PNL-4825

UC-98

18 L

# Intervention Adjustment to Data of the Joint Petroleum Reporting System

D. L. Hall
T. K. Gardenier
A. L. Slavich

October 1984

Prepared for the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

Pacific Northwest Laboratory
Operated for the U.S. Department of Energy
by Battelle Memorial Institute

Battelle

PNL-4825

## DISCLAIMER

PACIFIC NORTHWEST LABORATORY
*operated by*
BATTELLE
*for the*
UNITED STATES DEPARTMENT OF ENERGY
*under Contract DE-AC06-76RLO 1830*

Printed in the United States of America
Available from
National Technical Information Service
United States Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22161

NTIS Price Codes
Microfiche A01

Printed Copy

| Pages | Price Codes |
|---|---|
| 001-025 | A02 |
| 026-050 | A03 |
| 051-075 | A04 |
| 076-100 | A05 |
| 101-125 | A06 |
| 126-150 | A07 |
| 151-175 | A08 |
| 176-200 | A09 |
| 201-225 | A010 |
| 226-250 | A011 |
| 251-275 | A012 |
| 276-300 | A013 |

PNL-4825
UC-98

# INTERVENTION ADJUSTMENT TO DATA OF THE JOINT PETROLEUM REPORTING SYSTEM

D.L. Hall
T.K. Gardenier[a]
A.L. Slavich

J.A. Mahaffey, Project Manager

October 1984

Prepared for
the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

Pacific Northwest Laboratory
Richland, Washington 99352

---

[a] Energy Information Administration

# CONTENTS

## FIGURES

# TABLES

# INTERVENTION ADJUSTMENT TO DATA OF THE JOINT PETROLEUM REPORTING SYSTEM

## 1.0 INTRODUCTION

In January, 1983 the Joint Petroleum Reporting System (JPRS) was replaced by the Petroleum Supply Reporting System (PSRS). The PSRS integrates petroleum data collection and includes all the previous weekly, monthly and annual reporting forms. Prior to January, 1983 the Joint Petroleum Reporting System (JPRS) contained diverse data relating to crude oil, and its refining, transport and storage, and also containing data relating to other U.S. petroleum products. The system was centered around the four monthly data collection forms listed in Table 1. The table also shows the identification numbers of previous and current forms and the dates when they were administered. Since the data used in this report end in December, 1982, only the JPRS forms are considered in the following.

TABLE 1. JPRS Forms by Reference Number and Date of Implementation

| Title | 1983 | 1979 to 1982 | 1975 to 1979 |
|---|---|---|---|
| Refinery Report | EIA-810 | EIA-87 | FEA-P320-MO |
| Bulk Terminal Stocks of Finished Petroleum Products | EIA-811 | EIA-88 | FEA-P321-MO |
| Pipeline Products Report | EIA-812 | EIA-89 | FEA-P322-MO |
| Crude Oil Stocks Report | EIA-813 | EIA-90 | FEA-P323-MO |

The above forms are used to obtain monthly data on receipts, storage, production and shipment from refiners, storers and transporters of petroleum products and crude oil and to obtain some information on company operations and geographical origin of crude oil. The data collected through the JPRS were used in such things as the analysis of supply problems and the estimation of the volume of petroleum products supplied for domestic use. JPRS thus produces data that is used as a basis for many policy and regulatory decisions. Many EIA publications are based on JPRS data, including:

| | |
|---|---|
| DOE/EIA-0008: | "Quarterly Report to Congress: Energy Information" |
| DOE/EIA-0011: | "Monthly Petroleum Status Report" |
| DOE/EIA-0035: | "Monthly Energy Review" |
| DOE/EIA-0105: | "Availability of Heavy Fuel Oils by Sulfur Levels" |
| DOE/EIA-0108: (also 0109) | "Crude Petroleum Products and Natural Gas Liquids" |
| DOE/EIA-0115: (also 0116) | "Supply, Disposition and Stocks of All Oils by States by County, 1978" |
| DOE/EIA-0173/2: | "Annual Report to Congress; Volume Two: Data" |
| DOE/EIA-0208: | "Weekly Petroleum Status Report" |

The four surveys in JPRS collect information at different terminal points or nodes in the petroleum industry and information relating to one type of product may be collected on more than one form. EIA-87, Refinery Report, contains information on production, processing, supply and disposition of petroleum and petroleum products. EIA-87 collects data specifically for Puerto Rico, Guam, Virgin Islands and the Hawaiian foreign trade zone. EIA-88, Bulk Terminal Stocks Report, deals primarily with the inventories of petroleum products at bulk terminals. EIA-9D, Crude Oil Stocks Report, records the inventory of crude oil stocks. The inventories of petroleum products in pipelines is contained in EIA-89, Products Pipeline Report. A flow diagram illustrating nodes in the distribution of petroleum products and the JPRS collection points is given in Figure 1.

Since the JPRS survey forms collect data at different nodes, the sampling frames for each form are separate. Moreover, for any one form, the respondents also differed over time since sampling frames were revised as companies entered, merged or went out of business. Table 2 compares the number of respondents as of June 1982 (as quoted in the JPRS system documentation prepared by Orkand Corporation) and in 1978 [as quoted in an evaluation report by Transportation and Economic Research Associates, Inc. (TERA)].

As changes were implemented into the survey forms, the frame of respondents occasionally was modified. Sometimes new companies were added to the survey forms without changing the forms. Thus, it is to be expected that the JPRS data series have changed over time. In the following paragraphs, some of these changes and their expected effect on the data series are discussed further.

There are several ways in which data series such as those collected by the JPRS may change over time as follows:

1.) Respondents may be added to or deleted from the sampling frame. Deletion may happen because respondents drop out of the survey (due to plant shutdowns, mergers with other companies or with components of companies, etc.). As additions, approximately 100 bulk terminals were added to the frame of the EIA-88 forms in 1977; blending stations were added to the frame of EIA-87 in 1981 and 1982.
2.) Changes in energy supply, demand or transportation modes may impact the values reported, although the survey forms or the frame have not changed. For example, opening the Alaskan pipeline in 1977 substantially escalated reporting of crude oil stocks from Alaska.

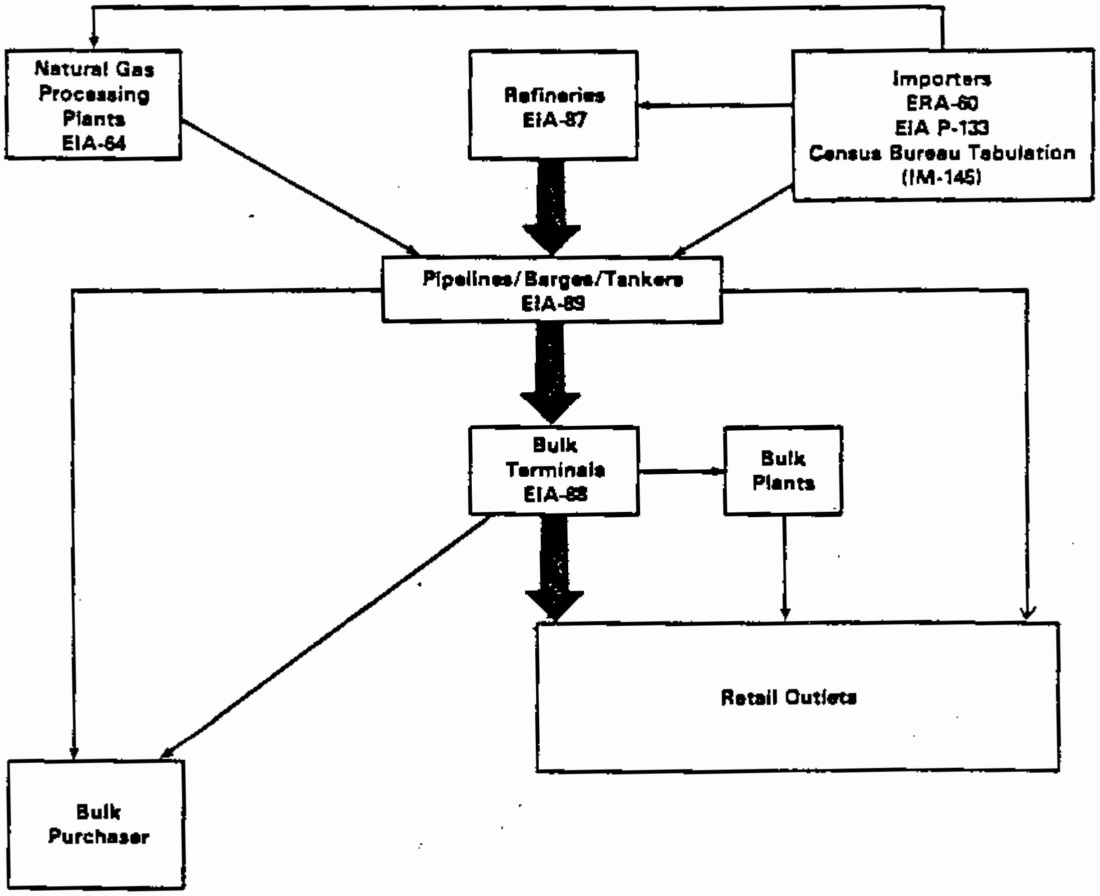

FIGURE 1. Simplified Flow Diagram of the Domestic Distribution of Petroleum Products and Measurement Points for the Reference Estimate

SOURCE: (Report in Progress) "An Assessment of the Principal Data Series of the Energy Information Administration"

TABLE 2. Number of Respondents by Date to JPRS Survey Forms

| | Number of Respondents/Month | |
|---|---|---|
| | 1982 | 1978 |
| EIA-87 | 353 | 319 |
| EIA-88 | 178 | 182 |
| EIA-89 | 78 | 77 |
| EIA-90 | 151 | 337 |

3.) The survey forms may be modified to correct for such things as misinterpretation, vagueness in wording or redefinition of products. For example, the country (if the origin of stocks was foreign) and/or state (if domestic) of origin of crude oil stocks often were not known by respondents to EIA-87 and EIA-90. Therefore, the 1981 version of EIA-87 and EIA-90 no longer requested this information.

4.) Procedures for computing quantities derived from the data collected on the forms may change. An example may be found in the way in which an apportionment of crude burned as fuel oil was made to residual and distillate fuel oils. In 1981 this adjustment was discontinued, since the relevant data were collected directly from that time on.

