forecasting the explanatory variables and then using their values in the model to forecast the demand. While the steps are reasonably straight forward, there are many pitfalls which must be avoided along the way and they are discussed more fully in subsequent sections.

2.3 Multiple Regression Using Least Squares

Since most demand models, including the one in this study, use least squares for calibration, the problems associated with least squares will be discussed. The modern computer has made the computational part of the regression analysis using least squares quick and relatively simple, but still leaves the analysis of results up to the user. The ability to process large volumes of data does not guarantee more reliable results. Least squares involves many assumptions that are frequently overlooked or not adequately tested. This is especially true when using time-series data, as discussed by John E. Meyers (3) in his excellent and detailed treatment of least squares methods. A summary of the assumptions and tests used in least squares analysis is given in Appendix A.

Most regression programs in addition to the regression equation will provide the following information that is needed in the analysis of the results: the mean and standard deviation of all the variables; the correlation matrix, which gives the correlation coefficients between all the combinations of the variables; the t-ratio of the coefficients which is a relative measure of the dispersion to the mean value; and the F-test and $\bar{R}^2$, which are measures of the goodness-of-fit. Some of the more complete programs will also provide the Durbin-Watson or Von Neumann statistics, which are used to indicate the degree of autocorrelation of the residuals, and the analysis of variance table which is very useful when dealing with a small number of data points. Many of the programs also provide the probabilities associated with t-ratios and the F-test. The probabilities are computed from t tables on the assumption that the residuals are random and normally distributed, which in many cases is not true. All of the above measures should be carefully reviewed in the analysis of the results to insure that the assumptions inherent in least squares are met and that tests of significance, if made, are valid.

The secular trends and the cyclical variations inherent in economic time-series data frequently invalidates the assumptions made in least squares analysis. If the secular trend is the dominant characteristic with relatively small cyclical variations around the trend, then high multicollinearity between the exogenous variables can be expected. If the cyclical variations, which are serially correlated, are not accounted for by the independent variables, then autocorrelation will be a problem. A series of regressions were run using simulated data to demonstrate these problems and the results are given in Appendix B.

One approach to avoid these problems is to take first differences to eliminate the trend and minimize the serial correlation. Another approach is to use detrended variables which minimizes the multicollinearity but not the serial correlation. The approach used in this study is a combination of both of the above and is discussed in Chapter 3.

2.4 Review of the Literature

2.4.1 General

The literature on demand models is quite extensive. A review was made of demand models in general as well as the specific application of demand models to air passenger transportation. This section will deal only with the specific application to air transportation. However, for more information on the general application of demand models, references which were reviewed are provided for the following areas: evaluation of statistical techniques (4); research in economics and business (5); transport planning (6); demand for electricity (7); demand for natural gas (8); the demand for air freight (9), (10); the demand for air travel in Canada (11); and the demand for air travel in the North Atlantic (12). The main emphasis in reviewing the specific applications was placed on the purpose, the type of model, the type and source of data, and the results obtained in the study. A summary of the factors used in the models and results of the regressions and forecasts are given in Appendix C.

2.4.2 Thesis by H. C. Bartlett (13)

The intent of the study was to develop long-term estimates of price elasticity as well as to reveal the major determinants of demand for air travel. The multiplicative functional form of the model was used with a log transformation. Log RPM per capita was used as the dependent variable and the following five independent variables were used: 1) log of a selected measure of business activity; 2) log of a measure of consumer spending; 3) log of airline average revenue per passenger mile; 4) log time; and 5) log of the disposable income per capita.

Twenty-six regressions were run in Phase One of the study using different measures of business activity and consumer spending with time-series data from 1947 through 1962. The price elasticity was very consistent throughout the series of regressions, it ranged from -1.705 to -2.084; however, the income elasticity was not consistent and ranged from -1.933 to +1.570. Only two of the regressions were considered reasonable by the author for further analysis.

Phase Two of the study selected the best model for additional modification and further tests. The model selected used producers' durable equipment outlays in constant dollars per capita for business activity and personal consumption expenditures for consumer spending. The modifications consisted of lagging the average revenue and disposable income by one year; using quarterly seasonally adjusted data; holding the proportion of coach travel to total travel constant at an average value for the time period considered; and adding an additional variable, the average revenue per passenger mile for Class 1 railroads to measure cross elasticity. The lagged variables did show an improvement of R^2 and also improved the t-ratio for the income elasticity. The use of quarterly data dropped the R^2 from .77 to .08 and was considered a failure. The use of a constant proportion also lowered the R^2. The use of the revenue variable for Class 1 railroads was not meaningful since the coefficient assumed the wrong sign.

A third phase of the study used the same yearly time-series data to test a set of city-pairs. Results were again inconsistent with R^2 ranging from .45 to .81. Bartlett concluded that the major determinants of demand were the average revenue per passenger mile, disposal personal income and outlays for producers' durable goods in that order; only the price elasticity was consistent and significant and the regressions were subject to high intercorrelation.

The major portion, 85%, of the thesis was devoted to a comprehensive analysis of the positions taken and findings of various authors on price elasticity; the regression analysis is only a small part of the total effort. The linear log model with time-series data is not the best choice to determine price elasticity. Most of the variables used have large secular trends relative to their cyclical variations and a high degree of multicollinearity could be expected. Correlation coefficients and D.W. statistics were not provided in the report so only the author's comments provide any clue as to the degree of multicollinearity or autocorrelation problems. A delta log model or the use of detrended variables would have eliminated the multicollinearity problem and possibly produced more consistent results.

2.4.3 Civil Aeronautics Board Study (14)

The purpose of this study was to forecast demand in terms of RPM for the years 1972 through 1981 using historical time-series data for the period 1946 through 1971. The multiplicative functional form of the model was used with delta log transformations. The dependent variable was delta log RPM per capita and the independent variables were: delta log of the fares per mile, delta log of disposable personal income per capita and log of a time trend. Data for 48 states was used from 1946 through 1962 and data for 50 states from 1963 through 1971. Since delta logs were used the absolute values of the data did not appear and the differences from both sets were used in a single regression to produce the following results,

$$\Delta \log \text{RPM/capita} = .0736 - 1.3498\, \Delta \log \text{FPM} + 1.0888\, \Delta \log \text{DPI/capita} - .0395 \log T \qquad (2.6)$$

with an R^2 of .559 and the D.W. statistic of 2.02. All of the coefficients carry the expected signs and are significant except for the trend. The R^2 of .559 is relatively good for a delta log model. Trends and values were projected for the independent variables with a range of values for FPM and forecasts for RPM were computed using the above equations. The forecasts of RPM in billions for 1975 ranged from 151.0 to 168.2 and for 1980 from 210.4 to 306.5.

The choice of a delta log model for forecasting is an interesting one. The more likely choice for forecasting should have been a linear log model which retains the secular trend; since the secular trend alone can account for 85% to 90% of the variation in the demand. The lack of significance of the trend is also to be expected since all the trends were removed in the process of making the delta log transformation of the variables. The D.W. statistic of 2.02 shows that delta log models do minimize autocorrelation while simultaneously removing the multicollinearity due to large secular trends. A demonstration of the forecasting ability of the model using only part of the data would have added considerable weight to the validity of the model.

2.4.4 Douglas Aircraft Co. A Contract Study (15)

This study forecasted demand to 1981 using time-series data from 1946 to 1972 for both Domestic and International Air Travel. The model used the multiplicative form with log transformations. RPM was selected as the dependent variable. The explanatory variables were; yield, trip length, the velocity of money, and an interest rate ratio of long-term rates divided by short-term rates. Both 48 state and 50 state data were combined in the regression with a dummy variable to correct for the difference after 1969. All the coefficients were significant at the 5% level with an R^2 of .9985, a D.W. statistic of 2.11, and an F ratio of 2803. The forecasts for RPM (billions) in 1975 were from 149.1 to 153.4 and in 1980 from 224.5 to 248.0.

The statistics are almost too good to be true. A review of the data indicates that yield, personal consumption expenditures and trip length all have large secular trends which would indicate potential multicollinearity problems, as does the very high R^2. However, no correlation coefficients are given in the study and no mention is made of any multicollinearity problem by the authors. Since the purpose of this study is to forecast, the multicollinearity and autocorrelation are not as critical as they would be in a model to determine elasticities. The assumption must be made that relationships which existed during the calibration time period will also exist in the future, but this is not unreasonable and forecasts could be very good. The report would have been stronger if an actual test of the forecasting ability using partial data had been presented. One of the major problems with using a large number of explanatory variables, especially ones such as velocity of money and interest rates, is the difficulty in predicting these variables which is required in making the forecast.

2.4.5 Philip Verleger - An Article on Demand Models (16)

Several models are discussed in the article; however, the emphasis is on a point-to-point model, which is a modified form of gravity model, that gives greater weight to income distribution rather than average income. This form is sufficiently different from the normal gravity model to warrant some explanation. The standard formulation is given by

$$T_{ij} = \alpha P^{\beta} \frac{M_i M_j}{d_{ij}^2} \qquad (2.7)$$

where T_{ij} is the travel between cities i and j, P is the price of travel, M_i and M_j are the populations of cities i and j and d_{ij} is the distance between the cities. Verleger presents a modified mass function where M_{ij} is a function of time, population and per capita income as given below

$$[M_i(t)M_j(t)]^y = ((\sum_{k=1}^{N} X_i^K(t)\ e^{b_i \bar{Y}_i^K(t)})(\sum_{k=1}^{N} X_j^K(t)\ e^{b_j \bar{Y}_j^K(t)}))^y \qquad (2.8)$$

where X_i^K represents the kth individual or group of individuals residing in i and $e^{b\bar{Y}^K(t)}$ represents their propensity to travel. $\bar{Y}^K$ is the average income for the group and affects the propensity to travel in an exponential fashion which gives greater weight to higher income levels within the population.

The model is non-linear and a special non-linear regression program was designed explicitly for the specification. The model is used to estimate price and income elasticities for 115 city-pairs using cross-section data from 1960 through 1967. Results show income distribution to be a significant factor; however, price elasticities are generally lower and income elasticities generally higher than results obtained in aggregate models.

These differences in elasticities are used by the author to support his argument that aggregate demand models generally do not provide valid elasticities. While this may be true it is interesting to note that less than 25 of the 115 price coefficients are statistically significant at the 5% level. Conclusions based on the total set of price coefficients are open to question. An equally valid assumption would be that price variations were not large enough to be a significant factor in the data set used, and the income coefficient absorbed the effects of other variables not explicitly stated in the equation. Point estimates or single average values for price and income elasticities were not provided in the article.

2.4.6 Boeing Company Paper (17)

This paper proposes and tests the concept of separating each variable into long-term trends and short-term variations; where the short term variations are used in a regression to determine price and income elasticities, and both short-term variations and long-term trends are combined to do forecasting. Quarterly time-series data is used for the period 1956 through 1970. The dependent variable is RPM per capita and the independent variables are current personal income and yield.

The variables were detrended by setting each variable equal to an exponential form of a polynomial in time, t, as shown in equation 2.9.

$$\text{Variable} = \exp(a + bt + ct^2) \qquad (2.9)$$

Both the linear and quadratic forms were tested. After establishing the long-term trend for each variable, the variables were detrended using the long-term trend and a regression was performed using the logs of the detrended variables. The equations and results of the regressions are shown below.

$$\log RPM^*/cap = 2.36 \log INC^*/cap \; -.40 \log Yield^* \qquad (2.10)$$
$R^2 = .64$ * Detrended by K exp (at)

$$\log RPM^*/cap = 2.968 \log INC^*/cap \; -.434 \log Yield^* \qquad (2.11)$$
$R^2 = .94$ * Detrended by K exp ($at + bt^2$)

Forecasts are made by combining the extrapolated long-term trend of RPM with short-term deviations which are determined by the regression coefficients and forecasts of income and yield.

The R^2 of .94 in equation 2.11 shows a very good explanation of the short-term variations in RPM; but tests of the forecasting ability using only part of the historical data showed a tendency for the forecast to drift away from the actual values. The major difficulty is that both models for long-term trend predict either a constant or increasing growth rate with time. This is counter to the normally decreasing growth rate to be expected as an industry matures. While the model fits the time period selected very well, a ten-year moving time period from 1945 to 1970 would have provided a better test of the model. An advantage of this model is that it has only two variables and forecasting requires only predictions of the changes in the variables, not their absolute values.

2.4.7 Air Traffic Forecasting at the N.Y. and N.J. Port Authority (18)

This paper presents several models, however, the main emphasis is on the "Port Authority Model" which is defined as a market research approach. This model divides the air travel market into a large number of travel "cells" for personal and business travel. The personal travel cells are classified by age, occupation, education and income; and the business travel cells are by industry, occupation and income. A typical matrix of personal travel has 134 individual cells showing the population in each cell.

A series of national household surveys conducted over a period of 15 years has provided a bank of data which is supplemented approximately every two years to perform an updated forecast. The main purpose of the survey is to determine whether a person is a "flier" or not, and if a "flier", how many trips are taken per year. Trends and growth rates are established for the number of fliers and number of trips per 1000 fliers in each cell.

Forecasts are made by first estimating population growth in each of the cells using sources prepared by experts in each field and then calculating the expected number of trips with the following equation.

$$\text{Number of trips} = \text{no. of people in popluation} \times \%\text{of fliers} \times \text{trips per 1000 fliers} \tag{2.12}$$

The estimates from each cell are summed and adjusted for elements not covered in the survey to produce totals for the forecasted years.

A weak point of the model, mentioned in the paper, is that it does not include any explicit expression of changes in air fares or other measure of the cost of travel. Because of this weakness other more conventional econometric techniques are used as supplementary forecasting tools. Results are given for both a delta log model and a linear log model. The models used deseasonalized quarterly U.S. domestic passenger miles data over the period 1949 to 1969, and both provided for distributed lags of the income and price variables.

The delta log used deflated national income, deflated average yield per passenger mile and a trend for independent variables and produced a price elasticity of -1.6 and an income elasticity of +1.38.

The linear log model used deflated GNP per capita, deflated average yield per passenger mile, a trend and three quarterly dummy variables as independent variables. Forecasts for billions of passenger miles in 1975 ranged from 140.2 to 168.4 and in 1980 from 206.8 to 334.7.

Although not discussed in this paper, the Port Authority has also performed three in-flight surveys in 1956, 1963 and 1967, each of a year's duration. These surveys are of New York routes (city-pairs) and are performed by airlines serving the New York area. They provide data similar to the market survey plus additional information on transportation modes used in travel to and from the airports (19).

This three-pronged approach of national surveys, in-flight surveys and econometric models is the most comprehensive and at the same time the most reasonable approach presented in any of the studies. Each of the methods provides complementary information which tends to overcome limitations in each of the individual methods. The national survey provides information on the total population, the in-flight survey provides an excellent sample of the local New York area, and the econometric models provide for an explicit expression of price and inflation.

CHAPTER 3

SPECIFICATIONS AND DATA SOURCES

3.1 Specifications

The explanatory variables selected for the model in this study were measures of consumer income, yield and inflation rates. RPM was selected as the dependent variable. Since high inflation rates have occurred infrequently over a long time period, yearly time-series data was selected to cover the time frame from 1940 to 1974.

Income and yield were selected as explanatory variables because most of the prior studies show them to be relatively consistent and significant determinants of demand. Several different variables were used for income and inflation to determine which best explained the variation in demand for the time periods selected.

The major decision was to select the functional form of the model. The ideal model would be one in which the variables are independent, the coefficients are stable and statistically significant, and the forecasting ability would allow the impact of inflation rates on demand to be measured. These attributes are not normally satisfied simultaneously in most models. The log model is normally used for forecasting but because of multicollinearity the coefficients are not stable and statistical tests are questionable. The delta log model generally provides statistically significant coefficients but is not really suitable for forecasting because the secular trend has been removed.

The approach selected in this study was to use a combination of the log, the delta log, and detrended variables in an attempt to provide the best compromise to satisfy both statistical tests and economic reasoning. The model is shown below, where the "*" indicates variables that were detrended by linear or non-linear secular trends.

$$\log RPM = \log YIELD^* + \log INC^* + \Delta \log INFL + TREND \quad (3.1)$$

The log of RPM was used to retain the secular trend for forecasting. The delta log and detrended form were used for the independent variables to remove trend and eliminate the multicollinearity problem. The delta log or first defference of the logs removes both trend and some of the cyclical variation to represent rate of change in the variable, therefore, it was selected for the inflation rate variable, INFL. Detrending also removes the trend but retains the total cyclical variation around the trend and as such should provide a better measure of the mid-range (2-10 yr.) variation in demand. Therefore, detrending was used for the yield and income, INC, variables even though it does not reduce the autocorrelation of the residuals as well as the delta log form.

This approach required a three step process. First, all the variables, except the measures of inflation, were detrended by performing a regression of a trend against the variable as shown in the example below.

$$\log RPM = TREND \quad (3.2)$$

The residuals from each of these regressions, which represent the detrended variables, were stored in files for step two. A second set of regressions were then run using the detrended variables and delta logs of the inflation factors, as shown below; where each of the different measures of income and inflation were tested to select the best measure for each factor.

$$\log RPM^* = \log YIELD^* + \log INC^* + \Delta \log INFL \quad (3.3)$$

After selecting the best measures the final regressions were run using the model in equation (3.1).

Since the TREND variable accounts for the secular trend in the RPM the coefficients of the other variables will adjust to account for the remaining cyclical variations in RPM. This means that the coefficients will be similar to those obtained in equation (3.3) where the trends were all removed, and will therefore provide reasonable measures of elasticity for yield and income. However, since the secular trend in RPM was retained the model also has the same forecasting ability of the standard log model.

To provide a test of some of the implicit assumptions made in selecting this model, regressions were also made for the linear, log, and delta log forms of the model to provide comparison results; and a moving time period was used to test the constancy of the coefficients with time. A dummy variable was also tested to account for the change in 48 state to 50 state data for RPM, but was not retained in the final form of the model. A detailed discussion of individual regressions and results is provided in Chapter 4.

3.2 Data Sources

Most of the data for RPM and yield were obtained from the "Handbook of Airline Statistics" (1) which is compiled by the Civil Aeronautics Board every two years. The 1973 edition contained summary yearly data up through 1972. Later data on RPM was obtained from CAB's "Air Carrier Traffic Statistics" a monthly issue. (21) Data on yield for 1973 and 1974 was obtained from periodicals which quoted CAB sources. The 1974 figure for yield is considered an estimate and is identified as such in the listing in Appendix D. The specific data used is defined as domestic operations for certified route air carriers. The data was summarized for 48 states up through 1969 and for 50 states thereafter. Figures for 48 states and 50 states are both given for 1969 for comparison purposes. Since there is an abrupt change in data in 1969, the data must be made into a single consistent set through the use of dummy variables or by manually adjusting the data. Both methods were tested in the study. In the manual adjustment the 50 state data was adjusted to estimated 48 state data by using a constant percentage shift, which was calculated from the dual 1969 values. The original data and adjusted values are listed in Appendix D.

The measures of income and inflation were obtained from the "Economic Report of the President" (20). This report provides quarterly and yearly historical data and is updated in February of each year. The measures of income were personal income, disposable personal income and personal consumption expenditures. Both the current dollars and constant 1958 dollars were used in the regression for all measures. The constant dollars were not available for personal income so were calculated using the same price deflator that was used for personal consumption expenditures (P.C.E.).

The two measures of inflation were the consumer price index (C.P.I.). and the implicit price deflator for personal consumption expenditures. The first difference of the logarithms of the yearly values which gives the percentage change per year was chosen for both measures to represent inflation rates. The C.P.I. was given for a base year of 1967 and was converted to a base year of 1958 to be consistent with the deflator for P.C.E. The data values are listed in Appendix D.

CHAPTER 4

CALIBRATION AND EVALUATION

4.1 General

This chapter presents the results of seven sets of regressions which were performed during this study. The first four sets were generated during the normal iterative process in developing the model to determine the best set of specific variables, data and functional form. The next two sets were special tests to compare results with standard functional forms, and test the coefficients of the models. The last set was to test the forecasting ability of the linear and non-linear models. Although regression sets 1 and 2 did not produce satisfactory results, they are included in the evaluation to show the problems encountered and as a background to the final solutions in regression sets 3 and 4.

Two of the library programs available on the Dartmouth Time Sharing System "TUCKREG" and "STAT22" were used in the regression analysis. The "STAT22" program allowed a selection for dependent and independent variables after providing the correlation coefficient matrix, which was very convenient for selection of the specific measure of income or inflation. The program "TUCKREG" provided an analysis of variance table and F-ratios which were useful in the analysis of the data. To facilitate processing the data a program was written to assemble and transform up to nine variables and provide any time segment of the total series in observation format. This allowed multiple regressions to be made with a single input file, and reduced the manipulation of data considerably. A complete set of detailed results is provided in Appendix E. In addition to the regression coefficients, the following data is also provided: the correlation matrix, t ratios, R bar squared, $\bar{R}^2$, the standard error of estimate, S_e, the Durbin-Watson statistic, D.W., and the degrees of freedom. Samples of this data will be repeated in the main text as required to illustrate format or compare results.

As mentioned in Chapter 3 the approach required a three-step process of 1) detrending the variables, 2) running a regression on the detrended variables to determine the best measure of income and inflation, and 3) the final regression with the best set of measures. Data will be presented for each step in the process with an analysis of the results. Variable names are simple abbreviations or initials with prefixes and suffixes to indicate some transformation of variable or change in the time period. The names for the basic variables are:

PERI = personal income,

DPI = disposable personal income,

PCE = personal consumption expenditures,

CPI = consumer price index, 1958=100

DID = disposable personal income implicit deflator,

YLD = yield,

RPM = revenue passenger miles.

Prefixes were added such as "L" to indicate the log to the base ten and "DL" to indicate delta log. The "*" is used to indicate variables detrended by a linear trend, and "**" to indicate detrending by a non-linear trend. A numeric suffix was added if some change was made from current to constant dollars or in the time period. For example, LPERI2*, is the log of personal income using constant dollars which was detrended with a linear trend. A complete list of the variables is provided in Appendix E,

Table E.1.

4.2 Regression Set 1

The first set of regressions used time series data from 1940 to 1972, current dollars for income and yield, and all detrended variables including the measure for inflation. Tests were also made with and without a dummy variable. The first step of detrending the variable was accomplished by running a regression of both the log and the linear form of the dependent variable against a linear time trend as the independent variable. A file was generated for each set of three variables such as LPERI, PERI, and TRD. Separate regressions were run on each pair and results obtained as shown below.