This report deals specifically with changes made to the survey forms in January 1981 and the resulting changes to the data series. Naturally, when a series has changed at some time point, the data after the change are no longer comparable to those before. In many cases, though, comparisons are desired that use pre- and post-intervention data as a series. It is thus necessary to have a methodolgy for "updating" the older data so that such comparisons can be made validly. To produce this methodology, the particular intervention must be modeled. However, when attempting to analyze one particular intervention, other types of interventions must be considered also. If effects of other interventions can be modeled, the overall variability of the series can be reduced and the intervention of interest can be better isolated. Thus, in the next section, we discuss (in addition to the format modifications of the forms) the trends and changes noted in the JPRS since January 1976 to December 1982. The year 1976 was chosen since it corresponds to the first year for which microdata are computerized in a "universal" format in the JPRS master files. We will discuss, in particular, changes to the data series for inventories of: a.) motor gasoline, b.) distillate oil, c.) residual fuel oil and d.) crude oil. These are the series studied in detail in subsequent sections of this report.

# 2.0 CHANGES IN THE JPRS

Since 1976 there have been several changes to the JPRS and in the petroleum industry as monitored by the JPRS that have affected the data series collected by forms EIA-87, EIA-88, EIA-89 and EIA-90. For instance, the Alaskan pipeline was completed in 1977, thus during the last six months of 1977, crude oil production in Alaska, as reported by the JPRS, jumped to an average of 737,000 barrels/day from 186,000 barrels/day. However, this report will focus on the major changes made to the JPRS in January, 1981.

In January 1981 several changes were made to the JPRS forms. These changes were made in response to an analysis of JPRS data, Federal Highway Administration (FHWA) data, and data on petroleum products from other external sources. It was concluded that changes had taken place in the petroleum industry and that these needed to be reflected in the EIA survey forms. In particular, the changes affected motor gasoline, distillate, residual fuel and crude oil, however, only the inventories of motor gasoline and crude oil should have been directly affected. These are discussed in greater detail in the following sections.

## 2.1 CHANGES FOR MOTOR GASOLINE

The changes to the survey items relating to motor gasoline were implemented after noting that the differences between JPRS and FHWA estimates of volume supplied for domestic use were steadily increasing. Prior to 1979 JPRS estimates of motor gasoline supply had been about 2% lower than those for FHWA. However, in 1979 the difference increased to about 4%, and in 1980 to about 5%. In May 1980 estimates by two other EIA surveys were also higher than those of JPRS; P-306 exceeded JPRS supply estimates by an average of 3.2%, EIA-25 by an average of 3.8%.

These discrepancies were attributed to two causes:

1.) The reported production of motor gasoline in refineries did not include blending component receipts which were blended into motor gasoline; e.g., butane or petrochemical feedstocks. These were treated as intermediate products and were not reported as motor gasoline. In many petrochemical plants, gasoline blending stocks were treated as a petrochemical processing byproduct and were sold to refineries which blended them to make gasoline. These receipts of intermediate products that were blended into finished gasoline were therefore not reported.
2.) Leaded motor gasoline was produced at downstream blending stations by adding tetra-ethyl lead to naphthas and butanes to raise the octane level. Environmental Protection Agency (EPA) regulations allowed the product of these stations to have five times as much lead content as the products of refiners. It was hypothesized that blending stations were taking advantage of these lead provisions and producing significant amounts of gasoline. The JPRS survey frame did not include blending facilities and thus would have missed their production and their stocks.

The following changes were made to JPRS to deal with the above issues:

1.) A new line "Gasoline Blending Components" was added to EIA-87 and gasoline was reported separately as leaded and unleaded.
2.) The open input column was expanded and respondents were instructed not to report blending components as unfinished oils. The open input

column also addressed the criticism that JPRS questions were oriented toward intended use rather than physical characteristics. In addition, companies were asked to report gross production and inputs. Thus a net production number was calculated instead of being reported directly by the companies.

3.) Many blending stations were added to the survey frame in January 1981. These accounted for an increase of about 2 to 3% in estimates of motor gasoline supplied for domestic use. As they were identified, more blenders were added: one in June 1981, four in October 1981 and nine in January 1982.

## 2.2 CHANGES FOR DISTILLATE AND RESIDUAL FUEL OIL

There were no specific changes in JPRS survey forms relating to distillate and residual fuel oil. However, in 1981 an adjustment formerly used to correct for a possible imbalance between unfinished oil supply and disposition was discontinued. When one refinery reported production of distillate or residual fuel oil and shipped it to another refinery, the latter refinery might have reported it as unfinished oil at receipt. This same fuel oil was then input and reprocessed for use or sale and was reported as whatever was produced. This practice produced a difference between supply and disposition of unfinished oil. Inputs to refineries, as reported, exceeded the reported supply. The adjustment mentioned earlier involved allocating 1/3 of this difference to residual fuel oil and 2/3 to distillate. Since this could not be supported by empirical evidence and since the reporting of inputs by product was intended to eliminate this difference, this adjustment was discontinued.

Since the series under consideration in this study deal with estimates of stocks rather than production and inputs, this data adjustment issue was not relevant to the major intervention adjustment analyses.

## 2.3 CHANGES FOR CRUDE OIL

In January 1981, while the motor gasoline changes were introduced to JPRS forms, another change was implemented to items relating to crude oil in EIA-87 and EIA-90. Many companies did not know the origin of crude oil held by them in different locations. Thus, EIA-87 and EIA-90 deleted the state of origin of domestic crude oil, and EIA-87 deleted the country of origin of foreign crude and unfinished oils. Alaskan crude oil in transit by water was reported separately on EIA-90 beginning in January 1981. In addition crude oil stocks at refineries had been reported on both EIA-87 and EIA-90 and the published data were taken from EIA-90. In January, 1981 refinery stocks were deleted from EIA and the published values were taken from EIA-87. This resulted in some confusion.

## 2.4 OTHER CHANGES TO JPRS

There were several other changes implemented to the JPRS in January 1981. Some of these did not deal specifically with the data series under consideration. However, these changes are included here for completeness.

1.) Upon checking operator names through a list prepared by the Maritime Administration, frame deficiencies were noted in the frames for EIA-88 and EIA-90, that is bulk terminals and crude stock holders. Marine terminal operators from this list, which were not introduced previously, were added to EIA-88 and EIA-90 after checking their eligibility.

2.) Crude oil stocks in transit by modes of transportation other than pipeline provided a vague category. Respondents indicated that they did not know specifics about shipment and expected arrival dates. Most did not report in-transit stocks (mostly crude oil). This lead to underestimating the stock level. A suggestion was made to combine in-transit stocks by mode into an overall category of "total estimated stocks in transit other than by pipeline".

3.) On EIA-87 the "shipments" category was separated from "refinery fuel use and losses". This was because it was observed that the latter were absorbed in production and were not reported. Shipments were, in fact, the only data items reported for that question.

## 3.0 THE SERIES TO BE ANALYZED

This section outlines the data series studied in the intervention analysis. Included in this section is a general description of each series. As noted previously, the four series under consideration are reported inventory levels of motor gasoline, distillate fuel oil, residual fuel oil and crude oil.

1.) Motor Gasoline: Total motor gasoline consists of finished leaded and unleaded motor gasoline, blending components of motor gasoline, and gasohol. Finished motor gasoline is composed of volatile hydrocarbons with or without additives which are blended to form a fuel. The 1981 survey forms do not provide a "total" for motor gasoline as the previous forms did. Previous forms did not provide the needed detail. Thus, the four components are tracked after 1981 in order to derive a total.

2.) Distillate Fuel Oil: Total distillate fuel oil consists of petroleum fractions and includes No. 1 and No. 2 heating oils, No. 1 and No. 2 diesel fuel oils and No. 4 fuel oil. No. 1 oils are used in vaporizing pot-type burners. No. 2 oils are used for domestic heating and in industrial burners. No. 4 oil is used in commercial burners which do not have preheating facilities. In general, distillate fuel oil is used in space heating, as highway diesel engine fuel, for agricultural machinery and in electric power generation.

3.) Residual Fuel Oil: Residual fuel oil includes No. 5 and No. 6 fuel oils and is used for space heating, industrial purposes and in electric power generation.

4.) Crude Oil: Crude oil consists of a mixture of hydrocarbons which are in liquid phase while in underground reservoirs and which remain in liquid phase under atmospheric pressure after passing through surface separating facilities.

The components for these four series exist in several locations on Survey Forms EIA-87, EIA-88, EIA-89 and EIA-90.

These series have attracted attention both from policy makers and from the general public. Stocks of these four series are tracked weekly in the Weekly Petroleum Status Report, showing the actual data as well as plots of average, range, minimum and seasonal patterns.

Production and refinery stock levels from these series for 1976-1978 were submitted for preliminary analyses in the study validating the Joint Petroleum Reporting System (JPRS) (ORNL 1980). Within the EIA, autoregressive integrated moving average (ARIMA) modeling and seasonal adjustment using the X-11 method have been applied, particularly to crude oil data dating back to 1967. Seasonality has been noted for data relating to distillate and residual fuel oil. The "State of Data" report, a study in progress at EIA for assessing various EIA data series and comparing them to external sources, also used estimates of total volume supplied to the domestic market for motor gasoline, distillate, kerosene and residual fuel oil. This study showed that motor gasoline monthly estimates of price as well as volume had relatively low variance a.) for data representing the pool of respondents per month and b.) for comparisons with external data sets. An exception was noted at intervention points. On the contrary, estimates for residual fuel oil were, in comparison, unstable from month to month. The two series (motor gasoline and residual fuel oil) thus represent series with essentially different characteristics over time.

As discussed previously, the 1981 motor gasoline interventions were major; yet the series were relatively free of intervention prior to 1981 in terms of changes related to items in the survey forms. Similarly, the changes in 1981 could have affected inventories of crude oil. However, the series for residual fuel oil and distillate fuel oil should not have reflected the 1981 interventions.

# 4.0 INTERVENTION ANALYSIS TECHNIQUES

Of the many intervention analysis techniques only a few appear promising for the problem at hand. This section furnishes a short description of several intervention analysis techniques and comments on their potential applicability to the inventory data series.

## 4.1 BOX-JENKINS/BOX-TIAO MODELING AND INTERVENTION ANALYSIS

This method for time-series modeling and intervention adjustment is outlined in the references by Hibbs (1977), Box and Tiao (1975, 1976) and Tiao, Box and Hamming (1975). Time series modeling can be performed on the EIA computer using available computer routines in the Statistical Analysis System (SAS) Econometric Time-Series Package (ETS).

A classical analysis by these methods proceeds by modeling the pre- and post-intervention series and testing if there are differences between the two sets of parameters. This requires that there be considerable amounts of data both before and after the intervention. In the present situation this is not the case, since there are only 60 observations before the intervention and 24 after.

As an alternative, the full series may be analyzed with transfer function techniques, using an additional intervention-related "dummy variable". Based on knowledge of the intervention, a model is fitted and tested for appropriateness. The process is then repeated until an acceptable model is obtained. In general, a considerable amount of data also is required for this approach to be successful.

Box and Tiao (1976) illustrated one method that seems particularly appropriate for the JPRS data. A model is fitted to the pre-intervention data and, based on this model, forecasts are made for the post-intervention period. The residuals from the forecasts are used to test if the original model fits the post-intervention data. If the test indicates that a change has taken place, residuals can be computed from a class of possible post-intervention models and compared to the residuals from the forecasts. Patterns in the residuals suggest possible changes, and least squares can be used to estimate the magnitude of the changes.

To be more specific, suppose that some known change occurs after time $t_0$ so that there are observations $z_1, \ldots, z_{t_0}$ before the intervention and $z_{t_0+1}, \ldots, z_N$ after the intervention. Further, suppose that the pre-intervention time series is modeled as

$$\Phi(B)\ w_t = \theta(B)\ a_t \tag{1}$$

where the $w_t$ are some differences of the $z_t$'s,

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p,$$

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

and where B is the backshift operator defined by $BX_t = X_{t-1}$. It is assumed that $\Phi(B)$ and $\theta(B)$ have all of their zeroes outside the unit circle and have no common zeroes and that $\{a_t\}$ is a sequence of independent, identically distributed random variables with mean zero and variance $\sigma^2$ (white noise). Note that some of the $\phi_i$'s or $\theta_i$'s may be zero or otherwise restrained, so that the model could be a multiplicative seasonal model. By performing the indicated divisions (at least formally) the above model may be written alternatively as

$$z_t = \psi(B)\, a_t \tag{2}$$

or

$$\pi(B)\, z_t = a_t \,. \tag{3}$$

If the $\hat{w}_t(1)$ is taken as the minimum mean-square error forecast of $w_{t+1}$ given $w_1, \ldots, w_t$, then it is well known that

$$w_t - \hat{w}_t(1) = a_t;$$

that is, the $a_t$ are also the one-step-ahead forecast errors.

Following Box and Tiao (1976) an overall statistic, Q, to test for an intervention can be derived based on a standardized sum of the $a_i^2$. Q is defined as

$$Q = \sigma^{-2} \sum_{i=1}^{m} a_{t_o+i}^2 \tag{4}$$

where $t_o+m = N$ and where $\sigma$ is the standard deviation of the residuals from the pre-intervention period. The appropriateness of the model for the post-intervention period may be tested by comparing the value of Q to a Chi-square with m degrees of freedom. When the number of pre-intervention observations, n, is small, Box and Tiao suggest that the value of Q/m be referred to an F table with m and N - p degrees of freedom. This test is also equivalent to an appropriate test for intervention applied to all of the lead l forecast errors for l = 1, ..., m.

The statistic Q may be decomposed into components to study separate effects of different types of hypothesized interventions. For example the following situations may be examined: a.) an overall change in mean level, and b.) a change in one or more of the stochastic parameters. The authors suggest that the first type of effect may be examined by substituting

$$z_t - \beta x_t \tag{5}$$

for $z_t$ in (3). $\beta$ is a parameter to measure the shift in mean and the $x_t$ are indicator variables related to the intervention. Thus, the $x_t$ are 0 before the intervention and are 1 after. The values of $\pi(B)x_t = X_t$ for $t = t_0+1, \ldots, N$ are compared with the series of one step forecast errors, $a_t$, to determine if any patterns match.

A change in a stochastic parameter is explored by differentiating the error series with respect to the given parameter. For example, if it is desired to examine the parameter $\theta_1$, the series

$$W_{1t} = \left.\frac{-\partial a_t}{\partial \theta_1}\right|_0 \tag{6}$$

is formed for $t = t_0+1, \ldots, N$. This series of predicted residuals is compared with the one-step forecast errors to look for patterns. After possible interventions have been identifed with these methods, the one-step forecast errors can be regressed on the appropriate $X_t$'s and $W_t$'s to estimate the magnitudes of the effects.

That is, an approximate model for the $a_t$'s, assuming one level parameter and two stochastic parameters is

$$a_t \cong X_t + (\theta_1' - \theta_1)\, W_{1t} + (\theta_2' - \theta_2)\, W_{2t} + n_t$$

where the $X_t$'s and $W_{it}$'s are as above, the $\eta_t$ are some iid errors and the primes indicate post-intervention parameter values. Thus this model may be fit by least squares to produce $\beta$, $\theta_1' - \theta_1$ and $\theta_2' - \theta_2$ which can then be used to estimate the post-intervention stochastic parameters.

## 4.2 ANALYSIS BY CRITICAL SURVEY ITEMS AND BY COMPANY TYPOLOGY

A second procedure for intervention adjustment is to compare the change attributable to each of the "critical" survey questions with the overall change noted for estimates of petroleum supplied for domestic use. "Critical" survey questions may be defined as those that were affected most by the intervention; e.g., those which were added after 1981 or those whose definitions or structure were significantly modified. Some of these were described previously.

The structure of these analyses may involve the following: a) an analysis of variance of pre- and post-intervention data indexed by time, responses to specific survey items and estimate of total volume of that product supplied for domestic use, and b) a simultaneous check of which specific companies responded at each time period to insure that overall shifts in pattern are not due to a respondee profile of additions/deletions from the frame. Then, the adjustment for intervention would be item-specific values, and adjustment would be only for those items which significantly impact total volume.

An analysis by "company typology" is also possible. If there are certain types of companies or certain geographical regions to which changes can be attributed, intervention adjustment would be feasible only for these subsets of the JPRS data.

Allusions to this type of analysis are made in the paper by Turk (1978).

## 4.3 REGRESSION-BASED FRAMEWORK USING SAS

This approach would involve implementing the SAS procedure REG to estimate the intercept and slope of the linear regression line fitted to the pre- and post-intervention data. Then, the residuals from regression are analyzed to detect abrupt or gradual changes in the series attributable to the intervention. The size of the intervention adjustment would be determined using the coefficients for slope, intercept and the average magnitude of the residuals in the proximity of the intervention.

The difficulty with this approach lies in the fact that successive data are autocorrelated. Also, it is difficult to establish a cutoff for intervention in the case of noninstantaneous interventions. The concept of a "washout" as used in clinical trials for pre- and post-intervention would not be feasible.

Thus, this method appears to offer no advantages over Box-Tiao modeling.

## 4.4 MODELING AND ADJUSTMENT FOR INTERVENTION USING CTSS

This criterion is based upon deriving an index of pre- and post-intervention change using a summation procedure obtained from transitions in successive observations. The Cumulative Transitional State Score (CTSS) and its characteristics are described in the paper by Gardenier (1979). The "states" correspond to partitions in alternative control regions; the regions are derived from the distributional characteristics on the previously observed trends.

The methodology of CTSS is reminiscent of cumulative sum (CUSUM) techniques used in quality control. However, CUSUM techniques aggregate deviations from a long-term process average while CTSS sums transitions between successive states. If proximity to an overall process mean is to be maintained and is considered favorable, then the transition toward the mean carries more weight than deviations away from the mean. The sum CTSS would yield a summary index for the post-intervention or post-stabilization change.

The magnitude of the intervention would be derived using the CTSS score and the range of the boundaries for the alternative control regions.

## 4.5 CAUSAL ANALYSIS OF INTERVENTION THROUGH STRUCTURAL EQUATION MODELS

These procedures are discussed in the bibliography references by Hibbs (1977) and Zellner and Palm (1974). The advantage of structural equation models is that they provide a semi-causal analysis by relating exogenous policy parameters or interventions to endogenous target variables. If data could be obtained for variables associated with the interventions, structural equation models would provide a very useful approach.

A serious difficulty to the use of structural equation models is the lack of external data of a time-series nature which could be incorporated into the model.

## 4.6 METHODS BY AKAIKE AND GRAY ET AL. FOR DIRECT ESTIMATION OF p, d, q

These methods provide a way of directly estimating the parameters of autoregressive (p), differencing (d) and moving average (q) parameters of the autoregressive integrated moving average (ARIMA) models. In themselves, these methods do not provide a separate method of intervention adjustment, but they do provide a different approach to ARIMA-based methods. The TIMSAC-78 programs that implement some of the techniques of Akaike are available on the EIA computer.

An advantage of these procedures is that they should save time. Furthermore, they eliminate the visual inspection of the autocorrelation and partial autocorrelation functions that is typically used to select the most appropriate model parameters.

## 4.7 METHODS TO BE USED ON THE JPRS DATA

Based on the availability of suplementary data, the length of the data series and the availability of computer programs, the main methods chosen for analyzing the JPRS data are the Box-Tiao and CTSS techniques. The methods of Akaike will be used as a check of the models obtained through Box-Tiao modeling procedures. The next section contains the results of these Box-Tiao analyses. The results of the CTSS modeling are contained in Section 6.

# 5.0 RESULTS OF THE BOX-TIAO MODELING

As noted above the four series chosen for analysis were inventories of distillate fuel oil, residual fuel oil, motor gasoline and crude petroleum for the years 1976 to 1982. The major intervention of interest was the change in the survey forms implemented in January 1981. The following sections discuss the analysis of each of these series.

## 5.1 DISTILLATE FUEL OIL

Though data were available on total distillate fuel oil minus No. 4 fuel oil and on No. 4 fuel oil by itself, the series analyzed consisted of the total monthly stocks of distillate fuel oil. A plot of this data is given in Figure 2, and Figure 3 gives the number of companies used in computing this data. The intervention of interest occurs at the 61st data point signified by an arrow in Figure 2. As stated above, the changes to the survey forms should not affect the distillate series directly in any significant way.