	Const.	TRD	Corr. Coeff.	$\bar{R}^2$	S_e	D.W.	D.F.
LPERI	2.020 (138)	.0287 (39.8)	.99	.9789	.0430	.296	33
PERI	-39.79 (-1.19)	25.84 (15.9)	.94	.8815	96.89	.077	33

This format is used throughout Appendix E and is especially useful to compare coefficients as additional variables are added to the regression. Variables names are given at the top of the table, the dependent variable is always on the left. The first row of numbers gives the values of the constant, the coefficient of the independent variable, and the statistical measures. The t-ratios are given in parentheses in the second row below the coefficients.

Results of the detrending are given in Tables E2.1 and E.2.2. Regressions on Yield and RPM were run with and without a dummy variable.

Yield and inflation factors gave about the same results for the log and linear forms. RPM and income factors both gave improved results with the log forms, indicating that the log transformation was the best overall functional form. All the variables except LYLD had $\bar{R}^2$ of greater than .9 indicating large secular trends where the trend alone can account for over 90% of the variation. The dummy variable did improve the $\bar{R}^2$ slightly for LRPM and was also used for step 3.

The results of the regressions using detrended variables, step 2, are given in Tables E.3.1 and E.3.2. LPCE* and LYLD* both had t-ratios greater than 2. LPERI* and LDPI* carried the wrong sign, and LCPI* and LDID* were either not significant or the wrong sign. The correlation coefficients between measures of income and measures of inflation were also quite high, ranging from .70 to .95. Despite these poor results, LPCE* and LCPI* were selected as the best measures of income and inflation and step 3 was carried out with the best results shown below.

$$\underset{}{LRPM} = \underset{(204)}{3.134} \underset{(-4.87)}{-2.084\ LYLD^*} \underset{(1.2)}{+.891\ LPCE^*} \underset{(2.04)}{+1.711\ LCPI^*} \underset{(71.5)}{+.065\ TRD} \underset{(-4.43)}{-.115\ DUM} \qquad (4.1)$$

Details of the step-wise regression are given in Table E.4. LCPI* has the wrong sign and neither LPCE* nor LCPI* had reasonable t-ratios. In the step-wise regression LPCE* was significant with a t-ratio of 12.2 until LCPI* was added. LYLD* and TRD were both significant and stable and the coefficients were very close to the values obtained in step 2.

To gain some insight into the problem graphs were plotted of LRPM with the trend, Fig. 1, and of the detrended form LRPM*, Fig.3. Fig. 1 shows that although the log form of RPM was better than the linear form it was still not a good fit over the total time period from 1940 to 1972. The increase in demand during the war years 1940 to 1945 was far greater than the following period and this caused a considerable distortion in the residuals in those years. Because of this effect and the high correlation between income and inflation several changes were made and they are covered in the next section.

4.3 Regression Set 2

The time period was shortened to the years 1946 to 1972 to eliminate the war years. To minimize the correlation between income and inflation, constant dollars were used for measures of income and the delta log form of inflation was used to represent rate of inflation. Results of the detrending are shown in Table E.5. The $\bar{R}^2$ for all the measures of income were greater than .98, which indicated that very little variation remained in these variables after the trend was removed. This is not surprising since the values for income and inflation were highly correlated and deflating the income was almost the equivalent of dividing the variable by itself.

The results of the regressions on the detrended variables are shown in Table E.6. In this set LPERI2* has the largest correlation coefficient rather than LPCE2*. The deletion of the 1940-45 data reduced the standard error of estimate for LRPM by a factor of two. When LPERI2* and DLCPI were used without yield the signs were correct but they only explained 6% of the variation in LRPM2*. LYLD2* explained 49% of the variation in LRPM2* all by itself, and when added to the regression with the other variables the sign of LPERI2* became negative.

The final regressions are shown in Table E.7, and the best solution is given below. Note that LYLD2*

$$\underset{}{LRPM2} = \underset{(231)}{3.730} - \underset{(-5.72)}{1.951}\ LYLD2^* - \underset{(-1.1)}{.707}\ LPERI2^* \ \underset{(-1.92)}{-1.077}\ DLCPI + \underset{(63.0)}{.051}\ TRD2 \quad (4.2)$$

and TRD2 are significant but LPERI2* and DLCPI are not and, in fact, LPERI2* has the wrong sign. The use of a deflation factor on income apparently overcompensated for inflation and caused the reversal of sign on the income variable. The use of the delta log form for inflation did decrease the correlation between income and inflation and was retained for the remaining regressions.

4.4 Regression Set 3

Two changes were made for the third set of regressions; current dollars were used for the measures of income as in set 1; and data was added to all variables for 1973 and 1974. The time period was therefore from 1946 to 1974. The delta log form of inflation was used as in set 2. Results of detrending are shown in Table E.8. The detrended variables are shown graphically in Fig. 10. The correlation coefficients for LYLD3 and TRD3 was .76, and for the income measures and TRD3 the coefficients were 1.0, .99, and .99. It is interesting to note at this point that if the standard log transformations of the independent variables were used in a regression, all the variables including inflation would have correlation coefficients between .76 and 1.0, and severe multicollinearity problems would exist. The use of detrended variables reduces this effect considerably as shown in Table E.10. The largest correlation coefficient between yield and measures of income or inflation is .28. The correlation coefficients between measures of income and inflation range between .51 and .63, which is higher than desired but much lower than the results from a standard log transformation.

Results of the regressions using detrended variables are shown in Table E.9, where the variables were added in a step-wise manner to show the effects of additional variables. The coefficient for yield is very stable varying from only -1.485 to -1.612 as additional variables are added to the regression. This was to be expected since the yield variable has the greatest correlation with LRPM* and very little correlation with the other independent variables. The measures of income had very similar coefficients ranging from .35 to .49 with DLCPI3 and from .37 to .53 with DLDID3. This again was expected since the correlation between pairs of measures of income was .97 or greater. Both measures of inflation also had very similar coefficients. The results show that the selection of any one of the measures of inflation or income used in this study is more a matter of personal preference rather than statistical significance. When all three variables are used they can at best explain 67% of the remaining variation in LRPM3*.

The final regressions of this set with LRPM3 are shown in Table E.11 and the best result is shown below.

$$\underset{}{LRPM3} = \underset{(2.67)}{3.745} - \underset{(-7.1)}{1.588}\ LYLD3^* + \underset{(1.63)}{.491}\ LPERI3^* \underset{(-3.02)}{-1.489}\ DLCPI3 + \underset{(71.6)}{.050}\ TRD3 \qquad (4.3)$$

The statistical measures were: $\bar{R}^2$ of .9946, S_e of .0315, and a D.W. statistic of .962. All the coefficients have the correct sign, and all the variables except LPERI3* have t-ratios greater than 3. However the probability of a given level of confidence cannot be stated with any certainty because the D.W. statistic of .962 indicates a fair degree of autocorrelation and probabilities based on a normal or t-distribution of the residuals may not apply. The low D.W. statistic also indicates a possible misspecification in either the functional form or the absence of one or more important variables. It was this result and a plot of LRPM with trend, Fig. 2, which motivated the change from a linear to a non-linear secular trend and resulted in regression set 4. Fig. 2 shows that the two periods of high inflation, 1946-1948 and 1972-1974, are both well below the straight line secular trend, and this condition could explain why the solution emphasized the inflation variable at the expense of the income variable. A non-linear secular trend was selected as the next step, however, before discussing set 4 some comments on the high $\bar{R}^2$ seem in order.

The $\bar{R}^2$ of .9946 is quite high but it is also somewhat misleading. Although the $\bar{R}^2$ of .9946 means that 99.46% of the variation was explained in the log of RPM, LRPM, it is not a direct measure of the variation in the original variable, RPM. However, the standard error of estimate, S_e, is a measure of the average percentage error which occurs about the regression line and can be directly related to RPM. In this case, the antilog of .0315 is equal to 1.075, which says that the average percentage residual error in RPM was 7.5%. Therefore, while 99.46% of the variation was explained for LRPM only 92.5% of the variation was explained for RPM. Since the values for $\bar{R}^2$ are generally high when the log transformation is used for the variables, the standard error should always be used as a check on the percentage variation in the original variable.

Using the antilog of the standard error, S_e, as a measure of the average percentage error, the amount that each variable contributes to the explanation of the variation in LRPM3 can be calculated from Table E.11. The percentage contribution for each of the variables is as follows: trend, 86.7%; LYLD3*, 4.8%; LPERI3*, 0%; and DLCPI, 1% for a total of 92.5%. This shows that yield, income and inflation accounted for about half of the residual variation after detrending. The detrended variable LRPM* and the residuals after the full regression are shown in Fig. 5.

4.5 Regression Set 4

A non-linear least squares regression was used to fit an exponential curve to LRPM for both the 1940 to 1974 and the 1946 to 1974 time periods. The general equation is given below. Where t is a time trend from 1 to 29

$$\text{LRPM} = a + b\,[1-\exp(-ct)] \tag{4.4}$$

or 1 to 35 depending on the time period and a,b and c are the parameters to be determined by the regression. Results are shown graphically in Figures 1 and 2 and in equation form in Table E.12. A comparison of the residuals from linear and non-linear detrending are shown in Figures 3 and 4.

The regression for LRPM from 1940 to 1974 was performed as a test to see how well the non-linear secular trend would fit the war year period 1940 - 1945. The idea was that if the distortions due to the linear trend could be avoided, it might be possible to rerun the regressions for the total time period. The solution did provide a reasonably good fit for 1940 through 1945; however, because of the high rate of change in the early years the projection for a growth rate in 1980 was only 5.3%. This low growth rate was not considered reasonable and no further work was performed with this solution.

The predicted values for the non-linear trend, NLRPM3, for the time period 1946 through 1974 were used to detrend LRPM. The linear trend was used to detrend yield and income. The rationale for this decision was that the trends for yield and income could continue at uniform rates virtually indefinitely, but the demand should show the typical exponential decline in growth rate which is normally encountered as an industry matures.

Results of the regressions with the detrended variables are shown in Table E.13. The $\bar{R}^2$ of .86 shows that 86% of the variation in LRPM3** was explained by the best regression as compared to 67% in the linear case. Again, the coefficients of yield, income and inflation were very stable and all of the measures of income and inflation were equally good.

The final regressions for set 4 are shown in Table E.14 and the best result is shown below. The statistical measures were: $\bar{R}^2$ of .9983, S_e of

$$\underset{}{LRPM3} = \underset{(.07)}{2.29E\text{-}3} \; \underset{(-12.3)}{-1.512\ LYLD3^*} \; \underset{(5.78)}{+.959\ LPERI3^*} \; \underset{(-1.85)}{-.501\ DLCPI3} \; \underset{(131)}{+\ 1.001\ NLRPM3} \quad (4.5)$$

.0173 and a D.W. statistic of 1.309. All the coefficients have the correct sign, but the income variable now has the larger coefficient and t-ratio compared to the inflation factor which is just the reverse of the linear case. The S_e of .0173 equates to an average percentage residual error in LRPM3 of 4.1% or about one-half the value for the linear case. LRPM3** and the residuals after the full regression are shown graphically in Fig. 6.

The contribution of the explained part of the variation in LRPM is as follows: NLRPM3, 88.8%; LYLD3*, 4.9%; LPERI3*, 2%; and DLCPI3, .2% for a total of 95.9%.

The results of regression sets 3 and 4 are somewhat contradictory for income and inflation but both sets show yield to be a consistent and significant determinant of demand. The yield coefficient of -1.588 for the linear case and -1.512 for the non-linear case are in good agreement with other studies using time-series data. The values of the price coefficients obtained in the studies reviewed ranged from -1.35 to -1.8 with an average value of -1.55. The income coefficient of .491 and .959 for the linear and non-linear cases are both low compared to the other studies which ranged form .8 to 1.93 with an average of 1.3. It should be pointed out that the values obtained from the other studies are in themselves averages of selected sets of data. For example, in Bartlett's thesis (13) he ran 26 regressions where the income coefficient ranged from -1.993 to +1.570 and the price coefficient ranged from -1.671 to -2.061. However, all the models with negative income coefficients were rejected as incorrect and were not included in the final summary. In general models using time-series data show the price coefficient to be the most stable. The income coefficient is on the other hand extremely variable and highly susceptible to the time period and other factors used in the model. The variation in the coefficients for income and inflation in this study was not surprising since their correlation with the variation in demand, LRPM3*, was .29 or less in both models. The correlation of yield with the variation in demand was however much higher with values of -.77 for the linear case and -.83 for the non-linear case.

4.6 Regression Set 5

Several sets of regressions were performed to test the stability of the coefficients with a moving time period for both the linear and non-linear secular trend models. Moving time periods of 9 and 14 points were used for the early tests with 1946 - 1972 data of 27 points and 15 and 20 points were used for the 1946 - 1974 data of 29 points. For all except one set the variables were detrended using the secular trend for the total time period. Regressions were then performed using subsets of the detrended variables, for example, data points 1 through 9 of all the variables, then data points 10 through 18, and so forth. For the 15-year moving time period, using the linear model the data was segregated into subsets and then detrended. This method treats each subset as an independent group and is a better test of the moving time period; however, it also requires nine times as many regressions to perform a complete set. Since the results proved to be very similar for both methods, the second method was not repeated for the rest of the sets. Detailed results of the regressions are shown in Table E.15 through E.19, a summary of the results is given below.

Points Used	LYLD*		LPERI*		DLCPI	
			RANGE OF COEFFICIENTS			
Linear-single Trend						
1-9,10-18,19-27	- .21	-1.76	.38	3.78	.46	1.57
1-14,6-19,11-24,14-27	- .863	-1.90	1.06	2.33	- .95	1.94
1-20,5-24,10-29	-1.41	-1.61	.32	2.0	-1.05	.737
Linear-Separate Detrending						
1-15,8-22,15-29	-1.32	-1.71	.40	1.59	- .53	1.39
Non-Linear-Single Trend						
1-15,8-22,15-29	-1.35	-1.62	.98	1.26	- .70	1.18
1-20,5-24,10-29	-1.25	-1.58	.91	1.55	- .82	.43

The nine-year moving time period produced the poorest results and is really too short a period since with the number of variables used there were only 4 degrees of freedom. For the data set used in the regressions the 14- and 15-year periods are also open to question. Fig. 4 shows that the variation in the detrended demand is almost sinusoidal and a complete cycle of variation runs from 12 to 14 years. If a trend is fitted to a single sine wave it will tend to adjust along a line between the two peaks rather than through the average or zero values. The results will be quite different than if the trend were fitted to two or more cycles of the variation. The 15-year moving period is equivalent to fitting a single sine wave and the results could be biased.

The yield and income coefficients always had the correct sign but varied over a wide range. The coefficient for inflation had both plus and negative values indicating that it was the least stable of the coefficients. A review of Fig. 10 will show that high rates of inflation occurred in only two periods, around 1946 and 1974, and were also relatively short in duration. The average effect of inflation in any 15 or 20-year period would, therefore, not be significant. This does not mean that high inflation rates are not a significant factor in determining demand in the years in which they occur, it simply means that there is insufficient data to provide a statistical determination of their significance for the overall time period. The best results in terms of the least variation in the coefficients of yield and income were obtained with the non-linear model. The average values of the coefficients of -1.45 and 1.175 for yield and income compare very well with the values of -1.512 and .959 for the overall 29 point time period.

4.7 Regression Set 6

This series of regressions were performed to compare the results of the detrended model with results of the standard log and delta log model. The total time period, 1946-1974, is used in all of the regressions in this set. The first set of regressions used a log transformation for all variables except inflation, which was retained as delta log, so that the results could be compared directly with the linear detrended model. Results are shown in Tables E.20.1 and E.20.2. Step wise addition of variables was used to see the effect on the coefficients. With LRPM3 as the dependent variable and LYLD3 as the only independent variable the coefficient for LYLD3 was +7.13. The secular trend within LYLD3 dominated the solution and produced a very large coefficient of the wrong sign. When income, LPERI3, was added to the solution the coefficients for yield and income were -1.542 and 2.030, respectively. There was a sufficiently large trend within LPERI3 to allow the coefficient of the yield to return to its normal value. The addition of the inflation factor did not change the coefficients of yield and income since it did not contain any trend. The final addition of a trend, TRD3, resulted in coefficients which are almost identical to the detrended linear model and both are shown below for comparison purposes. The only differences are in the

$$\underset{}{LRPM3} = \underset{(5.78)}{3.782} - \underset{(-7.1)}{1.588}\, LYLD3 + \underset{(1.63)}{.491}\, LPERI3 - \underset{(-3.0)}{1.489}\, DLCPI3 + \underset{(5.3)}{.043}\, TRD3 \qquad (4.6)$$

$$LRPM3 = \underset{(2.67)}{3.745} - \underset{(-7.1)}{1.588}\, LYLD3^{*} + \underset{(1.63)}{.491}\, LPERI3^{*} - \underset{(-3.0)}{1.489}\, DLCPI3 + \underset{(71.6)}{.050}\, TRD3 \qquad (4.7)$$

constant and the coefficient for the trend. The main difference is in the correlation coefficients of the independent variables. In the log model the correlation coefficients between yield, income and trend range from .77 to .99; but in the detrended log model they range from .00 to .56. The detrended model reduces the multicollinearity considerably.

The results above also lend support to the results obtained with the simulated data in Appendix B. That is, when the log model is used with variables which have large secular trends compared to the cyclical variations, the correlation coefficients and the regression coefficients are both dominated by the trends and results are suspect and difficult to interpret. However, if a trend variable is added to the set of variables the regression coefficients are then dependent on the cyclical variations and not the secular trends within the variables. Unfortunately, there is no way to prove this without running a detrended log model. Regressions using all detrended variables provide correlation coefficients which are directly related to the cyclical variations in the variables and are very helpful in determining the most effective variables to use in the final solution.

The second set of regressions were run using the delta log form and results are given in Table E.21. Neither income nor inflation were significant and the inflation coefficient had a positive sign. The coefficient for yield ranged between -1.38 and -1.56 and was very close to the values obtained in the log and detrended log models. The Durbin-Watson statistic as expected was the best of any of the models and ranged between

1.937 and 2.085. The correlation coefficients were all very low except between income and inflation where it was .51, about the same as in the detrended model.

The last set of regressions were not performed with a standard model but used the log of RPM and delta logs of the independent variables with a trend. This was simply a test to see if the delta log transformations correlated at all with the variations in LRPM. Results are given in Table E.22. None of the variables had t-ratios greater than 1.6, even though the $\bar{R}^2$ was .98. The trend accounted for virtually all of the explained variation. While the delta log model may provide reasonable estimates for elasticities it does not appear useful to actually forecast changes in demand as was attempted in the C.A.B. study (14).

4.8 Regression Set 7

The last set of regressions were performed to test the forecasting ability of the models. Regressions were run for the time period 1946 - 1969 for both the linear and non-linear models. Results are shown in Table E.23, the results of the full-time period 1946-1974 are also shown in the same table for comparison purposes. The regression equation was then used with the known data values for the variables and predictions were made for years 1970 through 1974. The calculated values are shown in Table E.24. The difference between actual and predicted values was converted to a percentage and the average percentage error was calculated for both models. The linear model gave an average error of 21.4% and the non-linear model gave an average error of 7.9%. The results are shown graphically in Fig. 11. Both models gave results that were higher than the actual data, which is not surprising considering the large decrease in the rate of growth in demand in the 1970 to 1974 time period.

These results also indicate that the rate of inflation which was relatively high during this period has a greater impact than that shown by the overall regression. Although there is no statistical evidence to support a higher coefficient for the rate of inflation in the non-linear model, it is not unreasonable to assume that the coefficient is biased downward because of the relatively low values of inflation over the whole time period; and that the actual coefficient should be closer to -1.0, which is the average of the -.5 of the non-linear model and the -1.5 of the linear model. Then the decrease in demand during the 1970 to 1974 time period would be explained very well by a real decrease in the rate of growth as suggested by the non-linear secular trend plus the impact of higher than normal inflation rates.

4.9 Forecasts

Forecasts were made using both the linear and non-linear models as shown below in 5 year increments out to 1990. Forecasting for even short

$$LRPM3 = 3.745 - 1.588\ LYLD3^{*} + .491\ LPERI3^{*} - 1.489\ DLCPI3 + .050\ TRD3 \qquad (4.8)$$

$$LRPM3 = 2.29E{-}3 - 1.512\ LYLD3^{*} + .959\ LPERI3^{*} - .501\ DLCPI3 + 1.001\ NLRPM3 \qquad (4.9)$$

periods is subject to significant errors, the 15 year period chosen here does not imply any special accuracy but was selected simply to show the rapid departure between the forecasts of the linear and non-linear models. Forecasting with a model which uses detrended variables requires predictions for only changes in the variables rather than the absolute values; however, to further simplify the process only the limits of the changes were predicted.

These limits were then added to the secular trend to give upper and lower bounds for the demand.

To determine the limits on the variables, graphs were constructed for the original variables yield, income and inflation, see figures 7,8 and 9, and for the detrended variables, see Fig. 10. The limits for each variable were then multiplied by the regression coefficients in equations 4.8 and 4.9 and summed to produce the upper and lower bounds as shown below. The value

	LYLD	LPERI	DLCPI	LIMITS	
Change	∓.035	∓.05	+.025		
Change x Coef					
Linear	∓.05	∓.03	-.04	-.12	=.04
Non-Linear	∓.05	∓.05	-.01	-.11	+.09

of -.12 represents a decrease of 33% and the value of +.09 represents an increase of 23% relative to the secular trend. These values are very close to the actual historical changes in demand which are shown in Fig. 4.

The secular trends were predicted for both the linear and non-linear models. The non-linear trend, NLRPM3, was predicted using the equation which resulted from the non-linear regression as shown below.

$$NLRPM3 = 3.6499 + 3.6618 \left[1 - \exp(-.0177\, t) \right] \qquad (4.10)$$

The linear trend, TRD3, was a simple extension of the time trend values, such as 30, 35, and so forth. These values for the trends were used with the trend coefficients and the constants in the regression equation 4.8 and 4.9 to calculate average values of the demand. The limits calculated above were then added to the average value to produce the upper and lower bounds. The results are shown in tabular form in Table E.25 and in graphic form in Fig. 12. The values for LRPM were converted back to RPM and a fixed percentage difference was then used to provide estimates of 50 state data. These conversions plus the forecasts for 1980 from other studies are also given in Table E.25.