From a visual inspection of the series, several aspects seem notable. First, there is a strong seasonality with a period of 12. Second, there appears to be a general downward trend. Third, the series seems to have changed at time point 49 when the magnitude of the seasonal swings shows a marked decrease. This last observation is of interest since it occurs a full year before the changes to the survey. Discussions with EIA personnel related this apparent shift to a shift in the general economic climate and in usage patterns of distillate fuel oil. Finally, the number of companies declines from time point 61 (January 1981) on. If the apparent change in the series at point 61 were simply due to the number of companies, Figure 4, a time-series

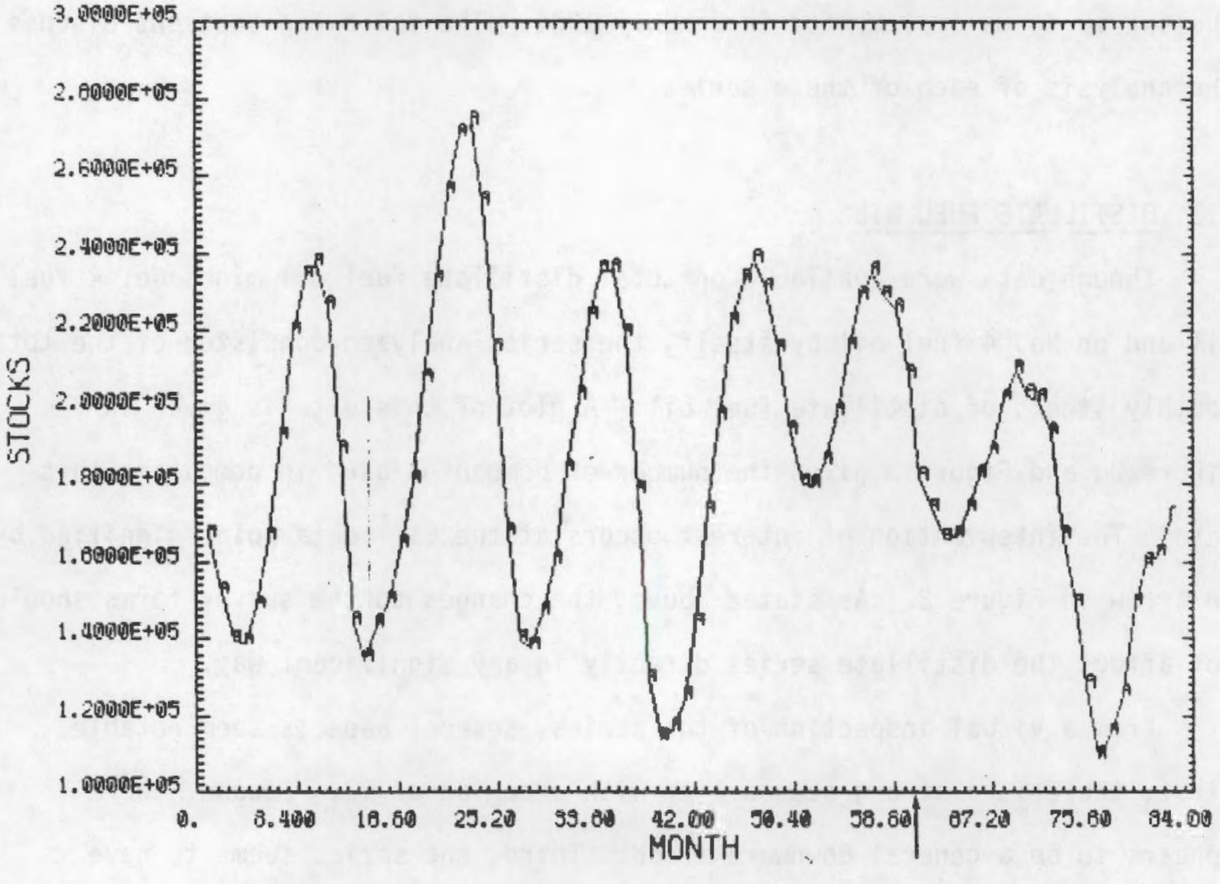

FIGURE 2. Total Distillate Stocks by Month (arrow denotes intervention)

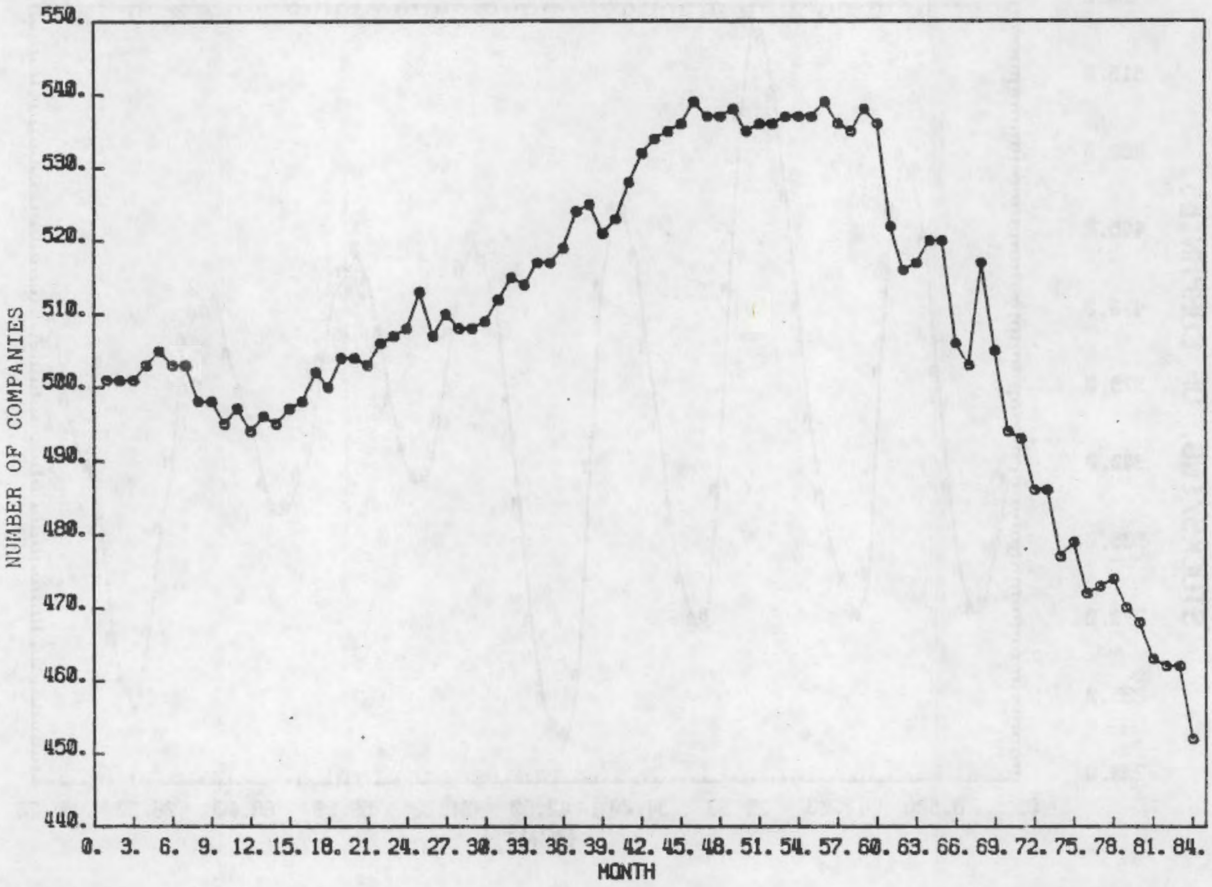

FIGURE 3. Number of Companies Used in Computing Total Distillate Stocks

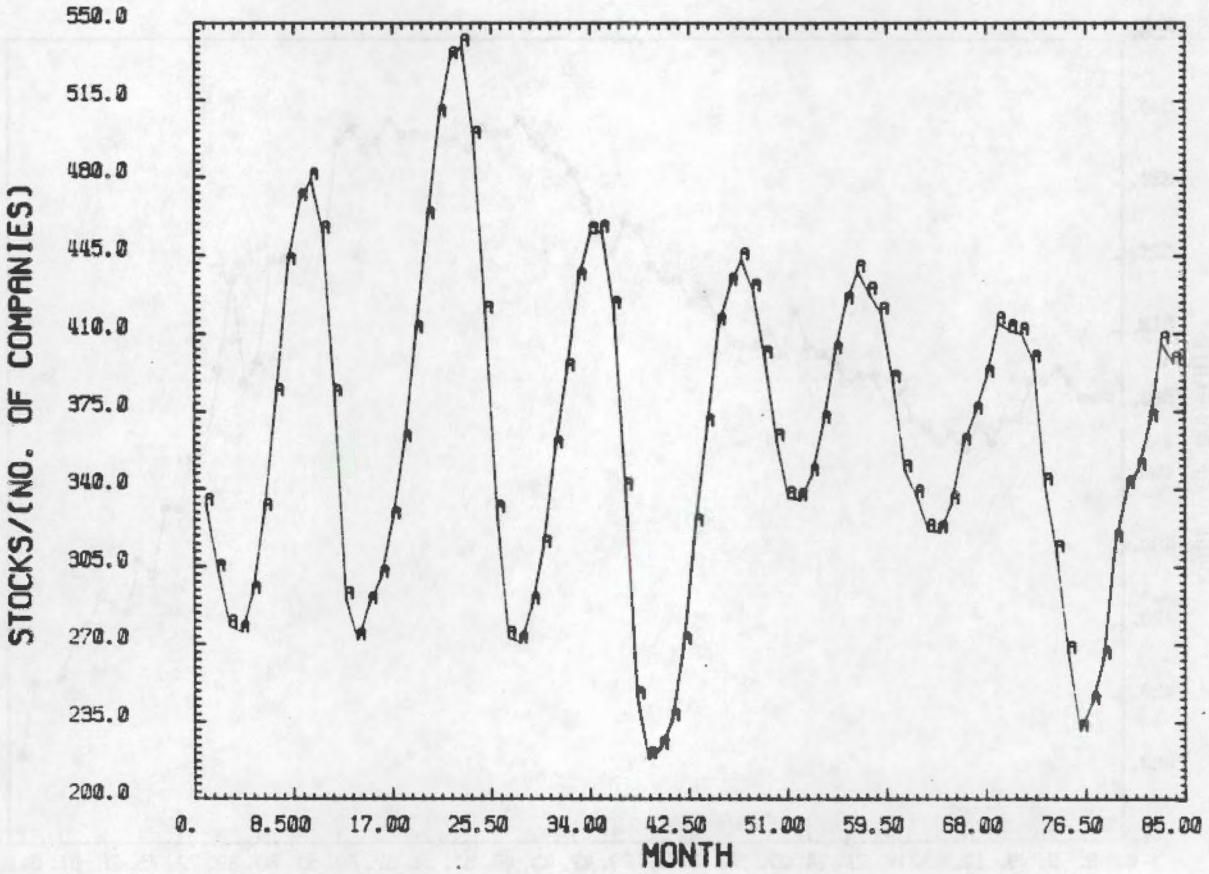

FIGURE 4. Total Distillate Stocks Divided by Number of Companies

plot of stocks divided by the number of companies, should not show the downturn. Since the change is apparent in this plot also (in fact, this plot is almost identical to Figure 2), the number of companies does not by itself explain the change.

To examine further the possible effect of the number of companies on distillate stocks several methods were tried. Following Bell and Hilmer (1983) the model

$$Z_t = \beta X_t + N_t \quad , \tag{7}$$

where $X_t$ is the number of companies reporting at time t and $N_t$ is an error series assumed to have some ARIMA structure, was fit by the process:

1. Use the $Z_t$ series and autocorrelation function (ACF) to decide on the differencing required to achieve stationarity, in this case orders 1 and 12.
2. Estimate $\beta$ from the least-squares fit of the model

   $$(1 - B)(1 - B^{12})\ Z_t = \beta(1 - B)(1 - B^{12})\ X_t + \varepsilon_t \ . \tag{8}$$

   where $\varepsilon_t$ is the differenced error series.
3. Model the residuals $\hat{\varepsilon}_t$ from the above fit by standard Box-Jenkins techniques.