The results show that by 1985 the forecasts for the linear and non-linear models no longer overlap and by 1990 the linear model forecasts are approximately twice the value of the non-linear model forecasts. It is difficult at this point in time to decide which model is better, but in another 5 to 10 years the choice should be obvious.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

While this preliminary investigation did not provide the definitive answer to the impact of inflation on the demand for air transportation, it did prove that: 1) high inflation rates are significant and inflation should be taken into account explicitly as a separate factor in the demand model; 2) regressions using all detrended variables provide more meaningful correlation coefficients and facilitate the selection of the best measures of the independent variables; and 3) the non-linear secular trend is not only a reasonable choice but a very effective one.

The results did not produce a single coefficient with a given level of confidence for the inflation factor, but they did show that the value of the coefficient should be between -.5 and -1.5. Using the average of -1.0 this indicates that everything else being equal a <u>10% rate of inflation could cause a 10% decrease in demand</u>. Needless to say a 10% decrease in demand is significant and the possibility of high inflation rates must be considered when forecasting for the future.

The use of the non-linear secular trend provided the best overall fit to the historical data with an average residual error in RPM of only 4.1%, the equation is shown below. The non-linear model also gave the best results

$$LRPM = 0.00229 - 1.512\ LYLD^* + .959\ LPERI^* - .501\ DLCPI + 1.001\ NLPM \qquad (5.1)$$

in the forecast test for the 1970-1974 time period with an average error of only 7.9% as compared to an average error of 21.4% for the linear model. The gradually decreasing growth rates predicted by the non-linear trend are also intuitively more satisfying since the past growth rates of 10% to 12% cannot be expected to continue indefinitely.

The combination of detrended variables and the non-linear secular trend proved to be very effective in explaining the variations in demand. The results from the moving time period test of 15 and 20 years agreed well with the overall period of 29 years and also tend to support the choice of this type of model. The results were equal to or better than the standard log or delta log models with the predictive ability of the log model and the stability of coefficients as in the delta log model.

Future efforts with this type of model could look for potential improvements in two areas, a reduction in the autocorrelation of the residuals and reduction in the correlation between the income and the inflation factor. The relatively low value of 1.3 for the Durbin-Watson statistic indicates that cyclical variations still remain. The use of measures of price and income different than the ones already tested or the use of additional variables might explain more of the remaining variation and bring the Durbin-Watson statistic up to a more respectable value of 1.8 to 1.9. Past studies have shown that time-series data provides reasonable price coefficients while cross-section data provides better income coefficients. If cross-section data were used to determine the income coefficient, which could then be employed in the time-series analysis, correlation between income and inflation would be greatly reduced.

APPENDIX A

Assumptions and Tests Used in Least Squares Analysis

The general form of the equation with several variables is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots \beta_n X_n + \varepsilon \qquad (A.1)$$

where Y is the dependent variable, $X_1 \ldots X_n$ are the independent variables, β_0 is a constant term, $\beta_1 \ldots\ldots \beta_n$ are the parameters, or true values relating the independent variables to the dependent variables, and ε is the disturbance term which is the sum of all other factors not explicitly included in the equation.

The regression equation is given by

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \ldots b_n X_n \qquad (A.2)$$

where $\hat{Y}$ is the predicted or estimated value of the dependent variable, b_0 is a constant, and b_1, b_2 ... b_n are the statistics or estimates of the parameters The least squares technique minimizes the sum of the squares of the residuals, denoted as

$$\min \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \qquad (A.3)$$

over the range of values from i = 1 to n. The characteristics of the residuals are very important in determining the validity of the statistical estimates of the parameters. The following assumptions must be true to obtain consistent and unbiased estimates of the parameters.

1. The expected value or mean of the residuals must be zero.
2. The variance of the residuals must be constant and independent of the value of X_j for all j's. Departure from this assumption is called heteroscedasticity.
3. All pairs of values of the residuals whether adjacent or not must not be correlated. A departure from this assumption gives rise to autocorrelation.

4. The probability density function of the residuals is assumed to be a normal distribution.

5. The exogenous variables must be independent, that is, not correlated with one another. A departure from this assumption is multicollinearity.

6. There must only be one-way causality from the independent variables to the dependent variable.

Various tests can be performed to measure the degree to which the assumptions have been met, the statistical significance of the coefficients of the independent variables and the overall goodness-of-fit of the equation.

Autocorrelation of the residuals is tested by computing the Durbin-Watson statistic. This measure (D.W.) can assume values from zero to 4 and if no autocorrelation is present will have an expected value of two. Serial or positive correlation will result in values less than 2 and negative correlation in values greater than 2.

Multicollinearity is measured by testing the correlation coefficients of the independent variables to determine if the degree of correlation is significant.

Heteroscedasticity cannot be measured by a simple test during the regression computation. If multiple values of the X_j's are available for each Y_j then the variance can be computed over the range of each X_j and tested for constancy. If multiple values are not available, a simple but effective test is to plot the residuals and the regression line and perform a visual check of the constancy.

The t-ratios can be computed as the ratio of the value of the coefficient, b, divided by the standard error (σ) of the coefficient. The t-ratio permits a test of the hypothesis that a given independent variable has no effect, based on the set of sample observations used for the regression analysis.

The coefficient of multiple correlation, R, its square, R^2, and the R^2 adjusted for the degrees of freedom, all show the degree of association of the dependent variable with the entire set of independent variables. For example, an R^2 adjusted, also called R bar squared, $\bar{R}^2$, of .97 indicates that 97% of the variation in Y around the mean value of Y has been "explained" by the equation.

The F-test is another significance test of the entire regression and is particularly valuable where $\bar{R}^2$ is low. The F-test permits a test of the hypothesis that all the coefficients are simultaneously equal to zero.

APPENDIX B

Regression Analysis Using Simulated Data

Although regression analysis using least squares is a widely accepted tool, well founded in statistical and mathematical theory, the results obtained using least squares are misunderstood, misused, and often the subject of considerable controversy. In order to demonstrate the least squares process, the use of standard tests, and point out some of the pitfalls, a series of regressions were run using simulated data. The data was selected to represent the secular trends and cyclical and seasonable variations normally found in time-series data, such as RPM, yield, income, and aircraft speed. Sine waves of different periods were selected to represent the cyclical variations because they are orthogonal and would not be correlated. Time trends were used to simulate the secular trend, and random noise was used to simulate residual factors not explicitly stated in the equation.

For the regressions the dependent and independent variables were generated using various combinations of the following components each containing 38 or 48 points:

1) a time trend, TT1, numbers 1 through 48,

2) a time trend, TT2, numbers .1 to .48,

3) a random noise component, RN, numbers from $\pm$ .5 or $\pm$ .3,

4) a one-cycle sine wave, SW1, where the value was equal to 1.sin θ and the total period was 360 degrees,

5) a two-cycle sine wave, SW2, with a total time period of 720 degrees,

6) a three-cycle sine wave, SW3, with the total time period of 1080 degrees.

The actual data is shown in Table B.1

A computer program was written to accept all of the above components, multiply each by constants selected at run time, and produce an observation format suitable for use by "TUCKREG" a library regression program on the Dartmouth Time Sharing System. The following table illustrates the process and will be used later to explain some test runs.

	TT1	TT2	RN2	SW1	SW2	SW3	Cont.
Y =	1	0	1	0	0	0	0
X1 =	0	1	0	2	0	0	0
X2 =	0	1	0	0	3	0	0
X3 =	0	1	0	0	0	4	0

The table shows that the dependent variable, Y, was simply the addition of 1·TT1 plus 1·RN2, or

$$Y_i = 1 \cdot TT1_i + 1 \cdot RN2_i \quad (B.1)$$

for all i from 1 to 48. In a similar manner

$$X1_i = 1 \cdot TT2_i + 2 \cdot SW1_i \quad (B.2)$$

$$X2_i = 1 \cdot TT2_i + 3 \cdot SW2_i \quad (B.3)$$

$$X3_i = 1 \cdot TT2_i + 4 \cdot SW3_i \quad (B.4)$$

This technique allowed the generation of variables composed of any combination of the six functions plus a constant. By running regressions of X1, X2 and X3 individually or together on Y and by varying the multiplilers it was possible to measure the effect that multicollinearity and autocorrelation had on the value and stability of the coefficients as well as the values and variation of the standard statistical measures. Output from the regression program "TUCKREG" included the F-test; Durbin-Watson statistic, D.W.; t-ratios; standard error of estimate, S_e, $\bar{R}^2$; correlation coefficient matrix, and other items which were used to evaluate the results.

Test #1 used TT1 plus RN1 for the dependent variable and TT2 plus different constants as shown in the table below.

	TT1	TT2	RN1	SW1	SW2	SW3	Cont.
Y =	1	0	5	0	0	0	0
X1 =	0	1	0	0	0	0	0
X2 =	0	1	0	0	0	0	10
X3 =	0	1	0	0	0	0	30

Results are shown below. The constant and coefficients are given in the first row and the t-ratios are given in parenthesis in the second row.

Reg. $\hat{Y}$ =	b_0	+ b_1 X	$\bar{R}^2$	D.W.	F	S_e
Y , X1	-0.111	10.045 X1	.9895	2.973	4414.	1.45
	(-0.26)	(66.4)				
Y , X2	-100.6	10.045 X2	.9895	2.972	4414	1.45
	(-53.0)	(66.4)				
Y , X3	-301.5	10.045 X3	.9894	2.963	4400	1.45
	(-61.3)	(66.3)				

Most of the results were as expected except for the Durbin-Watson statistic. Since TT2 is 10 times larger than TT1, b_1 should have been 10 and was 10.045. Since RN1 is small compared to TT1 one would expect $\bar{R}^2$ to be close to 1 and its value was .9895. The t-ratio for b_0 in the first equation indicates that the constant is not significantly different from zero and that is correct, it should have been zero. The constants added to X2 and X3 did not effect b_1 but were simply absorbed by the constant term b_0 in the regression equation. The one surprise was the D.W. of 2.97. The numbers from the random number table were obviously negatively correlated. To correct this problem a second set of random numbers, RN2, were generated, which by themselves produced a D.W. of 1.98. This second set, RN2, was used in tests #3 and #4 so that any significant deviations from 2 would be due to effects other than the random noise component.

Tests #2 used sine waves and noise to demonstrate the stability of the coefficients with truly independent exogenous variables as well as the effect of cyclical residuals on the Durbin-Watson statistic. The composite variables are shown in the table.

		TT1	TT2	RN1	SW1	SW2	SW3
Y	=	0	0	10	10	10	10
X1	=	0	0	0	10	0	0
X2	=	0	0	0	0	10	0
X3	=	0	0	0	0	0	10

The results are shown below.

Reg. $\hat{Y}$	b_0	$+ b_1X_1$	$+ b_2X_2$	$+ b_3X_3$	$\bar{R}^2$	D.W.	F	S_e
Y,X1	0	.992 X1			.302	.316	21	10.5
=	(0)	(4.6)						
Y,X1&X2	0	.992 X1	1.005 X2		.625	.553	40.2	7.7
	(0)	(6.3)	(6.4)					
Y,X1&X2&X3	0	.992 X1	1.005 X2	.997X3	.994	2.97	268	2.96
	(0)	(16.4)	(16.6)	(16.1)				

Correlation Coefficients

	X1	X2	X3	Y
X1	1	0	0	.56
X2	0	1	0	.57
X3	0	0	1	.55
Y	.56	.57	.55	1

Note that there were three regressions and a new variable was added each time. The correlation matrix shows that there is no correlation between any of the independent variables. Since the same multiplier was used for the sine wave components in Y and in the X's the coefficients of each variable should have been 1 and they ranged from .997 to 1.005. The coefficient of X1 remained the same throughout all three regressions which shows that it was not biased by the exclusion of the other variables in this special case. However, that would not always be true and with only equation number 1 using Y and X1 the validity of the coefficient is highly suspect because of the low $\bar{R}^2$ of .302 and the low D.W. of .316. Even without knowing how the data was generated one would speculate that there was a specification error and a look at the plot of residuals in figure B.1 would reenforce this assumption. The residuals are due to SW2, SW3 and RN1 and have a high degree of autocorrelation as indicated by the D.W. statistic. Note that the residuals in figure B.2 are fairly random despite the high D.W. of 2.97. This indicates that the D.W. statistic is very sensitive to even small amounts of autocorrelation; and while it is an excellent flag that potential problems may exist, a careful study of the residuals should be made before rejecting results solely on this one statistic. The t-ratios are all well above 2 which indicates that the coefficients are fairly well defined; however, hypothesis tests using t-tables are not possible because the high autocorrelation generally precludes the possibility of a normal distribution.

Test #3 used trends, random noise and sine waves to simulate typical economic time-series data. The .2 multipliers for the sine wave components of X1 and X2 were selected to make the cyclical variation small compared to the mean value of the trend as is normally the case with economic data.

	TT1	TT2	RN2	SW1	SW2	SW3
Y	1	0	1	0	1	1
X1	0	2	0	0	.2	0
X2	0	2	0	0	0	.2
X3	0	1	0	0	0	0

The results are shown below.

Reg $\hat{Y}$	b_0	b_1X_1	b_2X_2	b_3X_3	$\bar{R}^2$	D.W.	F	S_e
Y,X1&X2	.239 (1.4)	2.058X1 (5.2)	2.922X2 (7.4)		.9978	.38	8210	.525
Y,X1&X2&X3	.053 (.7)	4.95 X1 (17.6)	5.53 X2 (21.0)	-10.98X3 (-12.9)	.9996	1.69	30378	.223

Correlation Matrix

	X1	X2	X3	Y
X1	1	.995	.998	.997
X2	.995	1	.998	.998
X3	.998	.998	1	.997
Y	.997	.998	.997	1

The F value of 8210 and the $\bar{R}^2$ of .9978 look very good but are misleading. The correlation matrix shows that all the independent variables are highly correlated with correlation coefficients greater than .99. The values of b_1 of 2.06 and b_2 of 2.9 are determined almost entirely by the trend components of X1 and X2. The D.W. statistic of .38 shows high positive autocorrelation which is to be expected since half of the cyclical variation in Y still remains, see Figure B.3.

By adding an additional trend variable X3 which can compensate for the trend components of X1 and X2 the coefficients increase by a factor of 2, b_1 of 4.95 and b_2 of 5.53; and are now determined by the cyclical component. Note that with the three independent variables F and $\bar{R}^2$ have increased. The D.W. statistic has also increased to 1.69 because of a reduction in the cyclical residuals, see Figure B.4. The apparently good fit is misleading and the results are not statistically valid because of the high degree of multicollinearity. This problem of multicollinearity has lead to a practice of taking linear or logarithmic first differences to eliminate the trend and do the regression with only the remaining variations. An alternative to this approach is shown in test #4.

Test #4 also uses a trend, noise, and sine waves for the dependent variable as in test #3 but X1 is a trend only and X2 and X3 represent detrended variables, that is, only sine waves. The variable X3 with only one of the two required sine waves is to test the effect of excluding a variable.

	TT1	TT2	RN2	SW1	SW2	SW3
Y	2	0	3	0	10	10
X1	0	2	0	0	0	0
X2	0	0	0	0	3	3
X3	0	0	0	0	0	3

Results are given below.

Reg. $\hat{Y}$	b_0	b_1X_1	b_2X_2	$\bar{R}^2$	D.W.	F	S_e
Y,X1&X3	2.803 (1.25)	9.83X1 (19.8)	2.717X3 (5.3)	.917	.092	205	6.7
Y,X1&X2	.141 (.709)	9.956X1 (229)	3.385X2 (90)	.999	2.15	29109	.587

Correlation Matrix

	X1	X2	Y
X1	1	-.064	.927
X2	-.064	1	.315
Y	.927	.315	1

The process of detrending the variable has decreased the correlation between the independent variables to -.064. The D.W. statistic of .09 in the first regression indicates a problem as expected since SW2 is still a part of the residuals. The second regression, however, meets all the assumptions and has good statistics. The multicollinearity has been removed, the D.W. statistic of 2.15 indicates very low autocorrelation, the $\bar{R}^2$ and F values indicate a very good fit to the dependent variable and hypothesis tests would be valid for testing statistical significance. Since the original variable for Y is used, not the first difference, the results can also be used directly for forecasting.

TABLE B.1

TT1	TT2	RN1	RN2	SW1	SW2	SW3
1	.1	.15	.3	.0000	.0000	.0000
2	.2	-.10	-.1	.1305	.2588	.3827
3	.3	.32	.2	.2588	.5000	.7071
4	.4	-.45	.2	.3827	.7071	.9239
5	.5	-.11	-.1	.5000	.8660	1.0000
6	.6	-.34	.2	.6088	.9659	.9239
7	.7	.31	.1	.7071	1.0000	.7071
8	.8	-.03	.3	.7934	.9659	.3827
9	.9	.23	-.1	.8660	.8660	.0000
10	1.0	-.07	-.3	.9239	.7071	-.3827
11	1.1	.34	.1	.9659	.5000	-.7071
12	1.2	-.04	-.2	.9915	.2588	-.9239
13	1.3	.27	-.2	1.0000	.0000	-1.0000
14	1.4	.11	.1	.9915	-.2588	-.9239
15	1.5	-.42	-.2	.9659	-.5000	-.7071
16	1.6	-.11	-.1	.9239	-.7071	-.3827
17	1.7	.24	-.3	.8660	-.8660	.0000
18	1.8	-.50	.1	.7934	-.9659	.3827
19	1.9	.49	.3	.7071	-1.0000	.7071
20	2.0	-.26	-.1	.6088	-.9659	.9239
21	2.1	.22	.2	.5000	-.8660	1.0000
22	2.2	-.02	.2	.3827	-.7071	.9239
23	2.3	-.47	-.1	.2588	-.5000	.7071
24	2.4	.10	.2	.1305	-.2588	.3827
25	2.5	-.19	.1	0	.0000	.0000
26	2.6	.09	.3	-.1305	.2588	-.3827
27	2.7	-.21	-.1	-.2588	.5000	-.7071
28	2.8	-.19	-.3	-.3827	.7071	.9239
29	2.9	-.34	.1	-.5000	.8660	-1.0000
30	3.0	.33	-.2	-.6088	.9659	-.9239
31	3.1	-.10	-.2	-.7071	1.0000	-.7071
32	3.2	.34	.1	-.7934	.9659	-.3827
33	3.3	-.45	-.2	-.8660	.8660	.0000
34	3.4	.13	-.1	-.9239	.7071	.3827
35	3.5	.42	-.3	-.9659	.5000	.7071
36	3.6	-.09	.1	-.9915	.2588	.9239
37	3.7	.32	.3	-1.0000	.0000	1.0000
38	3.8	.02	-.1	-.9915	-.2588	.9239
39	3.9	-.41	.2	-.9659	-.5000	.7071
40	4.0	.16	.2	-.9239	-.7071	.3827
41	4.1	-.43	-.1	-.8660	-.8660	.0000
42	4.2	.49	.2	-.7934	-.9659	-.3827
43	4.3	-.10	.1	-.7071	-1.0000	-.7071
44	4.4	.23	.3	-.6088	-.9659	-.9239
45	4.5	-.37	-.1	-.5000	-.8660	-1.0000
46	4.6	.41	-.3	-.3827	-.7071	-.9239
47	4.7	-.01	.1	-.2588	-.5000	-.7071
48	4.8	.09	-.2	-.1305	-.2588	-.3827