When these steps were performed on the distillate series for times 1 to 84 the least-squares fit of $\beta$ produced a nonsignificant estimate. This method thus produced no evidence of a significant effect of the number of companies. After a discussion of the modeling of the distillate series other attempts at assessing an effect of the number of companies will be addressed.

The modeling of the distillate series proceeded along classical Box-Jenkins lines. Because of the intervention after time 60 and the apparent change at time 49, the series of times 1 to 48 and 1 to 60 were analyzed separately. The ACF indicated a need for differencing of orders 1 and 12. The ACF, partial autocorrelation function (PACF) and the inverse autocorrelation function (IACF) for the differenced series (orders 1 and 12) are given in Figures 5 to 10. As can be seen, a multiplicative integrated moving average model of orders 1 and 12 is suggested. Several candidate models were estimated, with the model (0,1,1) x $(0,1,1)^{12}$ producing the best fit.

```
AUTOCORRELATIONS
LAG COVARIANCE CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1        STD
  0  133705329    1.00000  |                    |********************|        0
  1   72953071    0.54563  |              .     |***********         |   0.169031
  2    7363473    0.05507  |            .       |*         .         |   0.213502
  3  -11193024   -0.08371  |            .     **|          .         |   0.213908
  4  -18820440   -0.14076  |            .    ***|          .         |   0.214642
  5  -12267954   -0.09175  |            .     **|          .         |   0.217461
  6    7962579    0.05955  |            .       |*         .         |   0.218564
  7   10775838    0.08059  |            .       |**        .         |   0.219027
  8   16772287    0.12544  |            .       |***       .         |   0.219973
  9    6622515    0.04953  |            .       |*         .         |   0.221908
 10  -15768910   -0.11794  |            .     **|          .         |   0.222224
 11  -33403663   -0.24983  |            .  *****|          .         |   0.224005
 12  -43015453   -0.32172  |            . ******|          .         |    0.23183
 13  -34084401   -0.25492  |           .   *****|           .        |   0.244253
 14  -11012687   -0.08237  |           .      **|           .        |    0.25174
 15   -6037098   -0.04515  |           .       *|           .        |   0.252508
 16  -13640324   -0.10202  |           .      **|           .        |   0.252739
 17  -12412832   -0.09284  |           .      **|           .        |   0.253913
 18   -7579135   -0.05669  |           .       *|           .        |   0.254881
 19   -2484197   -0.01858  |           .        |           .        |   0.255241
 20   -1952814   -0.01461  |           .        |           .        |   0.255279
 21   -8741665   -0.06538  |           .       *|           .        |   0.255303
 22   -7808912   -0.05840  |           .       *|           .            0.255781
 23    4610042    0.03448  |           .        |*          .            0.256162
 24   11480815    0.08587  |           .        |**         .            0.256294
                           '.' MARKS TWO STANDARD ERRORS
```

FIGURE 5. Plot of Autocorrelation Function for Months 1 to 48 for Differenced (1 and 12) Distillate Fuel Oil Series

```
PARTIAL AUTOCORRELATIONS
           LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
             1     0.54563 |            .      |***********        |
             2    -0.34549 |            *******|      .            |
             3     0.10407 |            .      |**    .            |
             4    -0.17955 |            .  ****|      .            |
             5     0.10512 |            .      |**    .            |
             6     0.07137 |            .      |*     .            |
             7    -0.06047 |            .     *|      .            |
             8     0.19434 |            .      |****  .            |
             9    -0.21828 |            .  ****|      .            |
            10     0.00018 |            .      |      .            |
            11    -0.22282 |            .  ****|      .            |
            12    -0.14412 |            .   ***|      .            |
            13     0.01404 |            .      |      .            |
            14    -0.09396 |            .    **|      .            |
            15    -0.07234 |            .     *|      .            |
            16    -0.16136 |            .   ***|      .            |
            17     0.03561 |            .      |*     .            |
            18    -0.01038 |            .      |      .            |
            19     0.02973 |            .      |*     .            |
            20    -0.04845 |            .     *|      .            |
            21    -0.11165 |            .    **|      .            |
            22    -0.00264 |            .      |      .            |
            23    -0.01523 |            .      |      .            |
            24    -0.00575 |            .      |      .            |
```

FIGURE 6. Plot of Partial Autocorrelation Function for Months 1 to 48 for Differenced (1 and 12) Distillate Fuel Oil Series

```
INVERSE AUTOCORRELATIONS
     LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
       1   -0.74180  |     ***************|      .             |
       2    0.46102  |             .      |*********           |
       3   -0.15021  |             .   ***|      .             |
       4   -0.10345  |             .    **|      .             |
       5    0.34175  |             .      |*******             |
       6   -0.45117  |           *********|      .             |
       7    0.44279  |             .      |*********           |
       8   -0.34527  |             *******|      .             |
       9    0.19020  |             .      |****  .             |
      10   -0.03749  |             .     *|      .             |
      11   -0.05818  |             .     *|      .             |
      12    0.15007  |             .      |***   .             |
      13   -0.11403  |             .    **|      .             |
      14    0.11246  |             .      |**    .             |
      15   -0.08667  |             .    **|      .             |
      16    0.07336  |             .      |*     .             |
      17   -0.01191  |             .      |      .             |
```

FIGURE 7. Plot of Inverse Autocorrelation Function for Months 1 to 48 for Differenced (1 and 12) Distillate Fuel Oil Series

```
AUTOCORRELATIONS
LAG COVARIANCE CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1        STD
  0  154054762     1.00000 |                    |********************|          0
  1   96140924     0.62407 |              .     |************        |  0.145565
  2   30434980     0.19756 |            .       |****   .            |  0.194549
  3    5339639     0.03466 |            .       |*      .            |  0.198772
  4  -10075118    -0.06540 |            .      *|       .            |  0.198901
  5  -10496848    -0.06814 |            .      *|       .            |  0.199358
  6    -817274    -0.00531 |            .       |       .            |  0.199852
  7   -5956811    -0.03867 |            .      *|       .            |  0.199855
  8   -8409286    -0.05459 |            .      *|       .            |  0.200015
  9  -14309492    -0.09289 |            .     **|       .            |  0.200331
 10  -21610967    -0.14028 |            .    ***|       .            |  0.201246
 11  -31078930    -0.20174 |            .   ****|       .            |  0.203315
 12  -43662728    -0.28342 |            . ******|       .            |  0.207531
 13  -34312740    -0.22273 |           .    ****|        .           |  0.215609
 14  -14430996    -0.09367 |           .      **|        .           |   0.22045
 15  -16130427    -0.10471 |           .      **|        .           |  0.221295
 16  -21262398    -0.13802 |           .     ***|        .           |  0.222347
 17  -19804394    -0.12855 |           .     ***|        .           |  0.224162
 18  -17352158    -0.11264 |           .      **|        .           |  0.225726
 19   -9069643    -0.05887 |           .       *|        .           |  0.226918
 20   -6691027    -0.04343 |           .       *|        .           |  0.227243
 21  -13782855    -0.08947 |           .      **|        .           |   0.22742
 22  -25094087    -0.16289 |           .     ***|        .           |  0.228167
 23  -31046105    -0.20153 |           .    ****|        .           |  0.230628
 24   -4161217    -0.02701 |           .       *|        .           |  0.234345
                           '.' MARKS TWO STANDARD ERRORS
```

FIGURE 8. Plot of Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Distillate Fuel Oil Series

```
PARTIAL AUTOCORRELATIONS
          LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
            1   0.62407   |              .     |************        |
            2  -0.31432   |              ******|     .              |
            3   0.12499   |              .     |**   .              |
            4  -0.16773   |              .  ***|     .              |
            5   0.10984   |              .     |**   .              |
            6  -0.01021   |              .     |     .              |
            7  -0.10719   |              .   **|     .              |
            8   0.05841   |              .     |*    .              |
            9  -0.16136   |              .  ***|     .              |
           10   0.01276   |              .     |     .              |
           11  -0.20100   |              . ****|     .              |
           12  -0.13108   |              .  ***|     .              |
           13   0.09349   |              .     |**   .              |
           14  -0.07660   |              .   **|     .              |
           15  -0.12787   |              .  ***|     .              |
           16  -0.07682   |              .   **|     .              |
           17  -0.05777   |              .    *|     .              |
           18  -0.01353   |              .     |     .              |
           19  -0.05839   |              .    *|     .              |
           20  -0.12163   |              .   **|     .              |
           21  -0.11572   |              .   **|     .              |
           22  -0.23290   |              .*****|     .              |
           23  -0.13903   |              .  ***|     .              |
           24   0.15405   |              .     |***  .              |
```

FIGURE 9. Plot of Partial Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Oistillate Fuel Oil Series

```
INVERSE AUTOCORRELATIONS
             LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
               1   -0.64719   |       *************|     .              |
               2    0.38617   |              .     |********            |
               3   -0.09984   |              .   **|     .              |
               4   -0.01974   |              .     |     .              |
               5    0.24303   |              .     |*****.              |
               6   -0.24336   |              .*****|     .              |
               7    0.25234   |              .     |*****.              |
               8   -0.09527   |              .   **|     .              |
               9    0.06703   |              .     |*    .              |
              10    0.03610   |              .     |*    .              |
              11   -0.01715   |              .     |     .              |
              12    0.09100   |              .     |**   .              |
              13    0.04125   |              .     |*    .              |
              14   -0.04726   |              .    *|     .              |
              15    0.04916   |              .     |*    .              |
              16    0.03404   |              .     |*    .              |
              17   -0.01969   |              .     |     .              |
              18    0.07071   |              .     |*    .              |
              19   -0.04609   |              .    *|     .              |
              20    0.06054   |              .     |*    .              |
              21   -0.01385   |              .     |     .              |
              22    0.00179   |              .     |     .              |
              23    0.05886   |              .     |*    .              |
```

FIGURE 10. Plot of Inverse Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Distillate Fuel Oil Series

The results of the estimation are given in Table 3.

The ACF and PACF of the residuals indicated no lack-of-fit for either model; thus, both were tentatively accepted. As a further check on these models, autoregressive models were fit to the differenced data using Akaike's AIC procedure to automatically choose the order of the model that minimized the AIC. For the models given above, we would expect the order chosen to be in the neighborhood of 13. For the time 1 to 48 series, the order chosen was 14. The time 1 to 60 series did not behave as well and the order obtained was 1.