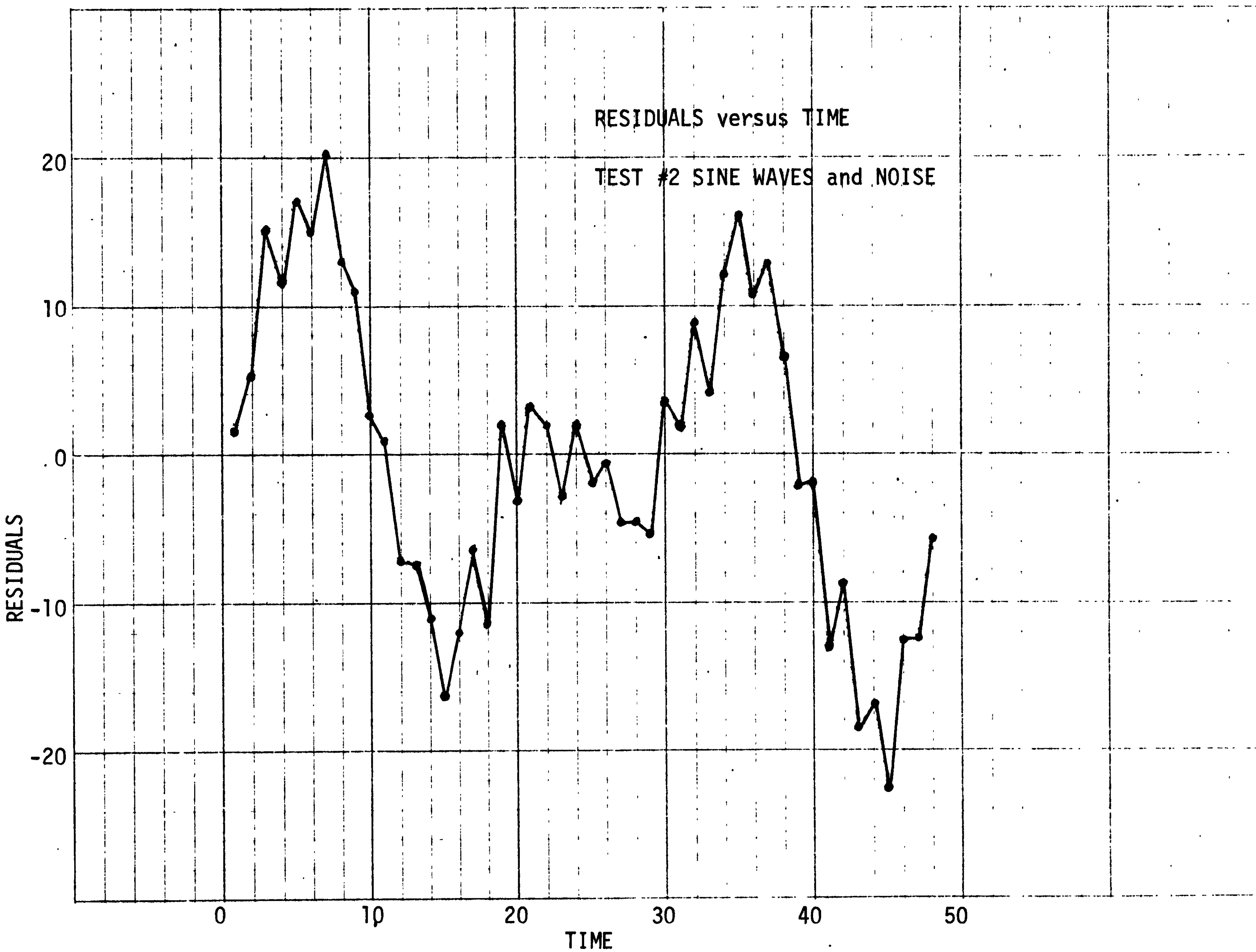

RESIDUALS versus TIME
TEST #2 SINE WAVES and NOISE
RESIDUALS
20
10
0
-10
-20
0
10
20
30
40
50
TIME

FIGURE B.1

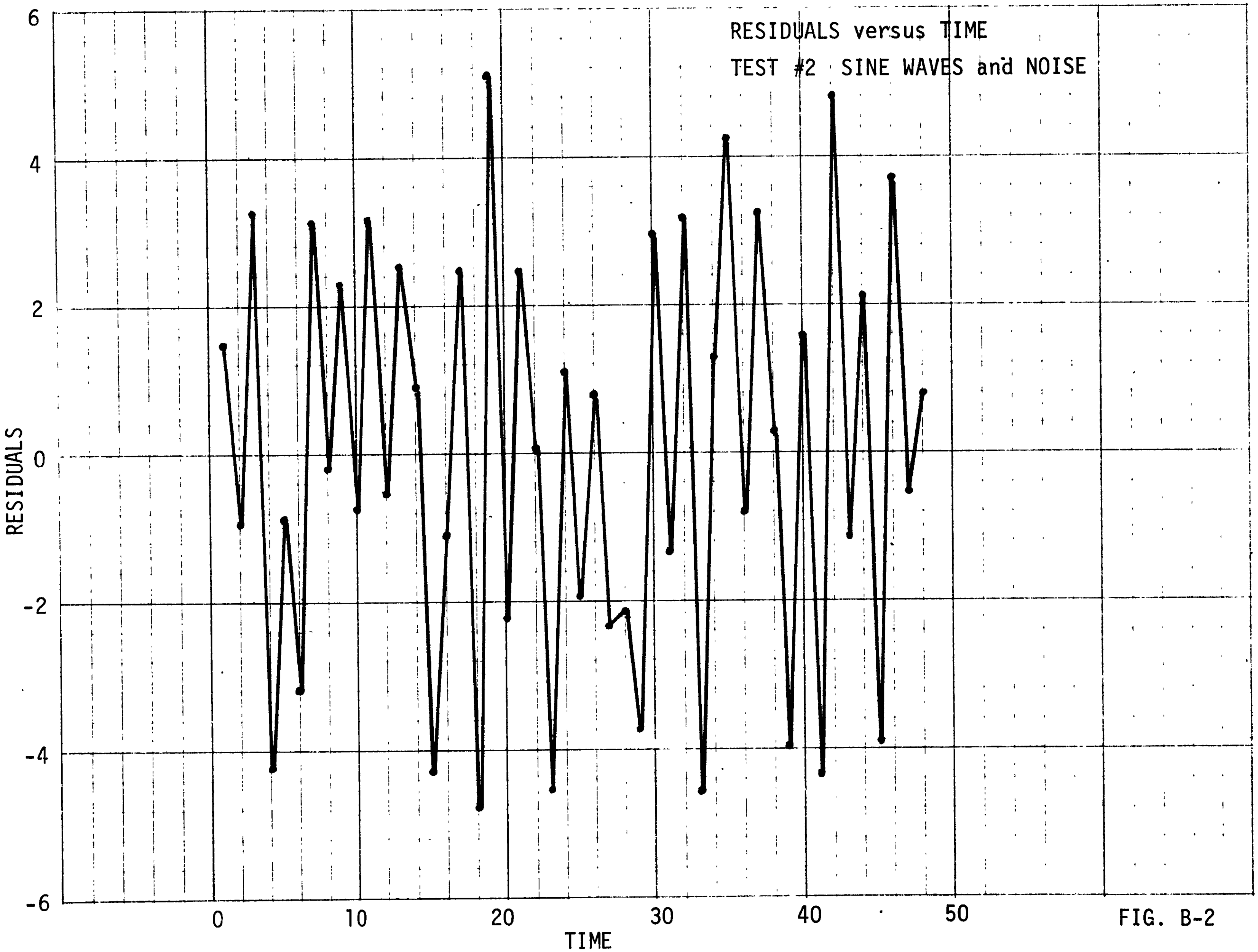

RESIDUALS versus TIME
TEST #2 SINE WAVES and NOISE
RESIDUALS
6
4
2
0
-2
-4
-6
0
10
20
30
40
50
TIME

FIG. B-2

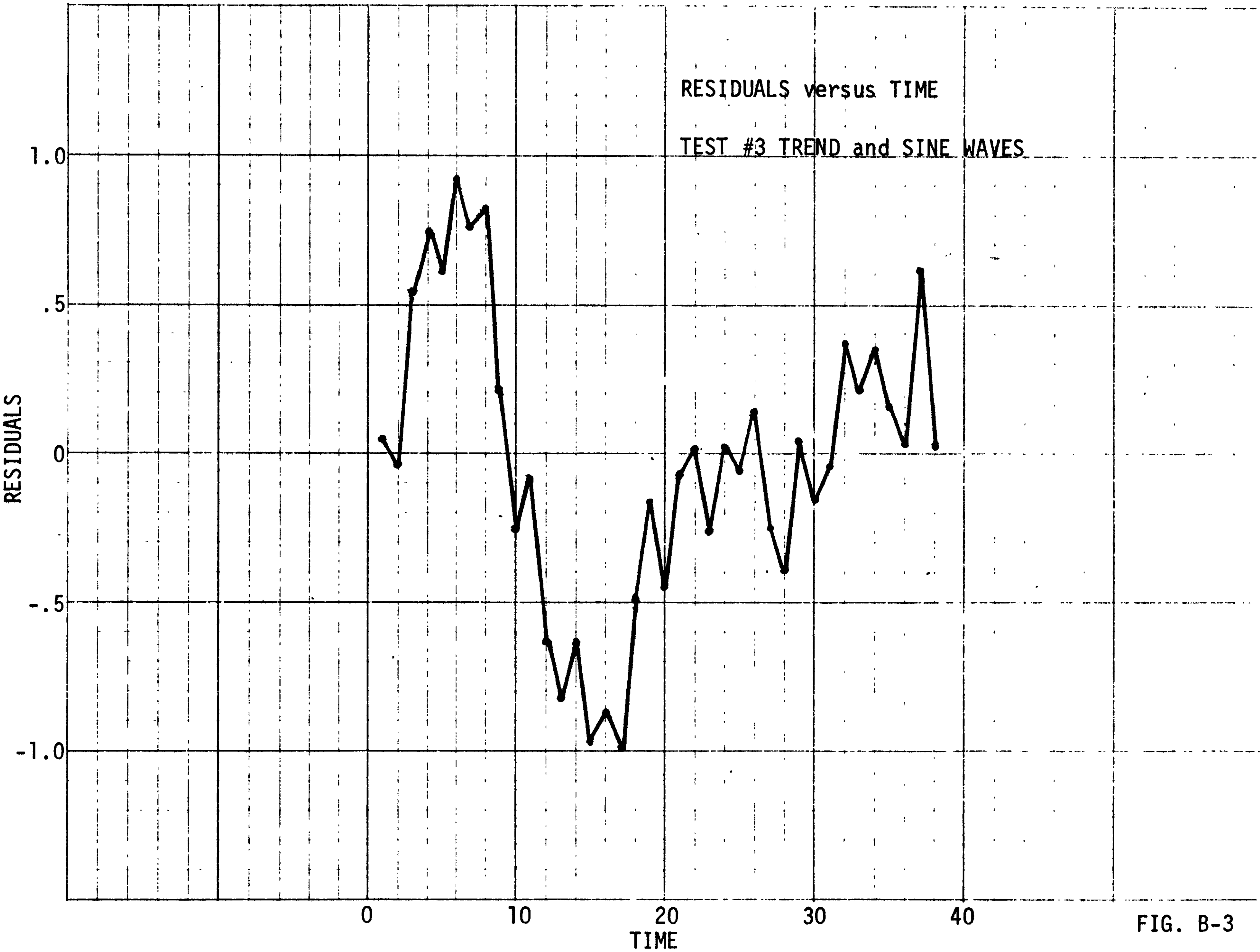

RESIDUALS versus TIME
TEST #3 TREND and SINE WAVES
RESIDUALS
1.0
.5
0
-.5
-1.0
0
10
20
30
40
TIME

FIG. B-3

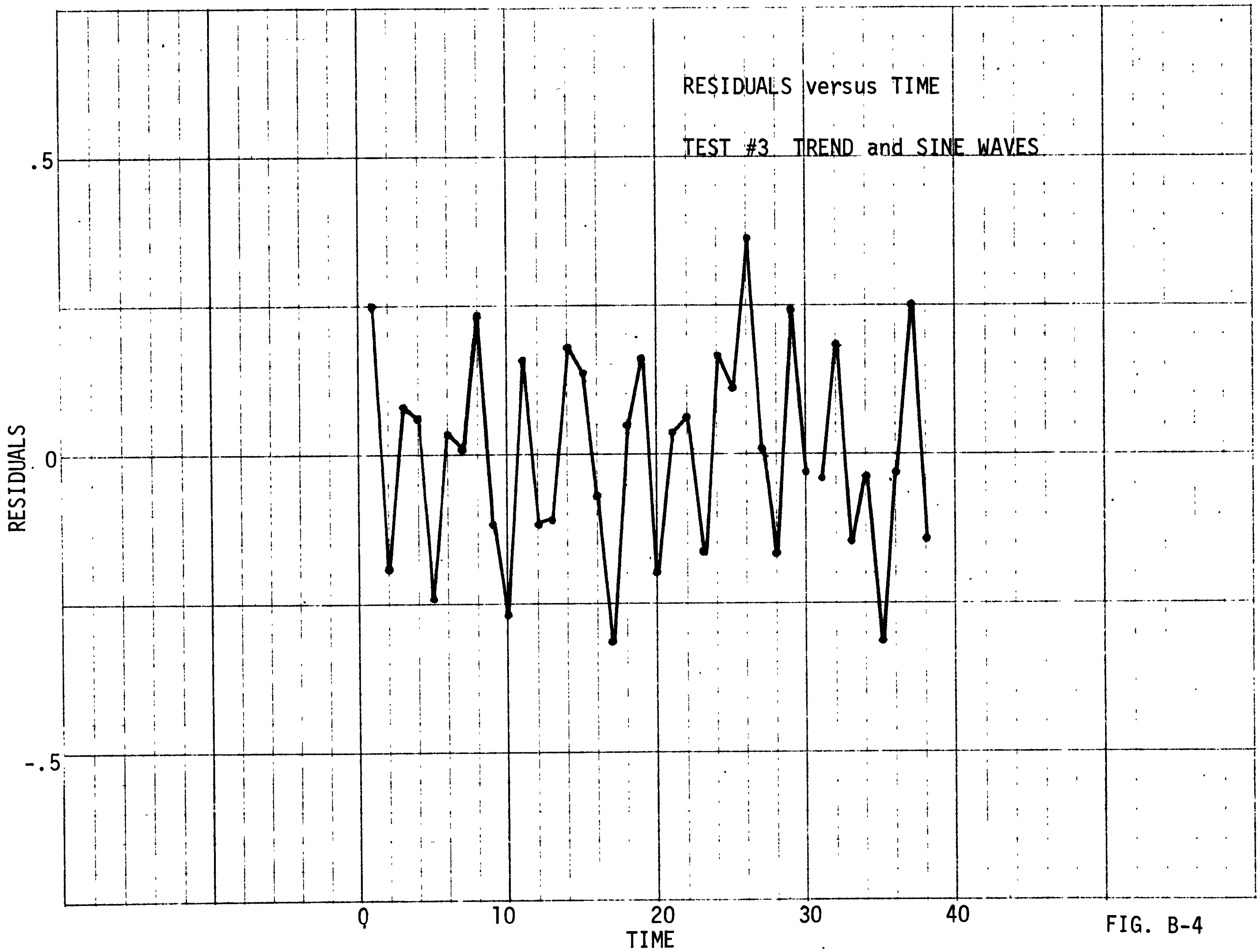

RESIDUALS versus TIME
TEST #3 TREND and SINE WAVES
.5
0
-.5
RESIDUALS
0
10
20
30
40
TIME

FIG. B-4

APPENDIX C SUMMARY OF RESULTS IN LITERATURE SURVEY

	CAB	BARTLETT	DOUGLAS	VERLEGER	BOEING CO.	PORT AUTHORITY	PORT AUTHORITY	THIS STUDY
	Δ LOG	LOG	LOG	GRAVITY	DETRENDED	ΔLOG	LOG	DETRENDED
DEPENDENT								
Passengers				X				
RPM			X			X	X	X
RPM/Capita	X	X			X			
INDEPENDENT								
INCOME								
GNP/CAP.							X	
Pers. Inc. Current $					X			
Disp. Pers. Inc. Real $								X
Inc. & Population				X				
National Inc.						X		
YIELD								
Aver. Fare/Pass. Mile Real $	X	X	X		X	X	X	
Price Index				X				
BUSINESS ACTIVITY								
Indust. Prod. Consumer Goods		X						
Prod. Durable Equip. Outlays		X						
CONSUMER SPENDING								
Pers. Consumption Expend.			X					X
P.C.E. Services		X						
P.C.E. Recreation		X						
Other								
Distance Between Cities				X				
Velocity of Money			X					
Interest Rate Ratio			X					
Time	X	X				X	X	X
Trip Length			X					

SUMMARY OF RESULTS IN LITERATURE SURVEY (continued)

		CAB	BARTLETT	DOUGLAS	VERLEGER	BOEING CO.	PORT AUTHORITY	THIS STUDY
Price Elasticity		-1.35	-1.8	-1.47		-.43	-1.61	-1.56
Income Elasticity		1.09	.8	1.93		2.5	1.38	.96
RPM Forecast 1975	Low	151.0		149.1			140.2	122
	High	168.2		153.4			168.4	193
1980	Low	210.4		224.5			206.8	186
	High	306.5		248.0			334.7	294

APPENDIX D

DATA

TABLE	ITEM
D.1	Revenue Passenger Miles, RPM Yield in cents per mile, current dollars Time Trend
D.2	Consumer Price Index, CPI Deflator for Personal Consumption Expenditures
D.3	Personal Income Disposable Personal Income Personal Consumption Expenditures
D.4	Non-Linear Secular Trend for RPM

TABLE D.1

Year		RPM (millions) 48 state	50 state		YIELD current $/¢ mile 48 state	YIELD 50 state	TREND
1940		1052.2			5.07		1
41		1384.7			5.04		2
42		1417.5			5.27		3
43		1632.5			5.35		4
44		2176.8			5.34		5
45		3360.3			4.95		6
46		5944.9			4.63		7
47		6105.3			5.05		8
48		5996.6			5.76		9
49		6767.6			5.78		10
50		8029.1			5.56		11
51		10589.7			5.61		12
52		12559.3			5.57		13
53		14793.9			5.46		14
54		16802.4			5.41		15
55		19852.1			5.36		16
56		22398.6			5.33		17
57		25378.7			5.31		18
58		25375.5			5.64		19
59		29307.6			5.88		20
60		30556.6			6.09		21
61		31062.3			6.28		22
62		33622.6			6.45		23
63		38456.6			6.17		24
64		44141.3			6.12		25
65		51887.4			6.06		26
66		60590.8			5.83		27
67		75487.3			5.64		28
68		87507.7			5.61		29
69		95945.9	102717.4		5.90	5.79	30
70	A	97281.1	104146.8	A	6.11	6.00	31
71	A	99422.	106438.4	A	6.45	6.33	32
72	A	110350.	118138.0	A	6.52	6.40	33
73	A	117990.	126317.4	A	6.76	6.63	34
74	A	121179.	129731.0	A	7.62	7.48*	35

*-Estimate
A-Adjusted from 50 states to 48 states by a constant percentage difference.

TABLE D.2

YEAR	CPI 1967=100	CPI 1958=100	Implicit Price Deflator for Personal Consumption Expenditures
1940	42.0	48.5	45.5
41	44.1	50.9	48.7
42	48.8	56.4	54.8
43	51.8	59.8	59.9
44	52.7	60.9	63.2
45	53.9	62.2	65.4
46	58.5	67.6	70.5
47	66.9	77.3	77.9
48	72.1	83.3	82.3
49	71.4	82.4	81.7
50	72.1	83.3	82.9
51	77.8	89.8	88.6
52	79.5	91.8	90.5
53	80.1	92.5	91.7
54	80.5	93.0	92.5
55	80.2	92.6	92.8
56	81.4	94.0	94.8
57	84.3	97.3	97.7
58	86.6	100.0	100.0
59	87.3	100.8	101.3
60	88.7	102.4	102.9
61	89.6	103.5	103.9
62	90.6	104.6	104.9
63	91.7	105.9	106.1
64	92.9	107.3	107.4
65	94.5	109.1	108.8
66	97.2	112.2	111.5
67	100.0	115.5	114.4
68	104.2	120.3	118.4
69	109.8	126.8	123.5
70	116.3	134.3	129.3
71	121.3	140.1	134.4
72	125.3	144.7	138.2
73	133.1	153.7	145.9
74	147.7	170.6	162.4

TABLE D.3

Personal Income		Disposable Personal Income		Personal Consumption Expenditures	
Curr $	Const $.58	Curr $	Const $.58	Curr $	Const $.58
78.3	172.1	75.7	166.3	70.8	155.7
96.0	197.1	92.7	190.3	80.6	165.4
122.9	224.3	116.9	213.4	88.6	161.4
151.3	252.6	133.5	222.8	99.3	165.8
165.3	261.6	146.3	231.6	108.3	171.4
171.1	261.6	150.2	229.7	119.7	183.0
178.7	253.5	160.0	227.0	143.4	203.5
191.3	245.6	169.8	218.0	160.7	206.3
210.2	255.4	189.1	229.8	173.6	210.8
207.2	253.6	188.6	230.8	176.8	216.5
227.6	274.5	206.9	249.6	191.0	230.5
255.6	288.5	226.6	255.7	206.3	232.8
272.5	301.1	238.3	263.3	216.7	239.4
288.2	314.3	252.6	275.4	230.0	250.8
290.1	313.6	257.4	278.3	236.5	255.7
310.9	335.0	275.3	296.7	254.4	274.2
333.0	351.3	293.2	309.3	266.7	281.4
351.1	359.4	308.5	315.8	281.4	288.2
361.2	361.2	318.8	318.8	290.1	290.1
383.5	378.6	337.3	333.0	311.2	307.3
401.0	389.7	350.0	340.2	325.2	316.1
416.8	401.2	364.4	350.7	335.2	332.5
442.6	421.9	385.3	367.3	255.1	338.4
465.5	438.7	404.6	381.3	375.0	353.3
497.5	463.2	438.1	407.9	401.2	373.7
538.9	495.3	473.2	435.0	432.8	397.7
587.2	526.6	511.9	458.9	466.3	418.1
629.3	550.1	546.3	477.5	492.1	430.1
688.9	581.8	591.0	499.0	536.2	452.7
750.9	608.0	634.4	513.6	579.5	469.1
808.3	625.1	691.7	534.8	617.6	477.5
864.0	642.9	746.4	555.4	667.1	496.4
944.9	683.7	802.5	580.5	729.0	527.3
1055.0	723.1	903.7	619.6	805.2	552.1
1150.4	708.4	979.7	603.2	877.0	539.9

TABLE D.4

Non-Linear Secular Trend for RPM

	NLRPM3	NLRPM
1940		3.01350
1941		3.12110
1942		3.22458
1943		3.32408
1944		3.41977
1945		3.51179
1946	3.71421	3.60028
1947	3.77738	3.68537
1948	3.83944	3.76720
1949	3.90042	3.84589
1950	3.96032	3.92157
1951	4.01917	3.99434
1952	4.07699	4.06431
1953	4.13379	4.13161
1954	4.18959	4.19632
1955	4.24442	4.25855
1956	4.29828	4.31839
1957	4.35119	4.37594
1958	4.40318	4.43128
1959	4.45425	4.48450
1960	4.50443	4.53567
1961	4.55373	4.58488
1962	4.60216	4.63221
1963	4.64974	4.67772
1964	4.69648	4.72148
1965	4.74240	4.76357
1966	4.78752	4.80403
1967	4.83185	4.84295
1968	4.87539	4.88038
1969	4.91817	4.91637
1970	4.96020	4.95098
1971	5.00150	4.98426
1972	5.04206	5.01626
1973	5.08192	5.04704
1974	5.12107	5.07663

APPENDIX E

Multiple Regressions Using Time-Series Data and Linear and Non-Linear Detrending

Table		Item
E.1	Variable names	
Regression Set 1 - Linear Detrending - Current Dollars - 1940-72		
E.2.1	Step 1	Detrend Variable
E.2.2		(continued)
E.3.1	Step 2	Select Best Measures
E.3.2		(continued)
E.4	Step 3	Final Form, Step-wise addition of variables
Regression Set 2 - Linear Detrending -Constant Dollars - 1946-72		
E.5	Step 1	Detrend Variables
E.6	Step 2	Select Best Measures
E.7	Step 3	Step-wise addition of variables
Regression Set 3 - Linear Detrending - Current Dollars - 1946-74		
E.8	Step 1	Detrend Variables
E.9	Step 2	Select Best Measures
E.10	Correlation Matrix for Step 2	
E.11	Step 3	Step-wise addition of variables
Regression Set 4 - Non-Linear Detrending - Current Dollars - 1946-74		
E.12	Results of Non-Linear Regression on RPM	
E.13.1	Step 2	Select Best Measure
E.13.2	Correlation matrix for E.13.1	
E.14	Step 3	Step-wise addition of variables
Regression Set 5 - Moving Time Period		
E.15	Single trend for total time period - 9 and 14 years	
E.16	Single time trend - Linear and Non-Linear	
E.17	Independent linear detrending for each period- 15 years	
E.18	Correlation matrices for E.19	
E.19	Step 3	Final form for each time period.
Regression Set 6 - Other Functional Forms		
E.20.1	Log RPM with Log (YLD and INC) plus DL (INFL)	
E.20.2	(continued)	
E.21	All variables Delta Log	
E.22	Log RPM and Delta Log for independent variables	

Regression Set 7 - Forecast Test

E.23	Regression equations with linear and non-linear trends for 1946 through 1969 and 1946 through 1974
E.24	Comparison of forecasts for 70 through 74 with actual data
E.25	Forecasts for 1975, 1980, 1985 and 1990 with both linear and non-linear trends

Symbols used in Tables

$\bar{R}^2$	Adjusted Coefficient of Multiple Determination
S_e	Standard error of estimate
D.W.	Durbin-Watson statistic
D.F.	Degrees of Freedom
( )	t-ratios in parenthesis below coefficients

E.1 VARIABLE NAMES

1940-72	1946-72	1946-74
CPI	CPI2	CPI3
DID	DID2	DID3
TRD	TRD2	TRD3
		NLRPM3
Current $	Constant $	Current $
DPI	DPI2	DPI3
PCE	PCE2	PCE3
PERI	PERI2	PERI3
RPM	RPM2	RPM3
Current $	Current $	Current $
YLD	YLD2	YLD3

E.2.1 REGRESSION SET 1

Current Dollars 1940-1972 Linear Detrending

Step 1 Detrend Variables

		CONST.	TRD	DUM	Corr. Coeff. TRD	Corr. Coeff. DUM	$\bar{R}^2$	S_e	D.W.	D.F.
1	LYLD	.703	.0028		.76		.5675	.0230	.645	31
		(85.7)	(6.6)							
2	YLD	5.035	.036		.76		.5692	.2961	.608	31
		(47.7)	(6.6)							
3	LYLD	.702	.0029	-.0048	.76	.40	.5546	.0234	.655	30
		(78.9)	(5.5)	(-.32)						
4	YLD	5.026	.0365	-.045	.76	.41	.5556	.3007	.615	30
		(43.9)	(5.5)	(-.23)						
5	LRPM	3.167	.062		.98		.9645	.1142	.221	31
		(77.9)	(29.5)							
6	RPM	-22736.	3308.		.91		.8156	15157	.083	31
		(-4.21)	(11.9)							
7	LRPM	3.141	.064	-.127	.98	.51	.9669	.1104	.291	30
		(74.7)	(26.2)	(-1.78)						
8	RPM	-13540.	2453.	44143.	.91	.79	.9341	9062.	.590	30
		(-3.92)	(12.2)	(7.5)						

E.2.2 REGRESSION SET 1

Current Dollars 1940-1974 Linear Detrending

Step 1 Detrend Variables

		CONST.	TRD	Corr. Coeff.	$\bar{R}^2$	S_e	D.W.	D.F.
9	LPCE	1.917	.0286	.99	.9833	.0383	.139	33
		(145)	(44.7)					
10	PCE	-25.250	19.975	.95	.9022	67.29	.080	33
		(-1.09)	(17.7)					
11	LPERI	2.020	.0287	.99	.9789	.0430	.296	33
		(136)	(39.8)					
12	PERI	-39.789	25.841	.94	.8815	96.89	.077	33
		(-1.19)	(15.9)					
13	LDPI	1.985	.0276	.99	.9833	.0368	.305	33
		(156)	(44.7)					
14	DPI	-23.436	21.875	.94	.8860	80.26	.078	33
		(-.85)	(16.3)					
15	LCPI	1.746	.0126	.97	.9316	.0350	.185	33
		(145)	(21.5)					
16	CPI	48.958	2.733	.97	.9302	7.665	.228	33
		(18.5)	(21.3)					
17	LDID	1.749	.0122	.96	.9182	.0374	.147	33
		(135)	(19.6)					
18	DID	50.528	2.581	.97	.9435	6.467	.250	33
		(22.6)	(23.9)					

E.3.1 REGRESSION SET 1

Current Dollars 1940-1972 Linear Detrending

Step 2 Select Best Measure

LRPM*	CONST.	LPCE*	LPERI*	LDPI*	LCPI*	LDID	$\bar{R}^2$	S_e	D.W.	D.F.
1	.0029	2.420				.0561	.6993	.0616	.529	30
	(.27)	(2.56)				(.06)				
2	.0019	3.610			-1.350		.7201	.0595	.412	30
	(.18)	(4.47)			(-1.5)					
3	.0014		-.916			3.258	.6828	.0633	.837	30
	(.13)		(-2.15)			(6.7)				
4	.0042		-.107		2.538		.5349	.0766	.720	30
	(.31)		(-.24)		(4.59)					
5	.0008			-1.167		3.328	.6896	.0626	.866	30
	(.07)			(-2.32)		(6.85)				
6	.0040			-.281		2.660	.5381	.0764	.729	30
	(.3)			(-.52)		(4.66)				

	LRPM*	LPCE*	LPERI*	LDPI*	LCPI*	LDID*
LRPM*	1.00	.85	.51	.50	.75	.81
LPCE*	.85	1.00	.76	.77	.94	.95
LPERI*	.51	.76	1.00	.99	.70	.79
LDPI*	.50	.77	.99	1.00	.73	.79
LCPI*	.75	.94	.70	.73	1.00	.96
LDID*	.81	.95	.79	.79	.96	1.00

* Linear Detrending

E.3.2 REGRESSION SET 1

Current Dollars 1940-1972 Linear Detrending

Step 2 Select Best Measure

LRPM*	CONST.	LYLD*	LPCE*	LCPI*	$\bar{R}^2$	S_e	D.W.	D.F.
7	-7.06E-6	-1.829			.1225	.1001	.212	31
	(0)	(-2.34)						
8	2.72E-3	-1.490	2.346		.8507	.0413	.884	30
	(.38)	(-4.6)	(12.3)					
9	3.92E-3	-2.053	.957	1.615	.8642	.0394	1.163	29
	(.57)	(-4.91)	(1.33)	(1.99)				

	LRPM*	LYLD*	LPCE*	LCPI*
LRPM*	1.00	-.39	.87	.78
LYLD*	-.39	1.00	-.08	.15
LPCE*	.87	-.08	1.00	.94
LCPI*	.78	.15	.94	1.00

* Linear Detrending

E.4 REGRESSION SET 1

Current Dollars 1940-1972 Linear Detrending

Step 3 Step-Wise Addition Of Variables

LRPM	CONST.	LYLD*	LPCE*	LCPI*	TRD	DUM	$\bar{R}^2$	S_e	D.W.	D.F.
1	3.167				.062		.9645	.1142	.221	31
	(77.9)				(29.5)					
2	3.161		2.480		.062		.9898	.0614	.529	30
	(144)		(8.79)		(55.3)					
3	3.161	-1.481	2.306		.062		.9928	.0516	.317	29
	(172)	(-3.66)	(10.1)		(65.7)					
4	3.139	-1.488	2.362		.064	-.108	.9952	.0420	.813	28
	(196)	(-4.51)	(12.2)		(69.0)	(-3.97)				
5	3.134	-2.084	.891	1.711	.065	-.115	.9957	.0398	1.167	27
	(204)	(-4.87)	(1.2)	(2.04)	(71.5)	(-4.43)				

	LRPM	LYLD*	LPCE*	LCPI*	TRD	DUM
LRPM	1.00	-.07	.10	.05	.98	.51
LYLD*	-.07	1.00	-.08	.15	.00	.00
LPCE*	.10	-.08	1.00	.94	-.06	-.08
LCPI*	.05	.15	.94	1.00	-.10	-.07
TRD	.98	.00	-.06	-.10	1.00	.57
DUM	.51	.00	-.08	-.07	.57	1.00

* Linear Detrending

E.5 REGRESSION SET 2

Constant Dollars 1946-1972 Linear Detrending

Step 1 Detrend Variables

		CONST.	TRD2	Corr. Coeff.	$\bar{R}^2$	S_e	D.W.	D.F.
1	LRPM2	3.710	.051	.99	.9847	.0509	.379	25
		(184)	(40.9)					
2	RPM2	-15544.	3934.	.93	.8687	12106	1.093	25
		(-3.24)	(13.2)					
3	LYLD2	.719	3.03E-3	.71	.4472	.0247	.537	25
		(73.4)	(4.97)					
4	YLD2	5.208	.039	.71	.4810	.3197	.523	25
		(41.2)	(5.01)					
5	LPCE2	2.270	.016	1.00	.9931	.0115	.635	27
		(517)	(63.6)					
6	PCE2	151.8	12.76	.98	.9617	21.66	.177	27
		(18.4)	(26.5)					
7	LPERI2	2.349	.018	1.00	.9899	.0149	.617	27
		(411)	(52.6)					
8	PERI2	172.6	17.33	.98	.9519	33.15	.159	27
		(13.7)	(23.6)					
9	LDPI2	2.304	.0167	1.00	.9902	.0141	.680	27
		(428)	(53.1)					
10	DPI2	160.9	14.39	.98	.9547	26.69	.183	27
		(15.8)	(24.3)					

E.6 REGRESSION SET 2

Constant Dollars 1946-1972 Linear Detrending

Step 2 Select Best Measures

LRPM2*	CONST.	LYLD2*	LPCE2*	LPERI2*	LDPI2*	DLCPI	DLDID	$\bar{R}^2$	S_e	D.W.	D.F.
1	.0169			1.446		-1.186		.0595	.0484	.520	24
	(1.14)			(1.85)		(-1.42)					
2	.0125				.827	-.717		.0323	.0507	.459	24
	(.68)				(.99)	(-.87)					
3	.0167			1.402			-1.324	.0308	.0491	.465	24
	(.97)			(1.68)			(-1.12)				
4	3.88E-7	-1.473						.4907	.0356	.342	25
	(0)	(-5.1)									
5	-7.22E-4	-1.975		-1.282				.5591	.0331	.712	24
	(-.11)	(-5.61)		(-2.21)							
6	.0134	-1.942		-.712		-1.016		.6606	.0315	.945	23
	(1.38)	(-5.79)		(-1.13)		(-1.87)					

	LRPM2*	LYLD2*	LPCE2*	LPERI2*	LDPI2*	DLCPI	DLDID
LRPM2*	1.00	-.71	.02	.24	.13	-.09	-.02
LYLD2*	-.71	1.00	--	-.65	--	-.33	-.38
LPCE2*	.02	--	1.00	.89	.93	.49	.52
LPERI2*	.24	-.65	.89	1.00	.96	.56	.62
LDPI2*	.13	--	.93	.96	1.00	.48	.54
DLCPI	-.09	-.33	.49	.56	.48	1.00	.99
DLDID	-.02	-.38	.52	.62	.54	.99	1.00

E.7 REGRESSION SET 2

Constant Dollars 1946-1972 Linear Detrending

Step 3 Step-Wise Addition Of Variables

LRPM2	CONST.	LYLD2*	LPERI2*	DLCPI	TRD2	$\bar{R}^2$	S_e	D.W.	D.F.
1	3.710				.0514	.9847	.0509	.379	25
	(184)				(40.9)				
2	3.710	-1.473			.0514	.9922	.0363	.343	24
	(258)	(-5.0)			(57.2)				
3	3.711	-1.980	-1.292		.0512	.9932	.0338	.717	23
	(277)	(-5.5)	(-2.17)		(61.1)				
4	3.730	-1.951	-.707	-1.077	.0509	.9939	.0320	.991	22
	(231)	(-5.72)	(-1.1)	(-1.92)	(63.0)				

	LRPM2	LYLD2*	LPERI2*	DLCPI	TRD2
LRPM2	1.00	-.09	-.04	-.21	.99
LYLD2*	-.09	1.00	-.65	-.33	.00
LPERI2*	-.04	-.65	1.00	.56	-.07
DLCPI	-.21	-.33	.56	1.00	-.20
TRD2	.99	.00	-.07	-.20	1.00

* Linear Detrending

E.8 REGRESSION SET 3

Current Dollars 1946-1974 Linear Detrending

Step 1 Detrend Variables

		CONST.	TRD3	Corr. Coeff.	$\bar{R}^2$	S_e	D.W.	D.F.
1	LRPM3	3.723	.0501	.99	.9840	.0544	.332	27
		(180)	(41.5)					
2	LYLD3	.710	3.75E-3	.76	.5561	.0276	.514	27
		(67.5)	(6.13)					
3	LPCE3	2.139	.0260	1.00	.9910	.0211	.222	27
		(266)	(55.6)					
4	LPERI3	2.218	.0272	.99	.9888	.0246	.258	27
		(236)	(49.8)					
5	LDPI3	2.173	.0264	.99	.9892	.0235	.232	27
		(243)	(50.6)					

E.9 REGRESSION SET 3

Current Dollars 1946-1974 Linear Detrending

Step 2 Select Best Measure

LRPM3*	CONST.	LYLD3*	LPCE3*	LPERI3*	LDPI3*	DLCPI3	DLDID3	$\bar{R}^2$	S_e	D.W.	D.F.
1	-1.13E-6	-1.511						.5728	.0349	.370	27
	(0)	(-6.21)									
2	-1.51E-6	-1.485	-.124					.5586	.0355	.385	26
	(0)	(-5.77)	(-.37)								
3	-1.07E-6	-1.508		-.026				.5565	.0356	.374	26
	(0)	(-6.02)		(-.09)							
4	-1.55E-6	-1.493			-.113			.5588	.0355	.383	26
	(0)	(-5.92)			(-.38)						
5	.0202	-1.612	.410			-1.339		.6480	.0317	.902	25
	(2.15)	(-6.88)	(1.15)			(-2.76)					
6	.0225	-1.590		.490		-1.491		.6663	.0308	.963	25
	(2.43)	(-7.26)		(1.66)		(-3.09)					
7	.0200	-1.582			.347	-1.326		.6468	.0317	.916	25
	(2.13	(-6.94)			(1.11)	(-2.73)					
8	.0203	-1.584	.415				-1.491	.6222	.0328	.784	25
	(1.9)	(-6.55)	(1.07)				(-2.32)				
9	.0237	-1.560		.525			-1.74	.6424	.0319	.847	25
	(2.23)	(-6.91)		(1.62)			(-2.69)				
10	.0205	-1.556			.370		-1.504	.6225	.0328	.801	25
	(1.91)	(-6.63)			(1.08)		(-2.32)				

E.10 REGRESSION SET 3

Current Dollars 1946-1974 Linear Detrending

CORRELATION MATRIX FOR E.9

	LRPM3	LRPM3*	LYLD3*	LPCE3*	LPERI3*	LDPI3*	DLCPI3	DLDID	TRD3
LRPM3	1.00	.12	-.10	-.03	-.01	-.02	-.02	.02	.99
LRPM3*	.12	1.00	-.77	-.26	-.12	-.19	-.27	-.26	.00
LYLD3*	-.10	-.77	1.00	.28	.14	.19	-.02	.02	.00
LPCE3*	-.03	-.26	.28	1.00	.97	.98	.51	.58	.00
LPERI3*	-.01	-.12	.14	.97	1.00	.99	.56	.63	.00
LDPI3*	-.02	-.19	.19	.98	.99	1.00	.52	.60	.00
DLCPI3	-.02	-.27	-.02	.51	.56	.52	1.00	.98	.01
DLDID3	.02	-.26	.02	.58	.63	.60	.98	1.00	.05
TRD3	.99	.00	.00	.00	.00	.00	.01	.05	1.00

E.11 REGRESSION SET 3

Current Dollars 1946-1974 Linear Detrending

Step 3 Step-Wise Addition Of Variables

LRPM3	CONST.	LYLD3*	LPCE3*	LPERI3*	LDPI3*	DLCPI3	DLDID3	TRD3	$\bar{R}^2$	S_e	D.W.	D.F.
1	3.723							.050	.9840	.0543	.332	27
	(180)							(41.5)				
2	3.723	-1.509						.050	.9931	.0356	.371	26
	(275)	(-6.08)						(63.4)				
3	3.723	-1.484	-.121					.050	.9929	.0362	.385	25
	(270)	(-5.65)	(-.35)					(62.3)				
4	3.723	-1.506		-.024				.050	.9929	.0363	.375	25
	(269)	(-5.89)		(-.08)				(62.2)				
5	3.723	-1.491			-.111			.050	.9929	.0362	.384	25
	(270)	(-5.8)			(-.37)			(62.4)				
6	3.742	-1.611	.411			-1.337		.050	.9943	.0324	.901	24
	(261)	(-6.73)	(1.12)			(-2.7)		(70)				
7	3.745	-1.588		.491		-1.489		.050	.9946	.0315	.962	24
	(267)	(-7.1)		(1.63)		(-3.02)		(71.6)				
8	3.742	-1.580			.348	-1.324		.050	.9943	.0324	.915	24
	(261)	(-6.79)			(1.09)	(-2.67)						
9	3.744	-1.559		.528			-1.744	.050	.9942	.0326	.849	24
	(251)	(-6.66)		(1.59)			(-2.64)	(69.2)				

E.12 REGRESSION SET 4

RESULTS OF NON-LINEAR REGRESSION ON LRPM

1946 to 1974 29 Points

$$3.6499 + 3.6618 \left[1 - \exp(-.0177\, t) \right]$$

1940 to 1974 35 Points

$$2.9016 + 2.9170 \left[1 - \exp(-.0391\, t) \right]$$

Note: A non-linear trend was used only for LRPM,
Yield and Income used linear detrending.

E.13.1 REGRESSION SET 4

Current Dollars 1940-1974 Non-Linear Secular Trend For LRPM3

Step 2 Select Best Measure

LRPM3**	CONST.	LYLD3*	LPCE3*	LPERI3*	LDPI3*	DLCPI3	DLDID3	$\bar{R}^2$	S_e	D.W.	D.F.
1	5.39E-5	-1.385						.6743	.0259	.614	27
	(.01)	(7.68)									
2	5.65E-5	-1.570	.872					.8192	.0193	.961	26
	(.02)	(-11.2)	(4.76)								
3	5.22E-5	-1.485		.784				.8461	.0178	.985	26
	(.02)	(-11.9)		(5.6)							
4	5.67E-5	-1.510			.767			.8196	.0193	.999	26
	(.02)	(-11.0)			(4.77)						
5	5.55E-3	-1.604	1.017			-.364		.8229	.0191	1.141	25
	(.98)	(-11.4)	(4.72)			(-1.24)					
6	7.63E-3	-1.512		.958		-.502		.8600	.0170	1.309	25
	(1.5)	(-12.6)		(5.89)		(-1.89)					
7	5.42E-3	-1.534			.890	-.355		.8229	.0191	1.207	25
	(.96)	(-11.2)			(4.72)	(-1.22)					
8	4.11E-3	-1.589	.979				-.298	.8164	.0194	1.072	25
	(.65)	(-11.1)	(4.26)				(-.78)				
9	7.14E-3	-1.501		.949			-.521	.8528	.0174	1.231	25
	(1.24)	(-12.2)		(5.36)			(-1.48)				
10	4.45E-3	-1.524			.871		-.323	.8176	.0194	1.137	25
	(.7)	(-11.0)			(4.29)		(-.84)				

* Linear Detrending ** Non-Linear Detrending

E.13.2 REGRESSION SET 4

CORRELATION MATRIX FOR E.13.1

	LRPM3**	LYLD3*	LPCE3*	LPERI3*	LDPI3*	DLCPI3	DLDID3
LRPM3**	1.00	-.83	.14	.29	.22	.14	.16
LYLD3*	-.83	1.00	.28	.14	.19	-.02	.02
LPCE3*	.14	.28	1.00	--	--	.51	.58
LPERI3*	.29	.14	--	1.00	--	.56	.63
LDPI3*	.22	.19	--	--	1.00	.52	.60
DLCPI3	.14	-.02	.51	.56	.52	1.00	--
DLDID3	.16	.02	.58	.63	.60	--	1.00

E.14 REGRESSION SET 4

Current Dollars 1946-1974 Non-Linear Secular Trend For RPM3

Step 3 Step-Wise Addition Of Variables

LRPM3	CONST.	LYLD3*	LPCE3*	LPERI3*	LDPI3*	DLCPI3	DLDID3	NLRPM3	$\bar{R}^2$	S_e	D.W.	D.F.
1	-1.92E-4							1.000	.9885	.0462	.449	27
	(0)							(49.0)				
2	2.89E-3	-1.386						.999	.9962	.0264	.614	26
	(.06)	(-7.54)						(85.7)				
3	-5.51E-3	-1.570	.873					1.001	.9979	.0196	.962	25
	(-.14)	(-11.0)	(4.67)					(115)				
4	-6.00E-3	-1.485		.785				1.001	.9982	.0181	.985	25
	(-.17)	(11.63)		(5.48)				(125)				
5	-5.57E-3	-1.510			.768			1.001	.9979	.0196	1.000	25
	(-.14)	(-10.83)			(4.68)			(115)				
6	5.40E-4	-1.604	1.018			-.363		1.001	.9979	.0195	1.141	24
	(.01)	(-11.13)	(4.62)			(-1.22)		(116)				
7	2.29E-3	-1.512		.959		-.501		1.001	.9983	.0173	1.309	24
	(.07)	(-12.31)		(5.78)		(-1.85)		(131)				
8	3.37E-4	-1.534			.891	-.355		1.001	.9979	.0195	1.207	24
	(.01)	(-11.0)			(4.62)	(-1.19)		(116)				
9	-2.57E-3	-1.590	.982				-.300	1.002	.9979	.0198	1.075	24
	(-.06)	(-11.0)	(4.18)				(-.77)	(114)				
10	-9.45E-4	-1.501		.952			-.524	1.002	.9983	.0177	1.235	24
	(-.03)	(-12.0)		(5.27)			(-1.46)	(128)				
11	2.41E-3	-1.524			.873		-.325	1.002	.9979	.0198	1.140	24
	(-.06)	(-10.8)			(4.21		(-.83)	(114)				

E.15 REGRESSION SET 5

Current Dollars 1946-1972 Moving Time Period Of 9 And 14 Years

Single Trend For Total Time Period

LRPM (points)	CONST.	LYLD*	LPCE*	DLCPI	TRD	$\bar{R}^2$	S_e	D.W.	D.F.
1-9	3.598	-1.520	.378	.465	.069	.9869	.0205	2.263	4
	(122)	(-2.45)	(.21)	(.52)	(13.8)				
10-18	3.741	-.205	3.78	1.57	.054	.9928	.0076	2.513	4
	(73.9)	(-.67)	(4.26)	(2.17)	(12.3)				
19-27	3.819	-1.760	.666	1.405	.045	.9923	.0124	2.906	4
1-14	3.631	-1.900	2.105	-.459	.062	.9978	.0186	2.010	9
	(198)	(-7.1)	(3.4)	(-1.2)	(36.1)				
6-19	3.755	-1.509	1.061	-.949	.051	.9914	.0172	.576	9
	(65.2)	(-4.93)	(.84)	(-1.47)	(9.73)				
11-24	3.802	-.863	2.327	-.483	.049	.9969	.0117	1.744	9
	(266)	(-3.7)	(4.1)	(-.42)	(39.3)				
14-27	3.920	-1.625	1.186	1.935	.041	.9980	.0098	2.803	9
	(150)	(-10.5)	(4.42)	(2.34)	(30.2)				