TABLE 3. Summaries from Modeling Distillate Fuel Oil

Series 1 to 48

Differences 1,12

| Type | Value | S.D. | t |
|---|---|---|---|
| MA1 | -0.726 | 0.127 | -5.72 |
| SMA1(12) | 0.781 | 0.182 | 4.28 |

Residual Mean Square = 54956992

Series 1 to 60

Differences 1,12

| Type | Value | S.D. | t |
|---|---|---|---|
| MA1 | -0.740 | 0.106 | -6.95 |
| SMA1(12) | 0.760 | 0.156 | 4.86 |

Residual Mean Square = 50379664

However, the plot of AIC versus order exhibited a dip at 13, after which time it continued to increase. Thus the AIC procedure furnished some evidence to accept the models.

Before examining the intervention, some other methods were tried to relate the apparent downturn to the decline in the number of companies after time point 60. First, using the model developed above for the series 1 to 60, the one-step forecast error for the times 61 to 84 were regressed on the number of companies reporting for these periods. This regression was not significant. The series 1 to 84 was considered by first modeling the time series and then regressing the residuals from the model on the number of companies. This regression was also not significant. Finally, distillate stocks were regressed on the number of companies and the residuals from the regression examined. These residuals showed the same structure, i.e., downturn, as the original series. It was thus concluded that merely the number of companies reporting does not explain the change in the series. Any possible company "effect" is more complicated and to begin to understand it would require an examination of the types of companies reporting before and after January 1981. It could also be argued that the decline in the number of companies is caused by, rather than caused, the change in the distillate series. Thus the apparent changes will be studied in the following.

Since there is not much post-intervention data, the method outlined in Box and Tiao (1976) was used to study the possible interventions at 49 and 61. The Q statistic, described previously, when calculated on the 1 to 48 series for the period 49 to 60 resulted in a value of 9.62 which is not significant at even the 0.1 level. However, when calculated on the same series for the period 49 to 84, Q was 47.55 which is very nearly significant at the 0.10 level. Thus there was a nearly significant change in the series at the time point 49. Moreover, when the 1 to 60 series was considered, the Q statistic for the period 61 to 84 is 43.31 which is significant at the .01 level. Taken by itself this could be interpreted as a significant change in the data at time point 61.

Assuming for the moment that the change at time point 61 is "real," can the nature of this change be determined? Figures 11 to 13 show the X, W1 and W12 series plotted with the one-step forecast errors to examine possible changes in level or in the parameters. These residual plots gave some weak indication that both the level and the first order moving average term changed, but when the magnitudes of the changes were estimated with least squares neither of the coefficients was significant. Thus these methods did not clearly identify the nature of the intervention.

There is considerable uncertainty as to how the "changes" in the data at the points 49 and 61 can be interpreted. Recall that the changes in the survey should have had no important direct effect on the distillate fuel oil data. Recall, further, that the decision to test the time point 48 was made after observing the data and noting the apparent change after that time. Thus the statistical near significance of the change at point 49 should be viewed with some skepticism. With this in mind, one should be very reluctant to attribute the "significant" change at point 61 to any unforeseen effect of the changed survey forms. Rather, the change at point 61 is better interpreted as a continuation of a change in the series that had begun earlier and coincidently includes the point 61.

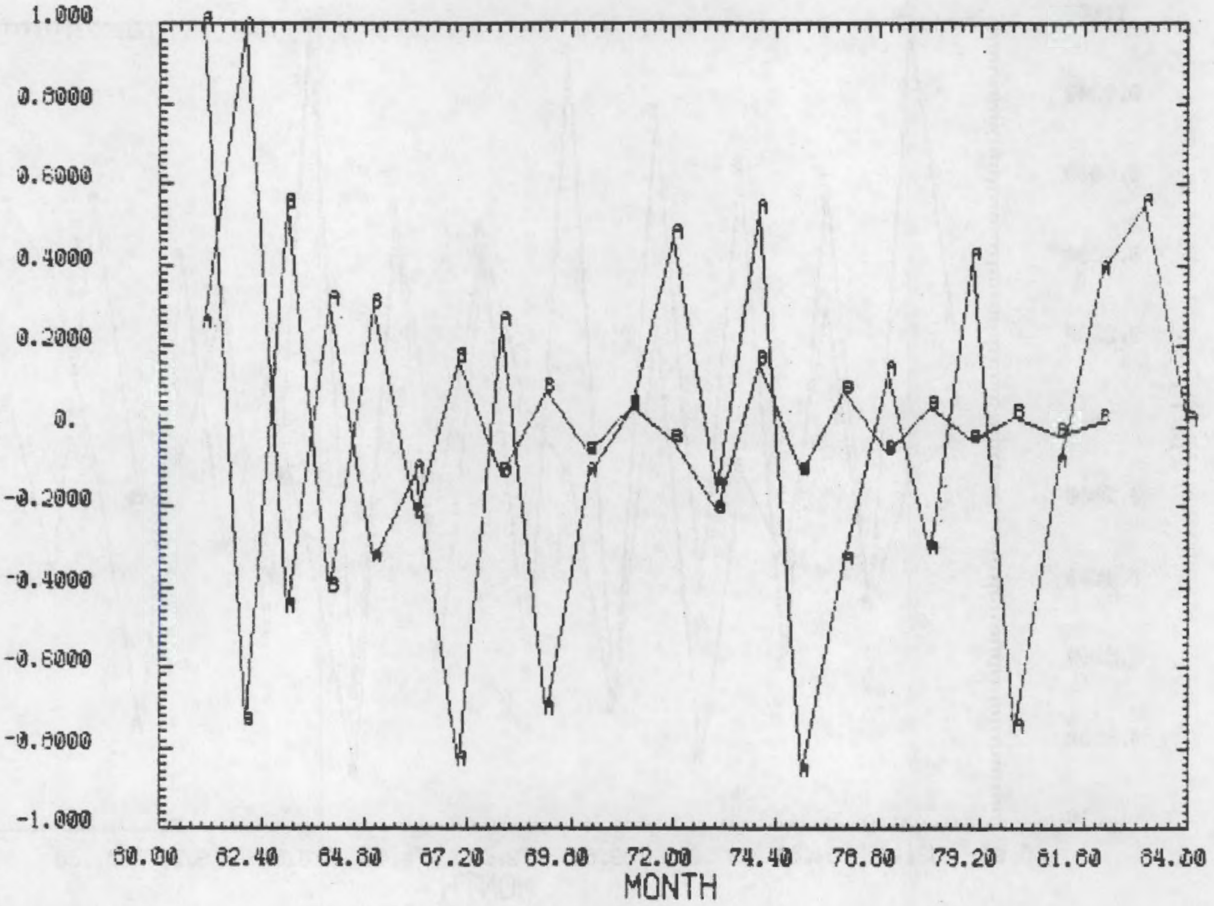

FIGURE 11. Plot of One-Step Forecast Errors (A) and Predicted Errors (B) Used to Test for Level Change at Time Point 61 for Distillate

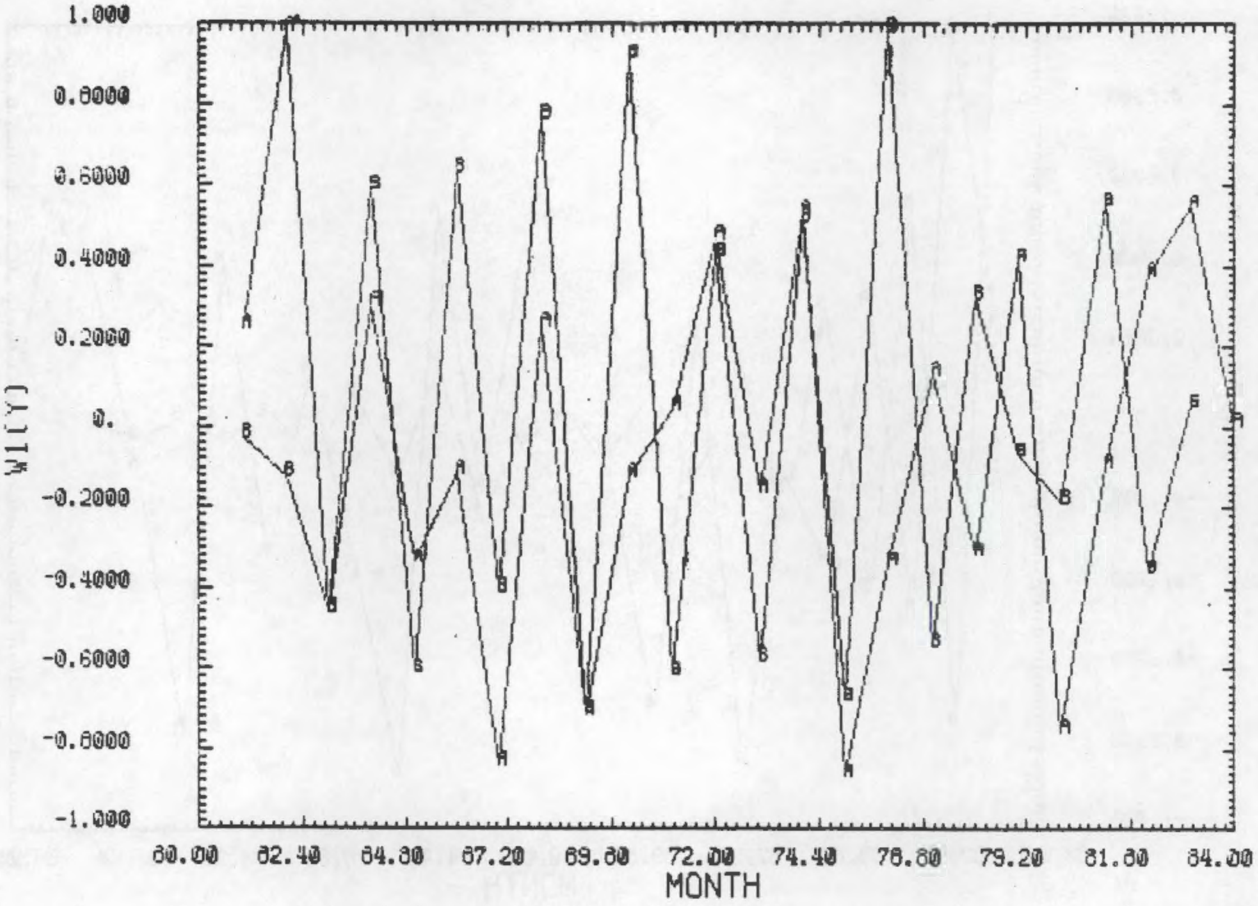

FIGURE 12. Plot of One-Step Forecast Errors (A) and Predicted Errors (B) Used to Test for MA1 Parameter Change at Time Point 61 for Distillate