E.16 REGRESSION SET 5

Current Dollars 1946-1974 Moving Time Period Of 20 Years

For Linear and Non-Linear and 15 years for Non-Linear

LRPM3 (points)	CONST.	LYLD3*	LPERI3*	DLCPI3	NLRPM3	$\bar{R}^2$	S_e	D.W.	D.F.
1-15	-.269	-1.545	1.116	-.158	1.067	.9842	.0335	.338	10
	(-1.3)	(-4.04)	(1.62)	(-.24)	(21.4)				
8-22	.157	-1.357	.978	-.704	.968	.8880	.0671	.055	10
	(.18)	(-.91)	(.35)	(-.13)	(4.65)				
15-29	.668	-1.623	1.256	1.182	.858	.7430	.1120	.026	10
1-20	-.172	-1.581	1.411	-.413	1.043	.9904	.0304	.404	15
	(-.99)	(-5.1)	(2.63)	(-.74)	(25.8)				
5-24	-.139	-1.247	1.553	-.817	1.036	.9830	.0384	.223	15
	(-.49)	(-2.15)	(1.63)	(-.57)	(15.4)				
10-29	.234	-1.403	.905	.433	.950	.9830	.0345	.265	15
	(1.2)	(-4.7)	(1.83)	(.278)	(21.5)				
	CONST.	LYLD3*	LPERI3*	DLCPI3	TRD3				
1-20	3.676	-1.613	2.014	-1.031	.057	.9940	.0237	1.060	15
	(181)	(-6.6)	(4.64)	(-2.41)	(32.9)				
5-24	3.751	-1.409	.887	-1.050	.050	.9919	.0269	.824	15
	(152)	(-3.5)	(1.37)	(-1.03)	(21.8)				
10-25	3.824	-1.458	.325	.737	.045	.9960	.0162	1.330	15
	(243)	(-10.5)	(1.42)	(1.02)	(45.9)				

E.17 REGRESSION SET 5

Current Dollars 1946-1974 Moving Time Period of 15 Years

Step 1 Detrending Variables

Points		CONST.	TRD3A	TRD3B	TRD3C	$\bar{R}^2$	S_e	D.W.	D.F.
1-15	LPERI3A	2.239	2.520			.9860	.0134	1.237	13
		(307)	(31.5)						
1-15	LYLD3A	.713	3.32E-3			.2129	.0254	.683	13
		(51.7)	(2.19)						
8-22	LPERI3B	2.252		.0239		.9879	.0118	.603	13
		(204)		(33.8)					
8-22	LYLD3B	.702		.0041		.3652	.0226	.391	13
		(33.2)		(3.0)					
15-29	LPERI3C	2.077			.0334	.9933	.0123	.431	13
		(126)			(45.5)				
15-29	LYLD3C	.721			.0033	.1388	.0309	.400	13
		(17.4)			(1.8)				

Correlation Matrices for [illegible]

Points 1-15

	LRPM3A	LYLD3A*	LPERI3A*	DLCPI3A	TRD3A
LRPM3A	1.00	-.14	.08	-.55	.99
LYLD3A*	-.14	1.00	-.02	-.23	.00
LPERI3A*	.08	-.02	1.00	.08	.00
DLCPI3A	-.55	-.23	.08	1.00	-.57
TRD3A	.99	.00	.00	-.57	1.00

Points 8-22

	LRPM3B	LYLD3B*	LPERI3B*	DLCPI3B	TRD3B
LRPM3B	1.00	-.16	.14	.51	.98
LYLD3B*	-.16	1.00	-.81	-.48	.00
LPERI3B*	.14	-.81	1.00	.39	.00
DLCPI3B	.51	-.48	.39	1.00	.44
TRD3B	.98	.00	.00	.44	1.00

Points 15-29

	LRPM3C	LYLD3C*	LPERI3C*	DLCPI3C	TRD3C
LRPM3C	1.00	-.19	-.07	.81	.98
LYLD3C*	-.19	1.00	.60	.21	.00
LPERI3C*	-.07	.60	1.00	.33	.00
DLCPI3C	.81	.21	.33	1.00	.84
TRD3C	.98	.00	.00	.84	1.00

E.19 REGRESSION SET 5

Current Dollars 1946-1974 Moving Time Period of 15 Years

Step 3 Regression for each Period

Points		CONST.	LYLD3A*	LPERI3A*	DLCPI3A	TRD3A	$\bar{R}^2$	S_e	D.W.	D.F.
1-15	LRPM3A	3.679	-1.554	1.596	-.534	.0577	.9971	.0143	2.580	10
		(309)	(-9.54)	(5.38)	(-1.92)	(54.6)				
		CONST.	LYLD3B*	LPERI3B*	DLCPI3B	TRD3B				
8-22	LRPM3B	3.832	-1.324	.401	.079	.0442	.9926	.0173	1.148	10
		(237)	(-3.42)	(.571)	(.059)	(37.1)				
15-29		CONST.	LYLD3C*	LPERI3C*	DLCPI3C	TRD3C				
	LRPM3C	3.787	-1.711	.895	1.349	.0460	.9951	.0156	1.369	10
		(110)	(-9.76)	(1.75)	(1.54)	(22)				

E.20.1 REGRESSION SET 6

Current Dollars 1946-1974

LOG(RPM,YIELD,INC) PLUS DELTA LOG(INFL)

LRPM3	CONST.	LYLD3	LPCE3	LPERI3	LDPI3	DLCPI3	DLDID3	TRD3	$\bar{R}^2$	S_e	D.W.	D.F.
1	-.986	7.127							.4640	.3145	.269	27
	(-.91)	(5.02)										
2	.363	-1.784	2.166						.9809	.0594	.169	26
	(1.72)	(-4.2)	(27)									
3	.324	-1.542		2.030					.9787	.0627	.164	26
	(1.45)	(-3.5)		(25.6)								
4	.310	-1.627			2.107				.9781	.0636	.184	26
	(1.37)	(-3.6)			(25.2)							
5	.422	-1.900	2.193			-2.638			.9891	.0449	.773	25
	(2.63)	(-5.9)	(36.1)			(-4.53)						
6	.392	-1.686		2.063		-2.932			.9889	.0453	.805	25
	(2.42)	(-5.27)		(35.7)		(-5.0)						
7	.373	-1.758			2.137	-2.789			.9873	.0483	.823	25
	(2.16)	(-5.11)			(33.4)	(-4.5)						
8	4.005	-1.611	.412			-1.338		.0455	.9943	.0324	.901	24
	(5.4)	(6.7)	(1.13)			(-2.7)						
9	3.782	-1.588		.491		-1.489		.0427	.9946	.0315	.963	24
	(5.78)	(-7.1)		(1.63)		(-3.0)		(5.3)				
10	4.107	-1.580			.348	-1.324		.0468	.9943	.0324	.916	24
	(6.1)	(-6.8)			(1.09)	(-2.67)						

E.20.2 REGRESSION SET 6

Current Dollars 1946-1974

LOG(RPM,YIELD,INC) PLUS DELTA LOG(INFL)

LRPM3	CONST.	LYLD3	LPCE3	LPERI3	LDPI3	DLCPI3	DLDID3	TRD3	$\bar{R}^2$	S_e	D.W.	D.F.
11	.386	-1.849	2.194				-3.24		.9892	.0446	.798	25
	(2.4)	(-5.8)	(36.3)				(-4.6)					
12	.354	-1.633		2.065			-3.67		.9895	.0441	.845	25
	(2.25)	(-5.3)		(36.7)			(-5.3)					
13	.336	-1.709			2.139		-3.518		.9879	.0472	.853	25
	(2.0)	(-5.1)			(34.2)		(-4.7)					
14	3.971	-1.583	.418				-1.494	.0426	.9939	.0335	.786	24
	(4.93)	(-6.4)	(1.05)				(-2.27)	(4.5)				
15	3.680	-1.559		.528			-1.744	.0417	.9943	.0326	.850	24
	(5.1)	(-6.8)		(1.59)			(-2.64)	(4.7)				
16	4.035	-1.555			.373		-1.506	.0462	.9939	.0335	.803	24
	(5.5)	(-6.49)			(1.06)		(-2.27)	(5.1)				

	LRPM3	LYLD3	LPCE3	LPERI3	LDPI3	DLCPI3	DLDID3	TRD3
LRPM3	1.00	.70	.98	.99	.98	-.02	.02	.99
LYLD3	.70	1.00	.78	.77	.77	.00	.05	.76
LPCE3	.98	.78	1.00	--	--	.06	.11	1.00
LPERI3	.99	.77	--	1.00	--	.07	.12	.99
LDPI3	.98	.77	--	--	1.00	.07	.11	.99
DLCPI3	-.02	.00	.06	.07	.07	1.00	--	.01
DLDID	.02	.05	.11	.12	.11	--	1.00	.05
TRD3	.99	.76	1.00	.99	.99	.01	.05	1.00

E.21 REGRESSION SET 6

Current Dollars 1946-1974

DELTA LOG

DLRPM3	CONST.	DLYLD3	DLPCE3	DLPERI3	DLDPI3	DLCPI3	DLDID3	$\bar{R}^2$	S_e	D.W.	D.F.
1	.0404	-1.458	.4464			.3559		.5822	.0201	1.937	24
	(3.1)	(-5.16)	(.79)			(.72)					
2	.0375	-1.381		.5730		.2338		.6092	.0194	2.010	24
	(3.9)	(-4.98)		(1.53)		(.54)					
3	.0443	-1.492			.2436	.4968		.5776	.0202	2.029	24
	(4.2)	(-5.39)			(.6)	(1.19)					
4	.0414	-1.542	.2882				.7066	.5884	.0197	2.026	24
	(3.36)	(-5.7)	(.54)				(1.23)				
5	.0378	-1.459		.4756			.5047	.6189	.0192	2.042	24
	(4.0)	(-5.5)		(1.26)			(.95)				
6	.0439	-1.560			.1580		.8087	.5962	.0198	2.085	24
	(4.3)	(-5.8)			(.39)		(1.6)				

	DLRPM3	DLYLD3	DLPCE3	DLPERI3	DLDPI3	DLCPI3	DLDID3
DLRPM3	1.00	-.76	-.01	.17	.02	-.35	-.32
DLYLD3	-.76	1.00	.29	.13	.20	.67	.67
DLPCE3	-.01	.29	1.00	--	--	.67	.66
DLPERI3	.17	.13	--	1.00	--	.51	.53
DLDPI3	.02	.20	--	--	1.00	.48	.49
DLCPI3	-.35	.67	.67	.51	.48	1.00	--
DLDID3	-.32	.67	.66	.53	.49	--	1.00

E.22 REGRESSION SET 6

Current Dollars 1946-1974

LOG AND DELTA LOG

LRPM3	CONST.	DLYLD3	DLPCE3	DLPERI3	DLDPI3	DLCPI3	DLDID3	TRD3	$\bar{R}^2$	S_e	D.W.	D.F.
1	3.745	-1.198	-.876			.394		.0504	.9834	.0536	.494	23
	(104)	(-1.59)	(-.56)			(.3)		(38.3)				
2	3.732	-1.107		-.146		-.010		.0503	.9832	.0539	.443	23
	(128)	(-1.43)		(-.13)		(-.01)		(37.5)				
3	3.743				-.823	.281		.0506	.9835	.0534	.520	23
	(125)	(-1.61)			(-.72)	(.25)		(37.9)				
4	3.744	-1.190	-.843				.459	.0504	.9834	.0536	.493	23
	(106)	(-1.61)	(-.56)				(.29)	(38.5)				
5	3.732	-1.122		-.164			.029	.0503	.9831	.0539	.443	23
	(128	(-1.46)		(-.15)			(.02)	(37.7)				
6	3.743	-1.184			-.830		.369	.0505	.9835	.0533	.522	23
	(125)	(-1.63)			(-.73)		(.27)	(38.0)				

	LRPM3	DLYLD3	DLPCE3	DLPERI3	DLDPI3	DLCPI3	DLDID3	TRD3
LRPM3	1.00	.00	.26	.31	.32	.07	.13	.99
DLYLD3	.00	1.00	.29	.13	.20	.67	.66	.05
DLPCE3	.26	.29	1.00	--	--	.67	.66	.28
DLPERI3	.31	.13	--	1.00	--	.51	.53	.32
DLDPI3	.32	.20	--	--	1.00	.48	.49	.34
DLCPI3	.07	.67	.67	.51	.48	1.00	--	.11
DLDID3	.13	.67	.66	.53	.49	--	1.00	.16
TRD3	.99	.05	.28	.32	.34	.11	.16	1.00

E.23 REGRESSION SET 7

Current Dollars 1946 to 1969 and 1974

Comparison of Coefficients for 24 and 29 Points

1946-1969

LRPM4	CONST.	LYLD4*	LPERI4*	DLCPI4	TRD4	NLRPM4	$\bar{R}^2$	S_e	D.W.	D.F.
1	3.721	-1.372	1.268	-1.496	.0527		.9941	.0288	.827	19
	(207)	(-4.97)	(2.98)	(-3.07)	(40.6)					
2	-.076	-1.469	1.273	-.547		1.020	.9981	.0163	1.303	19
	(-1.24)	(-9.46)	(5.29)	(-1.96)		(72.1)				

1946-1974

LRPM3	CONST.	LYLD3*	LPERI3*	DLCPI3	TRD3	NLRPM3	$\bar{R}^2$	S_e	D.W.	D.F.
1	3.745	-1.588	.491	-1.489	.0501		.9946	.0315	.962	24
	(267)	(-7.1)	(1.63)	(-3.02)	(71.6)					
2	2.29E-3	-1.512	.959	-.501		1.001	.9983	.0173	1.309	24
	(.07)	(-12.3)	(5.78)	(-1.85)		(131)				

E.24 FORECAST TEST

LINEAR

LRPM = 3.721 -1.372 LYLD* + 1.268 LPERI* -1.496 DLCPI + .0527 TRD

	Calculated	Actual	Difference Log	Difference %
70	5.0366	4.9880	-.0386	9.29
71	5.0741	4.9974	-.0767	19.32
72	5.1469	5.0428	-.1041	27.09
73	5.1912	5.0719	-.1193	31.61
74	5.1622	5.0834	-.0788	19.89

Average Error = 21.4%

NON-LINEAR

LRPM = -.0763 -1.469 LYLD* + 1.273 LPERI* -.5470 DLCPI + 1.020 NLRPM

	Calculated	Actual	Difference Log	Difference %
70	5.0085	4.9880	-.0205	4.83
71	5.0273	4.9975	-.0299	7.12
72	5.0847	5.0428	-.0419	10.13
73	5.1273	5.0718	-.0555	13.63
74	5.0991	5.0834	-.0157	3.68

Average Error = 7.9%

E.25 FORECASTS FOR LINEAR AND NON-LINEAR SECULAR TRENDS

LINEAR

	LRPM (48 state)			RPM Billions (48 state)			RPM Billions (50 state)		
	Low		High	Low		High	Low		High
1975	5.1276	5.2476	5.2876	134	177	194	143	189	207
1980	5.3781	5.4981	5.5381	239	315	345	255	336	368
1985	5.6286	5.7486	5.786	425	561	615	435	597	655
1990	5.8781	5.9991	6.0391	755	998	1094	805	1064	1166

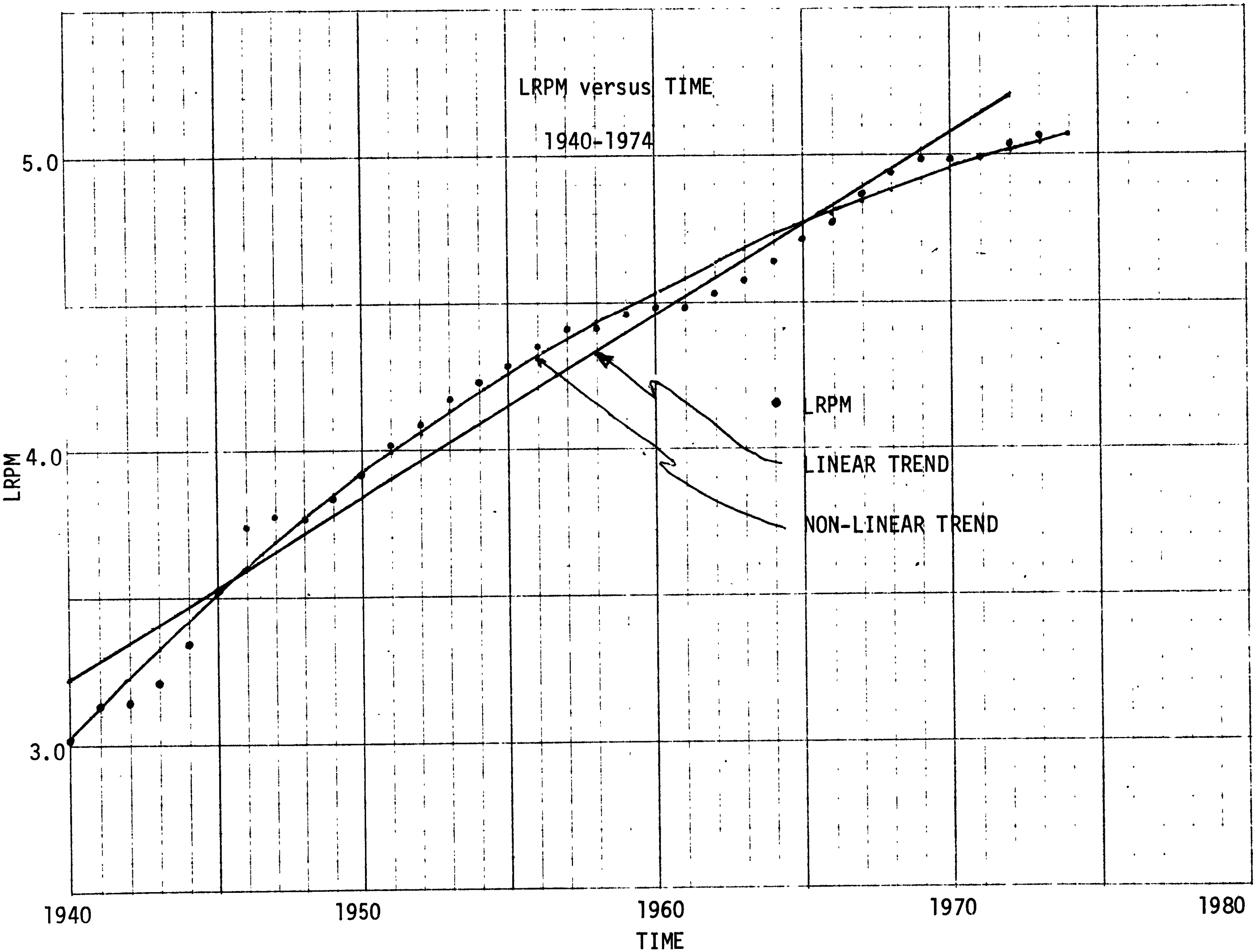
NON-LINEAR

	LRPM (48 state)			RPM Billions (48 state)			RPM Billions (50 state)		
	Low		High	Low		High	Low		High
1975	5.0570	5.1670	5.2570	114	147	181	122	157	193
1980	5.2407	5.3507	5.4407	174	224	276	186	239	294
1985	5.4078	5.5178	5.6078	256	329	405	273	351	432
1990	5.5609	5.6709	5.7609	364	469	577	388	500	615

Comparison of 1980 Forecasts

	This Study		Other Studies		
	Linear	Non-Linear	CAB	Douglas	N.Y. Port Auth.
Low	255	186	210	224	206
High	368	294	307	248	334

FIGURES



LRPM versus TIME
1940-1974
5.0
4.0
3.0
LRPM
1940
1950
1960
1970
1980
TIME
LRPM
LINEAR TREND
NON-LINEAR TREND

FIGURE 1

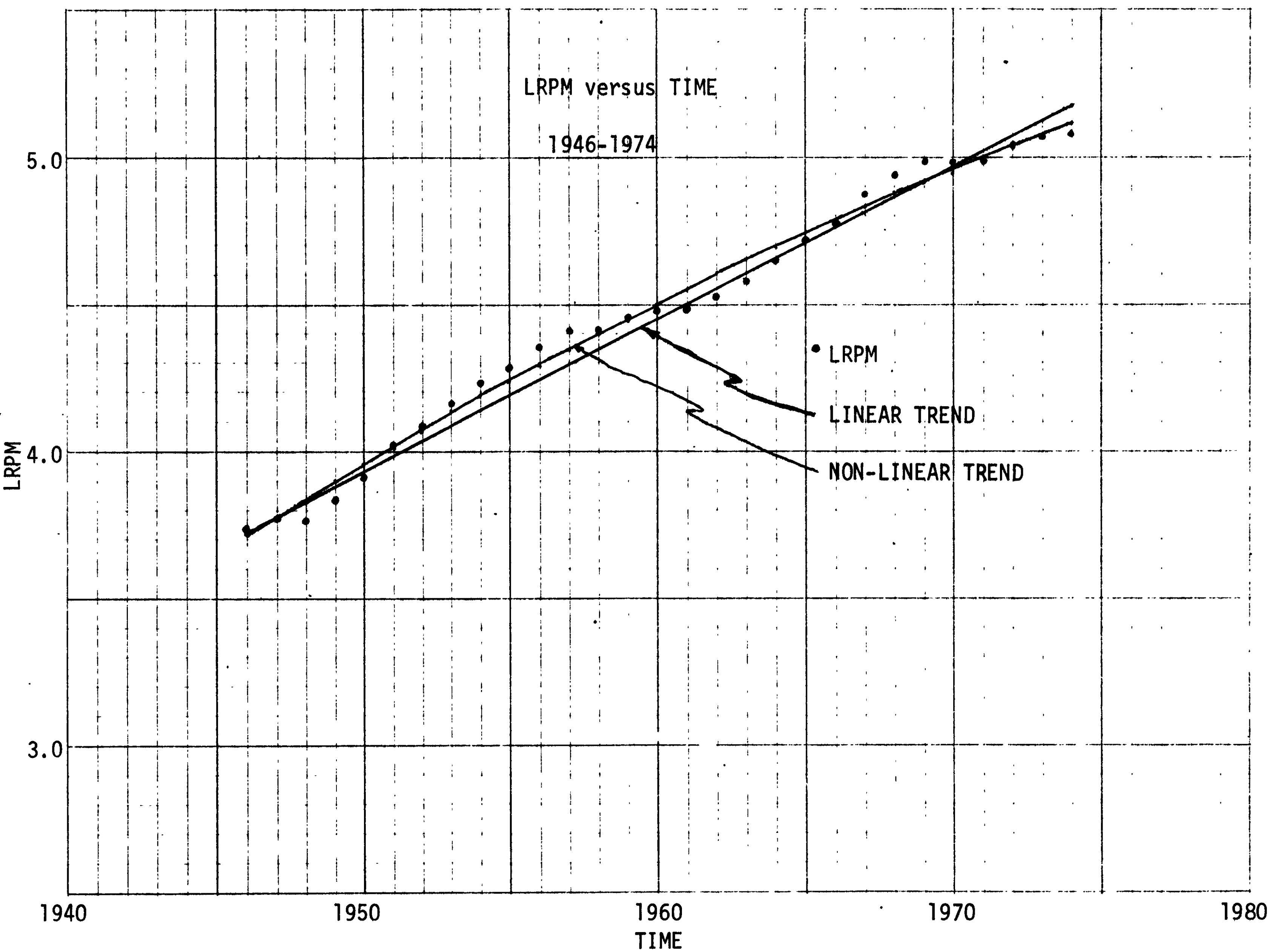
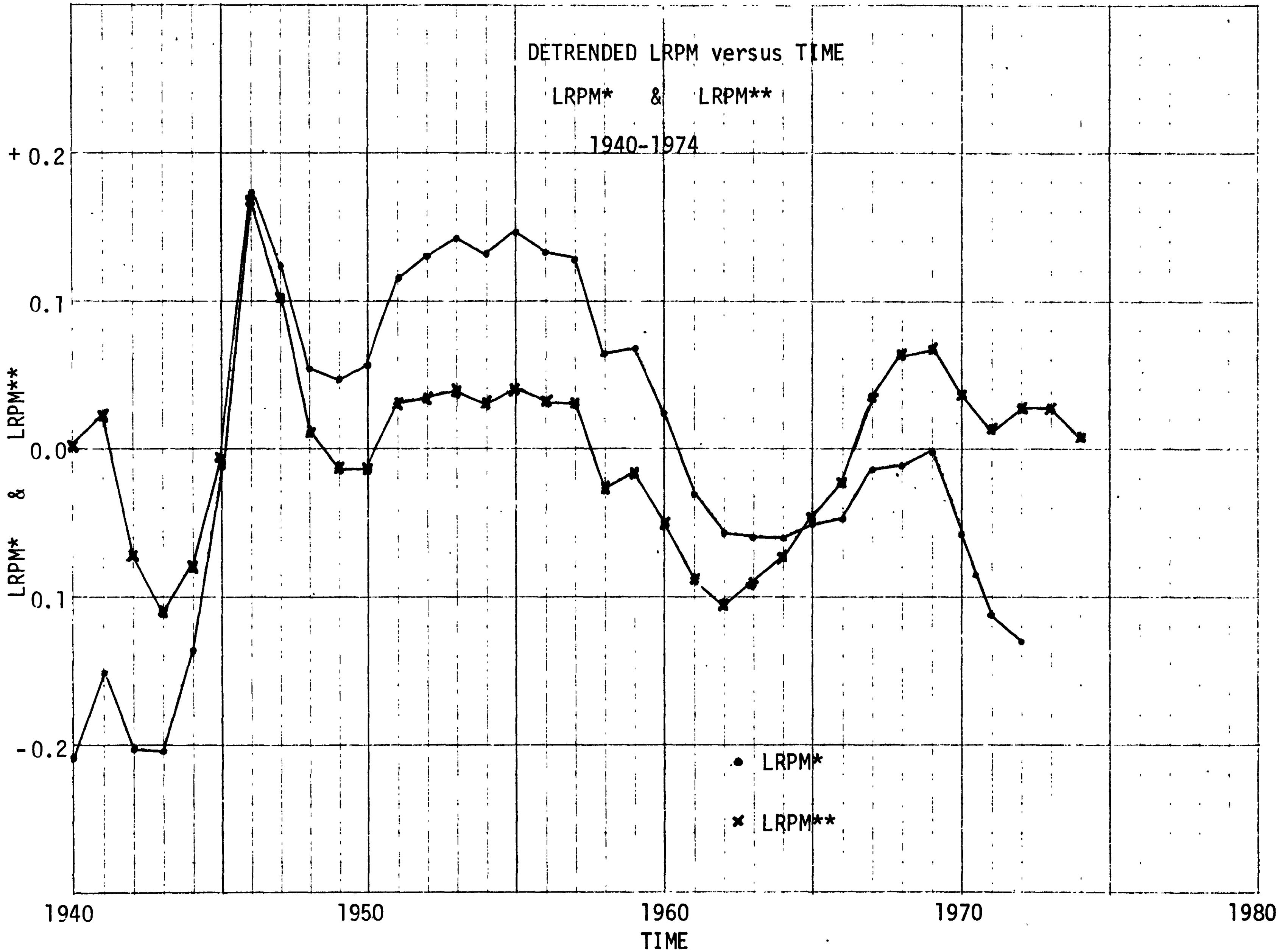

LRPM versus TIME
1946-1974
LRPM
LINEAR TREND
NON-LINEAR TREND
5.0
4.0
3.0
LRPM
1940
1950
1960
1970
1980
TIME

FIGURE 2



DETRENDED LRPM versus TIME
LRPM* & LRPM**
1940-1974
+0.2
0.1
0.0
-0.1
-0.2
LRPM* & LRPM**
1940
1950
1960
1970
1980
TIME
LRPM*
LRPM**

FIGURE 3

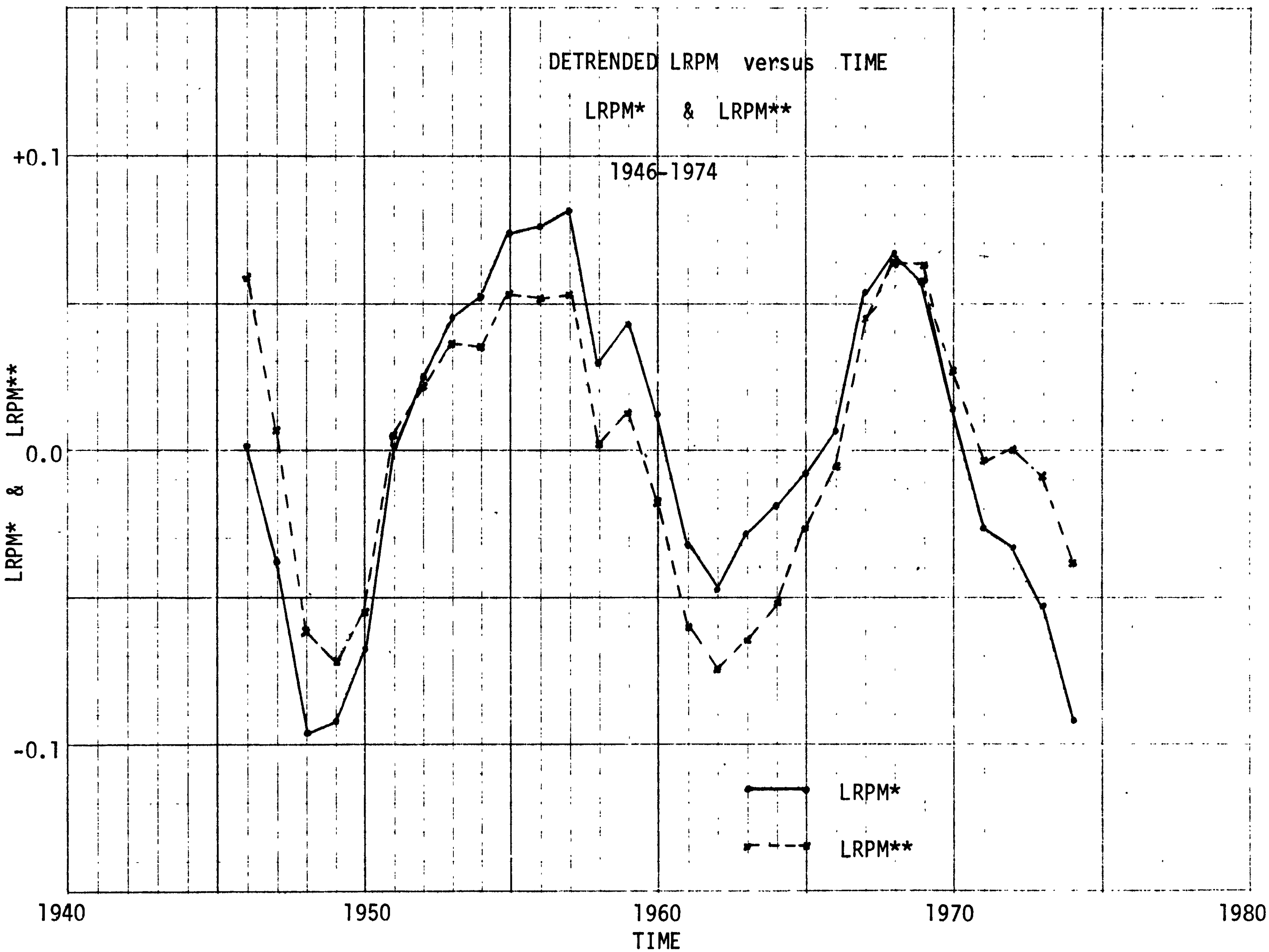

DETRENDED LRPM versus TIME
LRPM* & LRPM**
1946-1974
+0.1
0.0
-0.1
LRPM* & LRPM**
1940
1950
1960
1970
1980
TIME
LRPM*
LRPM**

FIGURE 4

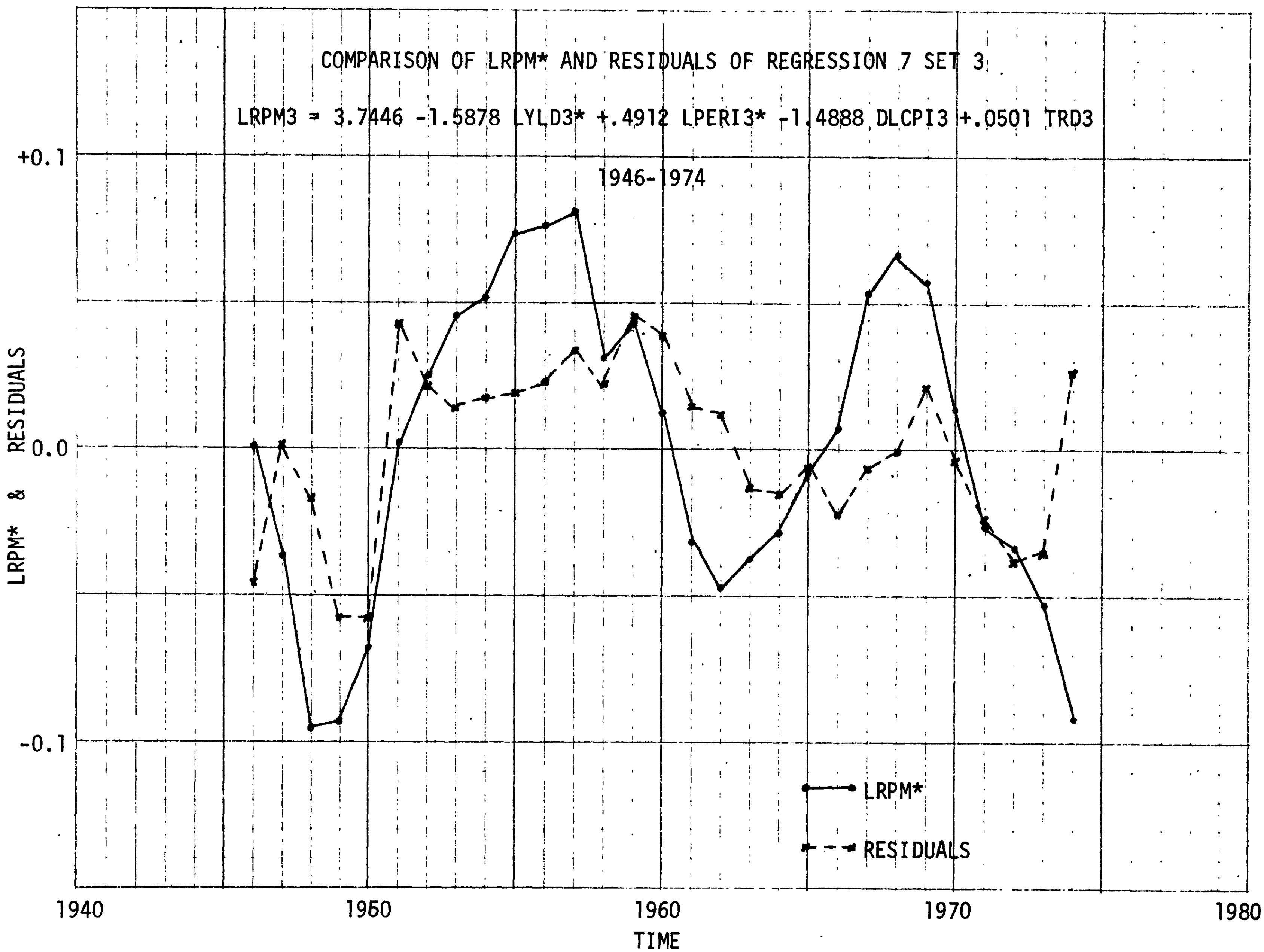

COMPARISON OF LRPM* AND RESIDUALS OF REGRESSION 7 SET 3
LRPM3 = 3.7446 -1.5878 LYLD3* +.4912 LPERI3* -1.4888 DLCPI3 +.0501 TRD3
1946-1974
+0.1
0.0
-0.1
LRPM* & RESIDUALS
1940
1950
1960
1970
1980
TIME
LRPM*
RESIDUALS

FIGURE 5

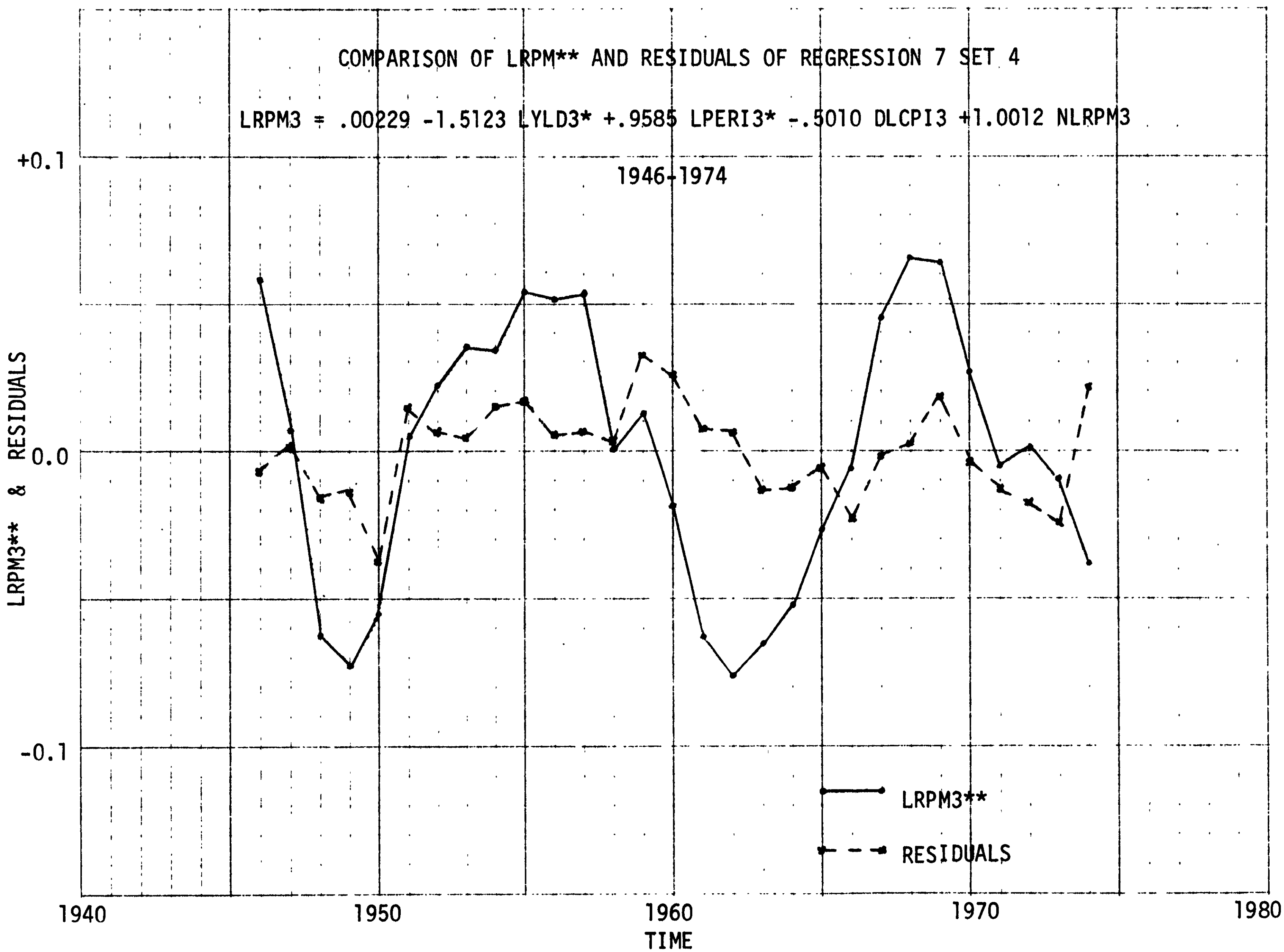

COMPARISON OF LRPM** AND RESIDUALS OF REGRESSION 7 SET 4
LRPM3 = .00229 -1.5123 LYLD3* +.9585 LPERI3* -.5010 DLCPI3 +1.0012 NLRPM3
1946-1974
LRPM3** & RESIDUALS
+0.1
0.0
-0.1
1940
1950
1960
1970
1980
TIME
LRPM3**
RESIDUALS

FIGURE 6

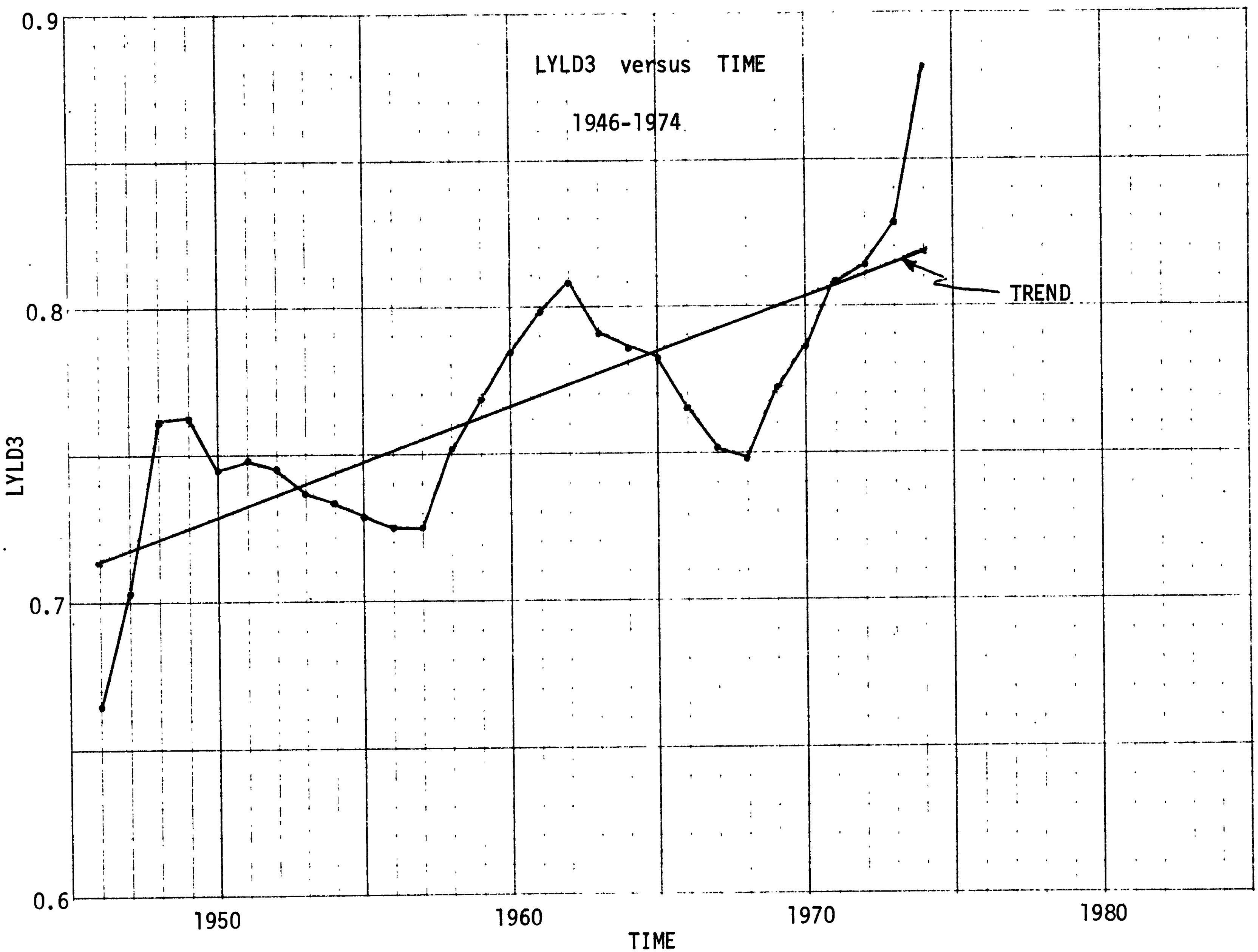

LYLD3 versus TIME
1946-1974
TREND
LYLD3
TIME
0.9
0.8
0.7
0.6
1950
1960
1970
1980

FIGURE 7

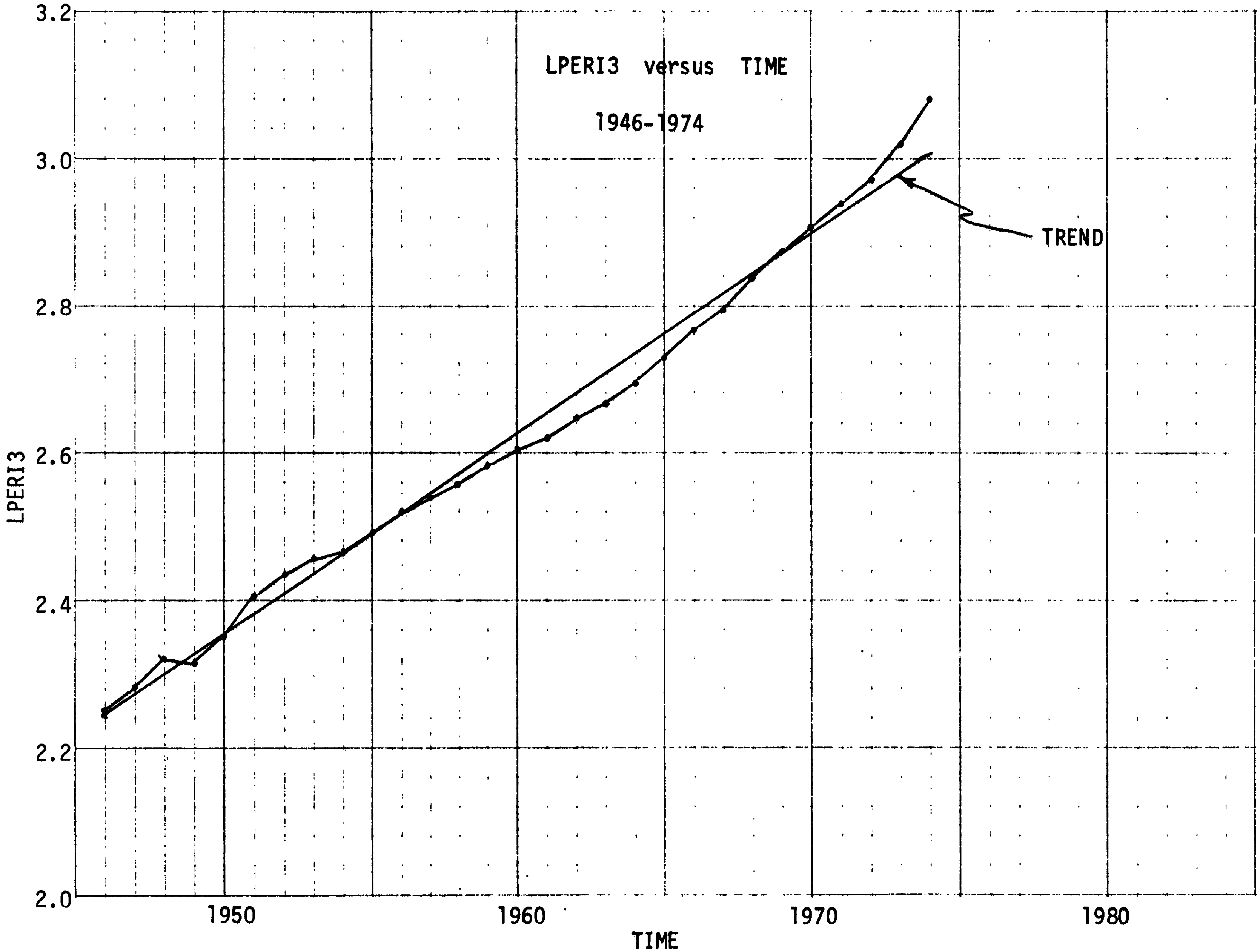

LPERI3 versus TIME
1946-1974
TREND
LPERI3
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1950
1960
1970
1980
TIME