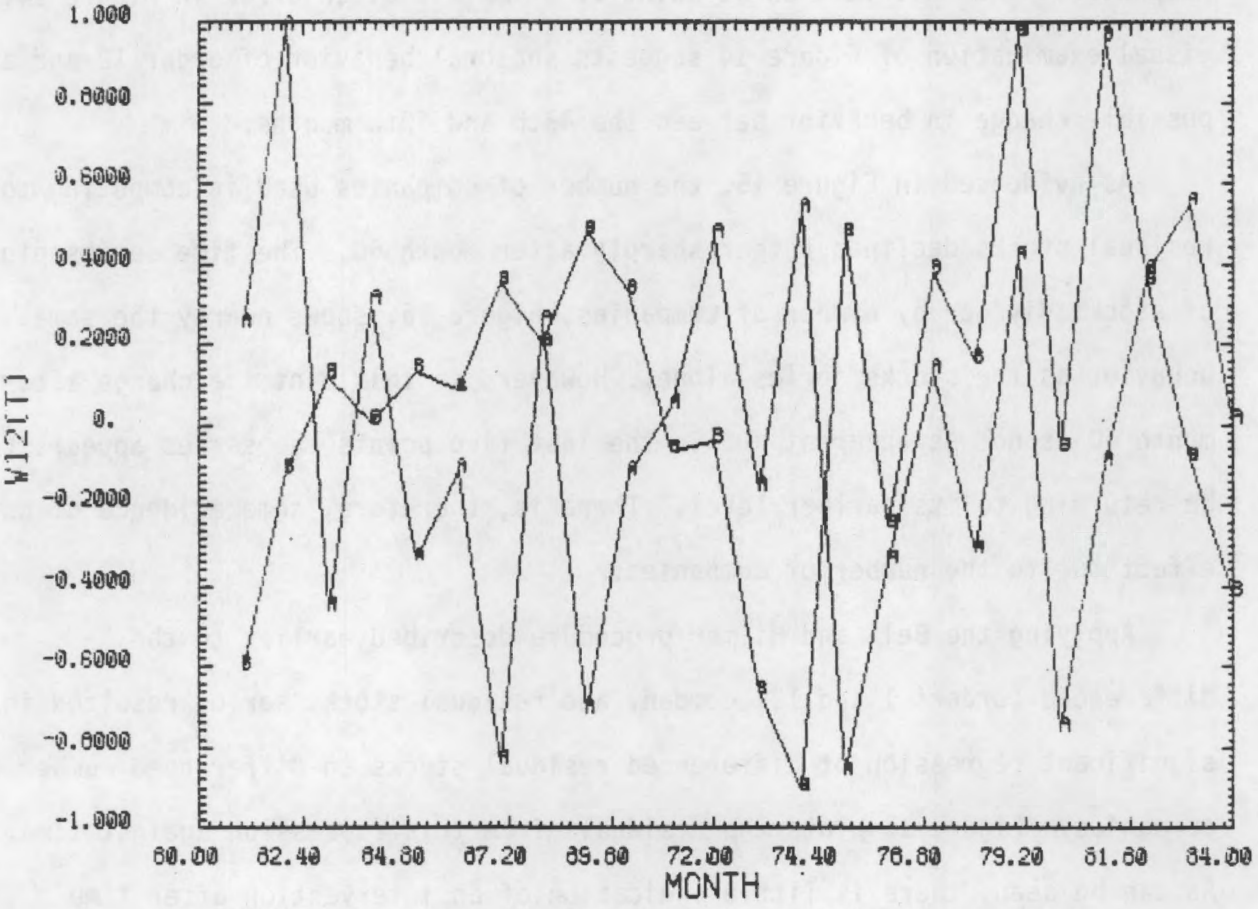

FIGURE 13. Plot of One-Step Forecast Errors (A) and Predicted Errors (B) Used to Test for MA12 Parameter Change at Time Point 61 for Distillate

## 5.2 RESIDUAL FUEL OIL

Figure 14 is a plot of the total monthly stocks of residual fuel oil for the years 1976 to 1982, Figure 15 shows the number of companies used in computing these stocks, and Figure 16 plots the stocks divided by the number of companies. The intervention at point 60 is marked by an arrow in Figure 14. A visual examination of Figure 14 suggests seasonal behavior of order 12 and a possible change in behavior between the 48th and 60th months.

As evidenced in Figure 15, the number of companies used in computing total residual stocks declines rather sharply after month 60. The time series plot of stocks divided by number of companies, Figure 16, shows nearly the same behavior as the stocks series alone. However, in this plot the change after month 60 is not as apparent and in the last five points the series appears to be returning to its earlier level. There is, therefore, some evidence of an effect due to the number of companies.

Applying the Bell and Hilmer procedure described earlier to the differenced (orders 1 and 12) company and residual stocks series resulted in a significant regression of differenced residual stocks on differenced number of companies. Figure 17 plots the residuals from this regression against time. As can be seen, there is little indication of an intervention after time point 60 in this series.

Figures 18 and 19 show the ACF and PACF respectively, for the time points 14 to 60 of the time series of the residuals from the regression. A seasonal MA model is suggested by the plots. The best fitting model was a $(0,0,0) \times (0,0,1)^{12}$ and the estimation results are given in Table 4. This model exhibited no strong indications of lack of fit and was judged acceptable.

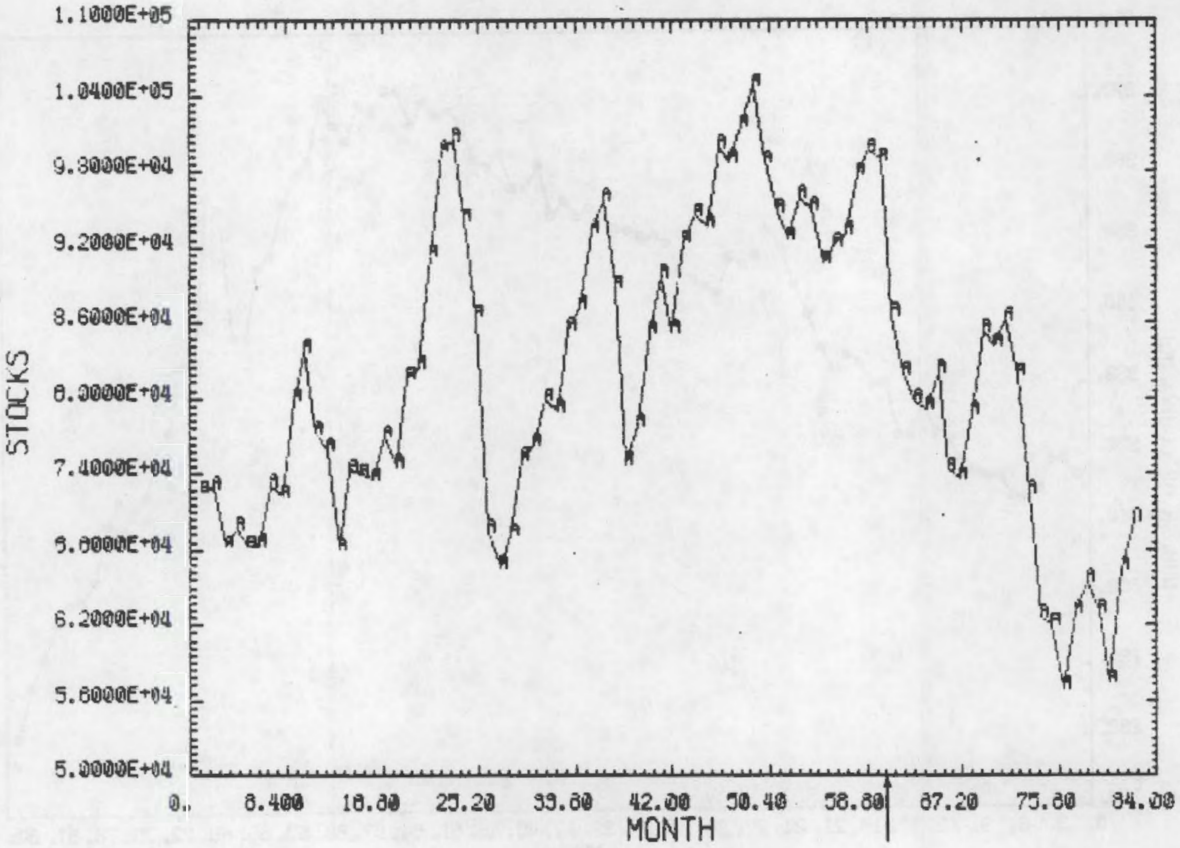

FIGURE 14. Total Residual Stocks by Month (arrow denotes intervention)