FIGURE 8

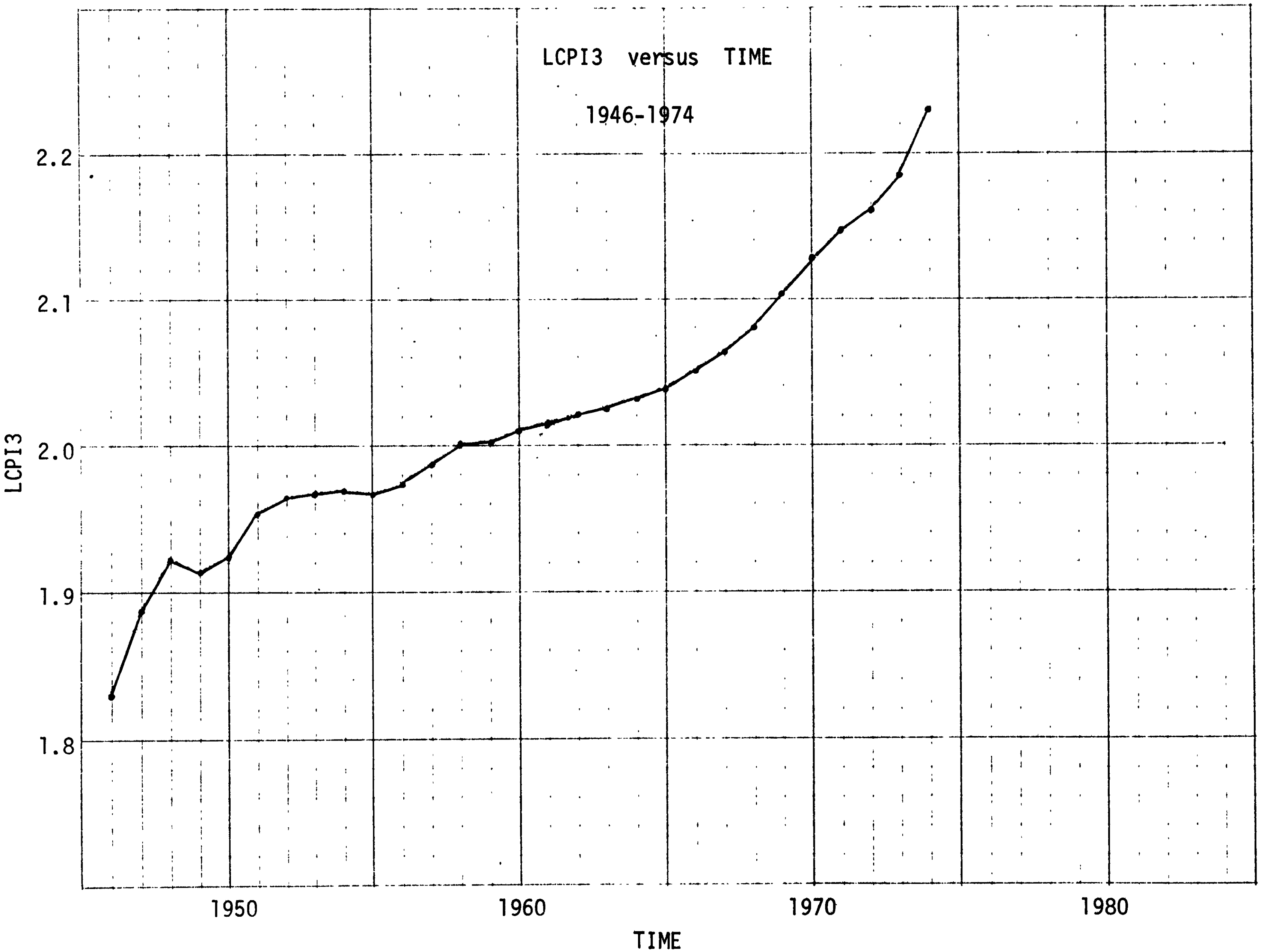

LCPI3 versus TIME
1946-1974
2.2
2.1
2.0
1.9
1.8
LCPI3
1950
1960
1970
1980
TIME

FIGURE 9

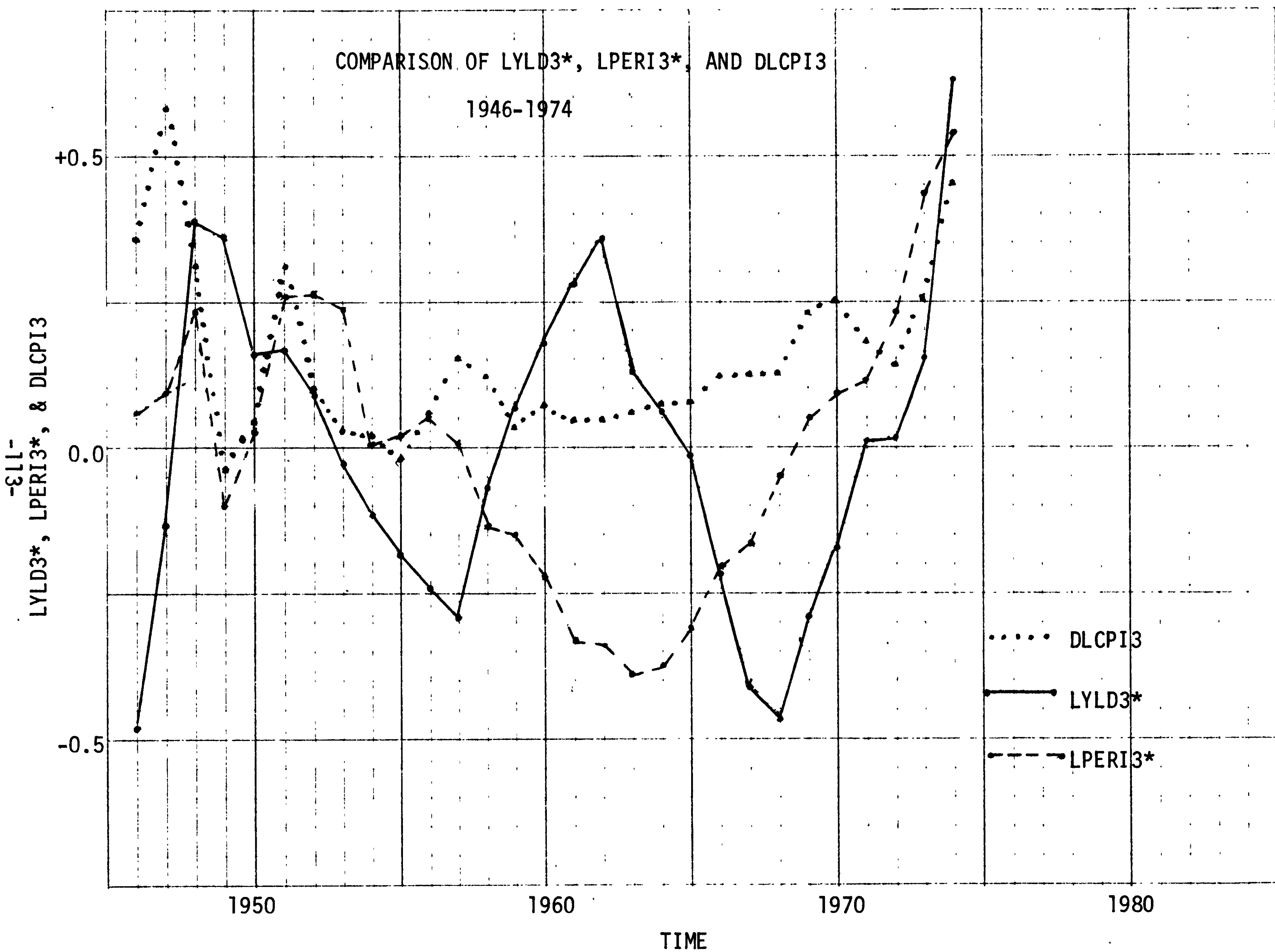

COMPARISON OF LYLD3*, LPERI3*, AND DLCPI3
1946-1974
+0.5
0.0
-0.5
LYLD3*, LPERI3*, & DLCPI3
1950
1960
1970
1980
TIME
DLCPI3
LYLD3*
LPERI3*

FIGURE 10

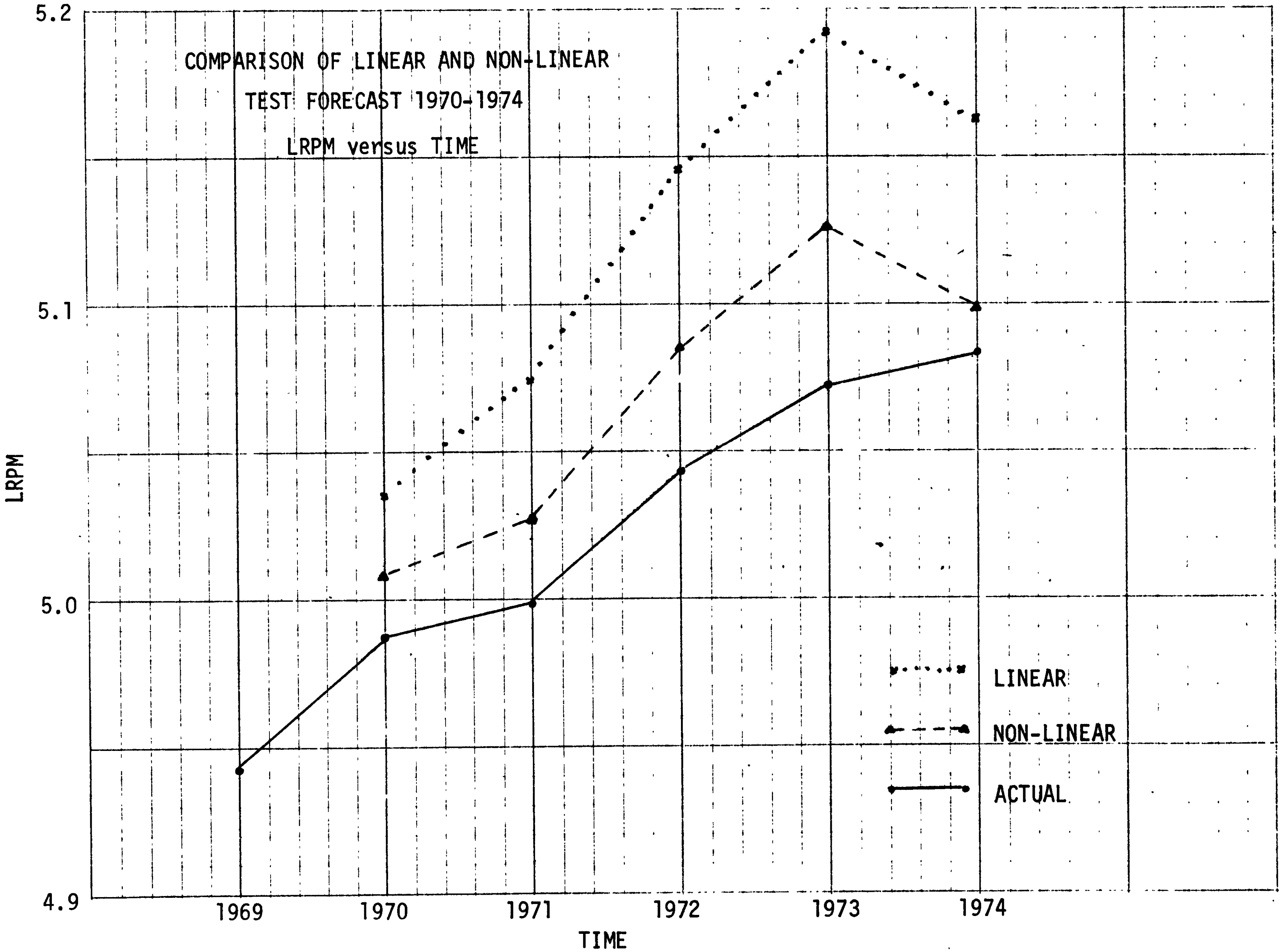

COMPARISON OF LINEAR AND NON-LINEAR
TEST FORECAST 1970-1974
LRPM versus TIME
LRPM
5.2
5.1
5.0
4.9
1969
1970
1971
1972
1973
1974
TIME
LINEAR
NON-LINEAR
ACTUAL

FIGURE 11

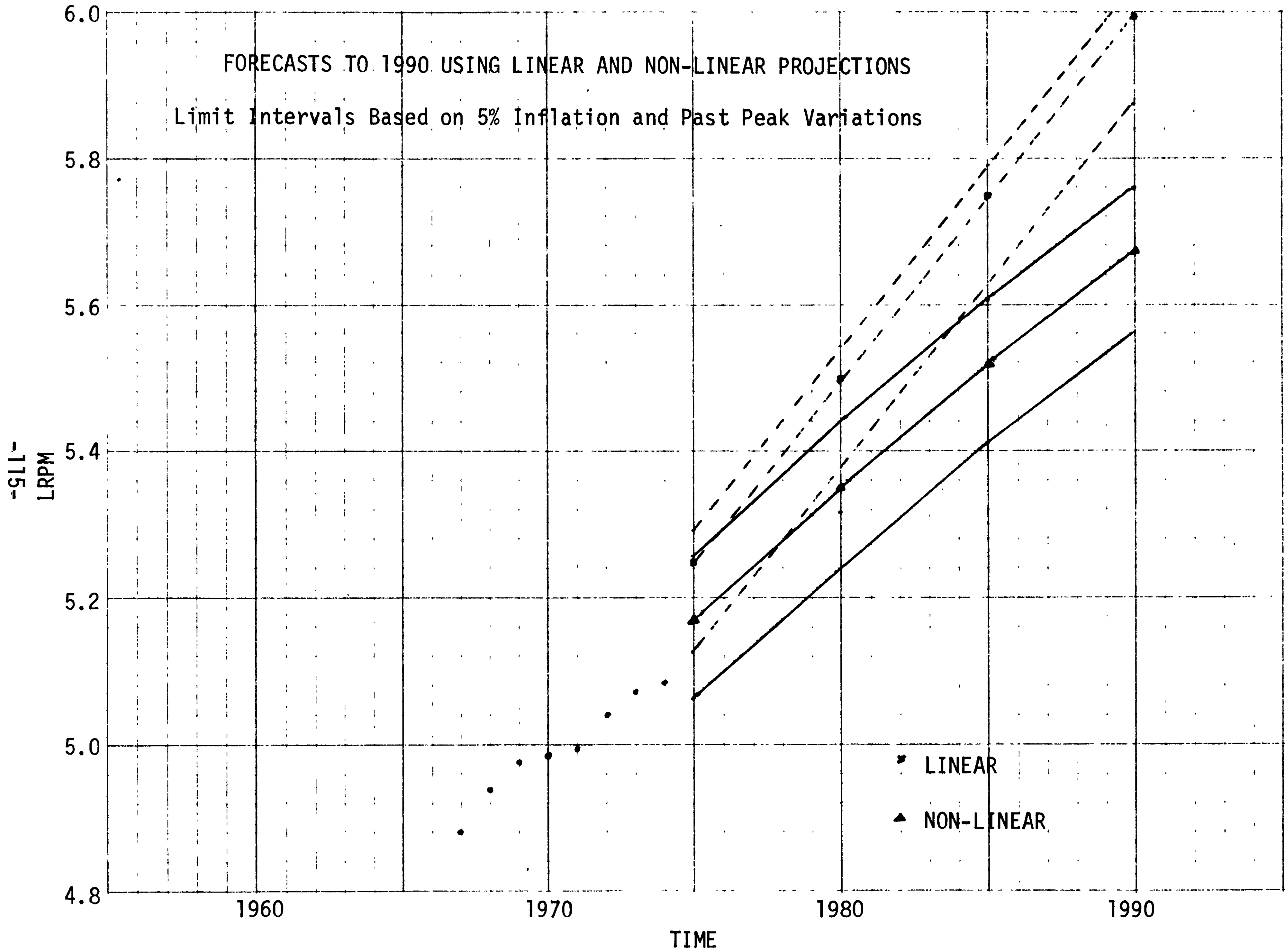

FORECASTS TO 1990 USING LINEAR AND NON-LINEAR PROJECTIONS
Limit Intervals Based on 5% Inflation and Past Peak Variations
LRPM
6.0
5.8
5.6
5.4
5.2
5.0
4.8
1960
1970
1980
1990
TIME
LINEAR
NON-LINEAR

FIGURE 12

REFERENCES

1. Civil Aeronautics Board, Handbook of Airline Statistics, 1973 Edition, Government Printing Office, Washington, D.C., 1974.

2. Quandt, R.E., The Demand for Travel: Theory and Measurement, Heath Lexington Books, Lexington, Mass., 1970.

3. Meyers, J.G., Methods and Techniques of Business Forecasting, Prentice-Hall, Inc., Englewood Cliffs, New Jersey , 1974.

4. Hooper, J.W. and Nerlove, M., Selected Reading in Econometrics from Econometrica, The MIT Press, Cambridge, Mass., 1970.

5. Ferber, R. and Verdoorn, P.J., Research Methods in Economics and Business, Macmillan Co., N.Y., 1962.

6. Meyer, J.R., and Straszheim, M.R., Pricing and Project Evaluation Vol. 1 of Techniques of Transport Planning, Brookings Institute, Washington, D.C., 1971.

7. Fisher, F.M., A Study in Econometrics: The Demand for Electricity in the United States, North-Holland Publishing Co., Amsterdam, 1962.

8. Balestra, P., The Demand for Natural Gas in the United States, North-Holland Publishing Co., Amsterdam, 1967.

9. Cotterman, W.W., "An Analysis of the Demand for Domestic Air Freight," Ph.D. Thesis, University Microfilsm, Inc., Ann Arbor, Michigan, 1969.

10. McKinnell, H.A., Jr., "An Econometric Analysis of the U.S. Air Freight Market," University Microfilms, Inc., Ann Arbor, Michigan, 1969.

11. Dunn, M.J., "A Statistical Analysis of the Demand for Domestic Air Passenger Transportation in Canada: 1960-1969," Ph.D. Thesis, University Microfilms, Ann Arbor, Michigan, 1972.

12. Alcaly, R.E., "An Economic Study of the Demand for Air Passenger Service Over the North Atlantic," University Microfilms Inc., Ann Arbor, Michigan, 1969.

13. Bartlett, H.C., "The Demand for Passenger Air Transportation: 1947-1962," University Microfilms, Inc., Ann Arbor, Michigan, 1965.

14. Watkins, W., Kaylor, D., and Richards, D., "Forecast of Scheduled Domestic Air Travel for the 50 States: 1972-1981," Civil Aeronautics Board Study, November 1972.

15. Aureille, Y.G., and Norris, C.T., "Short and Long Term Forecasting Models of the U.S. Domestic and International Traffic and Forecasts to 1981," Third Edition, Douglas Aircraft Co. Report No. C1-805-3084.

16. Verleger, P.K., Jr., "Models for the Demand for Air Transportation," The Bell Journal of Economics and Management Science, Autumn 1972, Volume 3, No. 2, pp. 437-457.

17. Lippke, B.R., and Stewart, J.R. "Economic Impact on Domestic Air Traffic, An Econometric Approach," Boeing Co. Paper No. 69-113, 1969.

18. Augustinus, J.G., "Air Traffic Forecasting at the Port Authority of New York and New Jersey," Port Authority, Aviation Economics Division, July 1972.

19. Port of New York Authority, "New York's Domestic Air Passenger Market," June 1967 - May 1968, Aviation Department, New York, N.Y.

20. Council of Economic Advisors, Economic Report of the President, February 1975, U.S. Government Printing Office, Washington, D.C.

21. Civil Aeronautic Board, Air Carrier Traffic Statistics, Monthly, Government Printing Office, Washington, D.C.