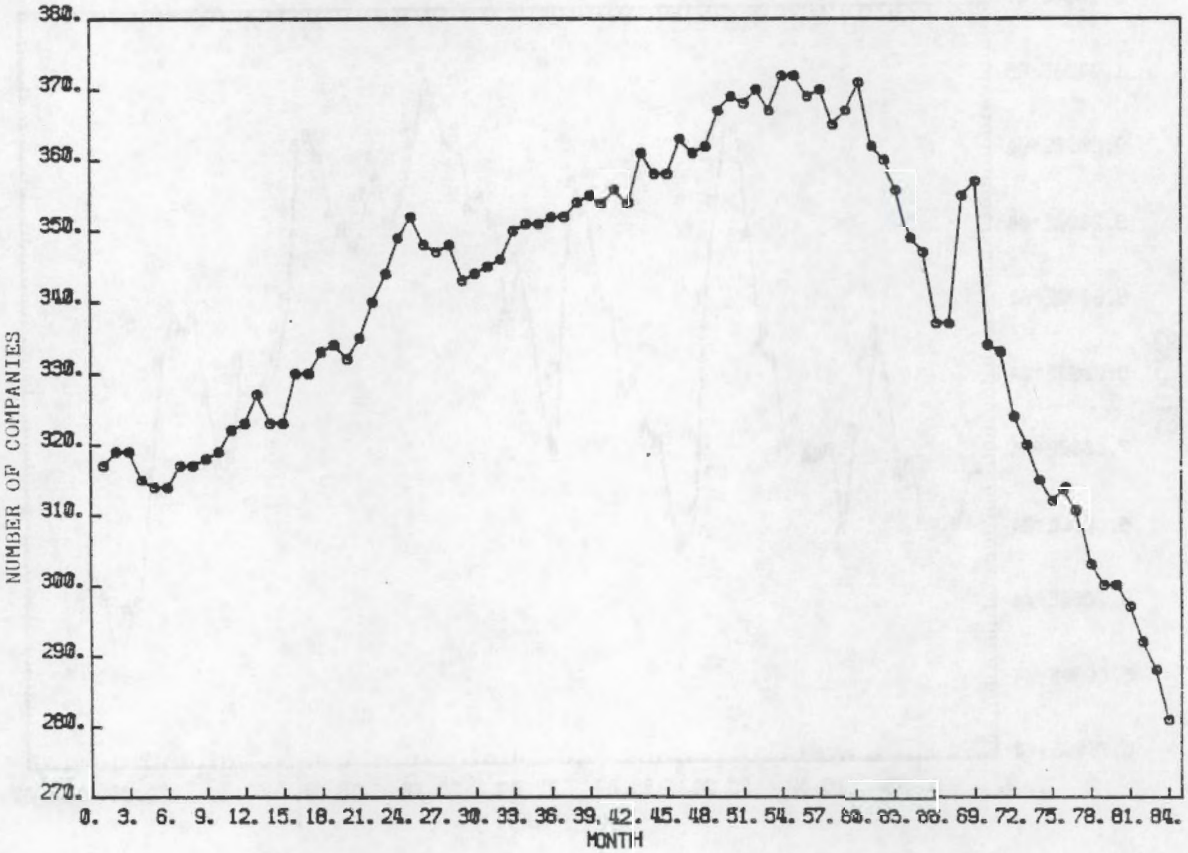

TABLE 15. Number of Companies Used in Computing Total Residual Stocks

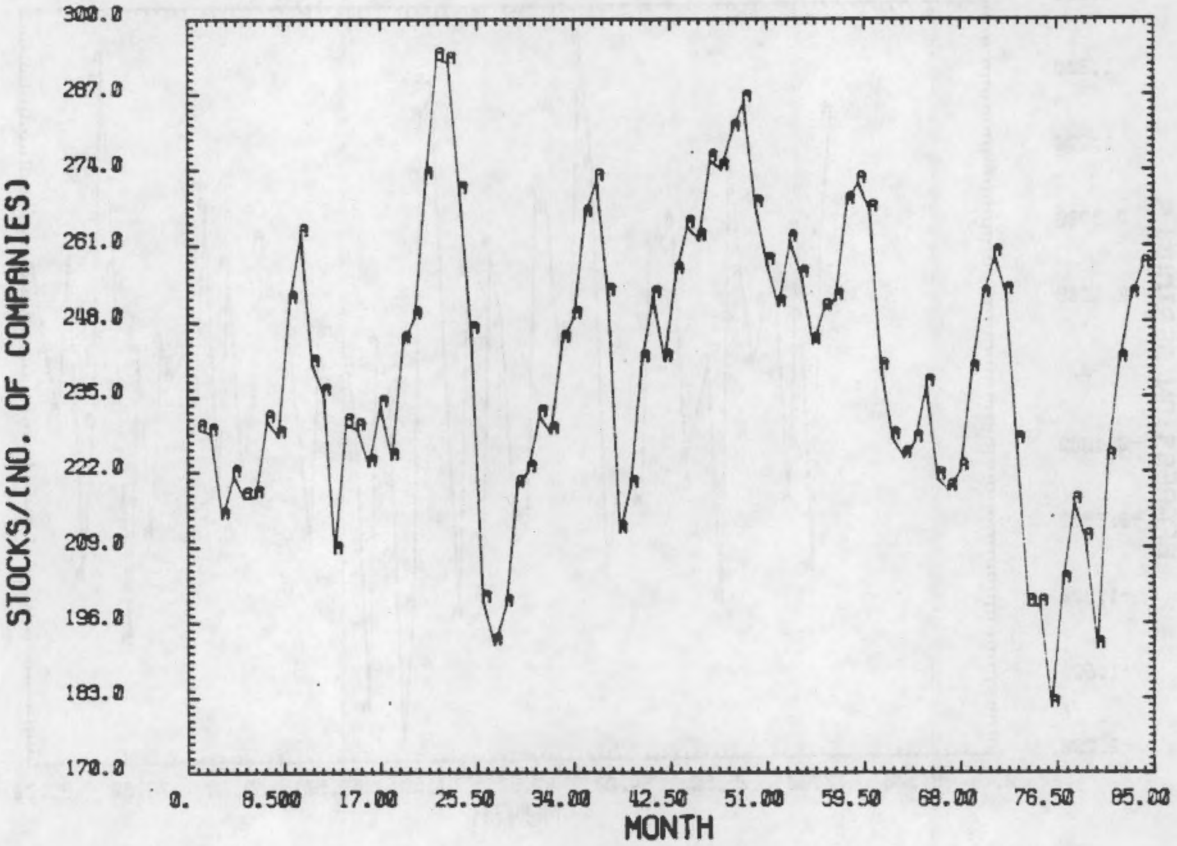

FIGURE 16. Total Residual Stocks Divided by Number of Companies Reporting

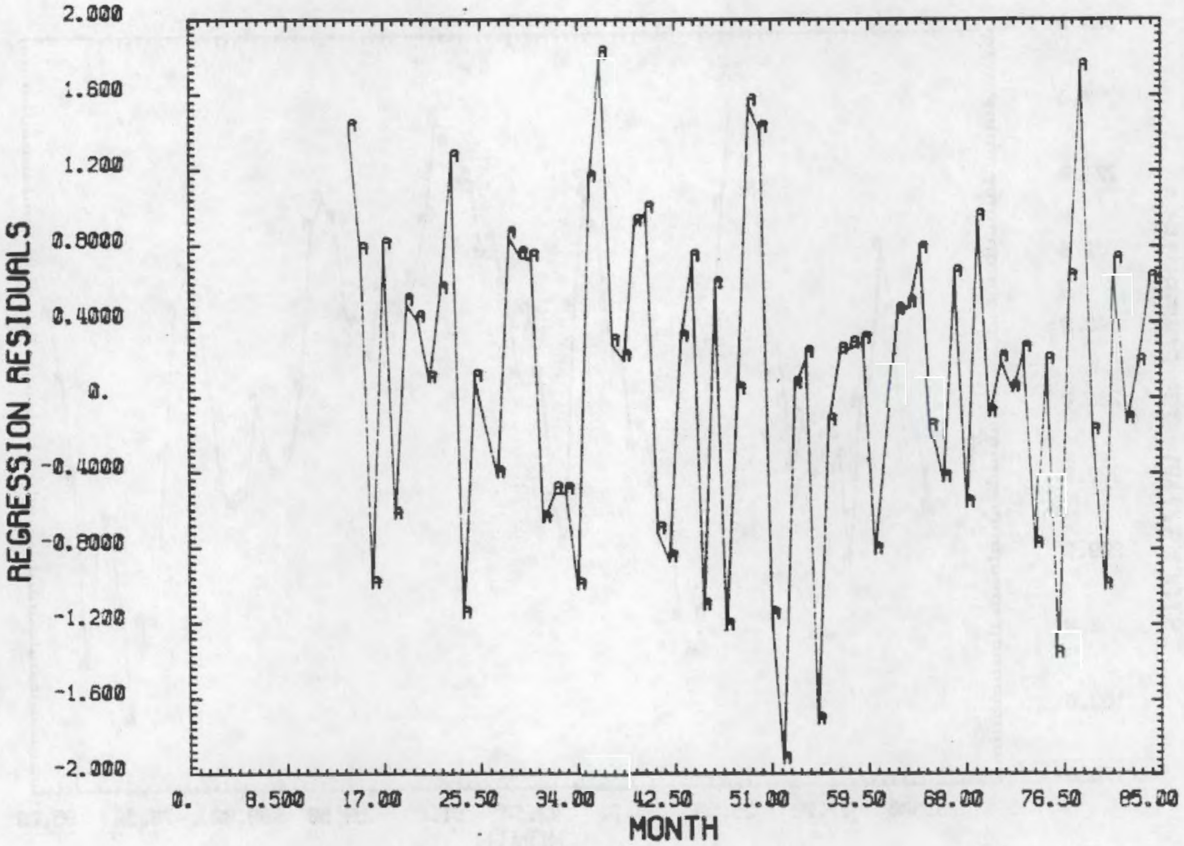

FIGURE 17. Time Series of Standardized Residuals from the Regression of Differenced Residual Fuel Oil Stocks on the Differenced Number of Reporting Companies

```
         -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
           +----+----+----+----+----+----+----+----+----+----+
 1   0.065                          XXX
 2  -0.152                      XXXXX
 3  -0.175                      XXXXX
 4  -0.092                        XXX
 5   0.004                          X
 6  -0.026                         XX
 7   0.018                          X
 8   0.222                          XXXXXXX
 9  -0.115                       XXXX
10   0.042                          XX
11  -0.118                       XXXX
12  -0.289                   XXXXXXXX
13   0.031                          XX
14   0.061                          XXX
15   0.044                          XX
16  -0.043                         XX
```

FIGURE 18. Plot of Autocorrelation Function of Regression Residuals Series for Months 14 to 60

```
         -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
           +----+----+----+----+----+----+----+----+----+----+
 1   0.065                          XXX
 2  -0.157                      XXXXX
 3  -0.158                      XXXXX
 4  -0.100                        XXX
 5  -0.040                         XX
 6  -0.087                        XXX
 7  -0.017                          X
 8   0.203                          XXXXXX
 9  -0.167                      XXXXX
10   0.130                          XXXX
11  -0.120                       XXXX
12  -0.301                  XXXXXXXXX
13   0.054                          XX
14  -0.062                        XXX
15  -0.099                        XXX
16  -0.125                       XXXX
```

<u>FIGURE 19</u>. Plot of Partial Autocorrelation Function of Regression Residuals Series for Months 14 to 60

Proceeding as before, the Q statistic was calculated for the period 61 to 84 and had value 14.00. This is not even close to significant at the 0.10 level. Thus the 1981 change in the survey form cannot be said to have had a significant effect on the residual fuel oil stock series. This was as expected since these changes did not address residual fuel oil.

It is instructive to compare these results to the results of modeling the raw (i.e., not adjusted for number of companies) residual fuel oil stocks series. The best fitting model for this data was a $(0,1,0) \times (0,1,1)^{12}$ model with the MA12 parameter equal to 0.763. This is virtually the same model as before since the regression residual series was already differenced by orders 1 and 12. However, the residual mean square for this model was 22961888, an increase of over 30%. The Q statistic for time periods 61 to 84 for this model was 18.49 also not significant at the 0.10 level but the series exhibits an apparent dramatic change after time point 60. Figure 20 plots the data and the forecasted series. These appear quite different since the forecasts continue upwards while the real data decline after month 60. This difference, though, is not judged to be significant due to the large variability in the data.

TABLE 4. Summary from Modeling Residual Fuel Oil

Series 1 to 60

| Type | Value | S.D. | t |
|---|---|---|---|
| MA12 | 0.920 | 0.097 | 9.50 |

Residual Mean Square = 17387688

Accounting for the number of companies thus has two effects. First, it removes the apparent change in the series after month 60 and, second, reduces the overall variability in the series. Discussions with EIA personnel related

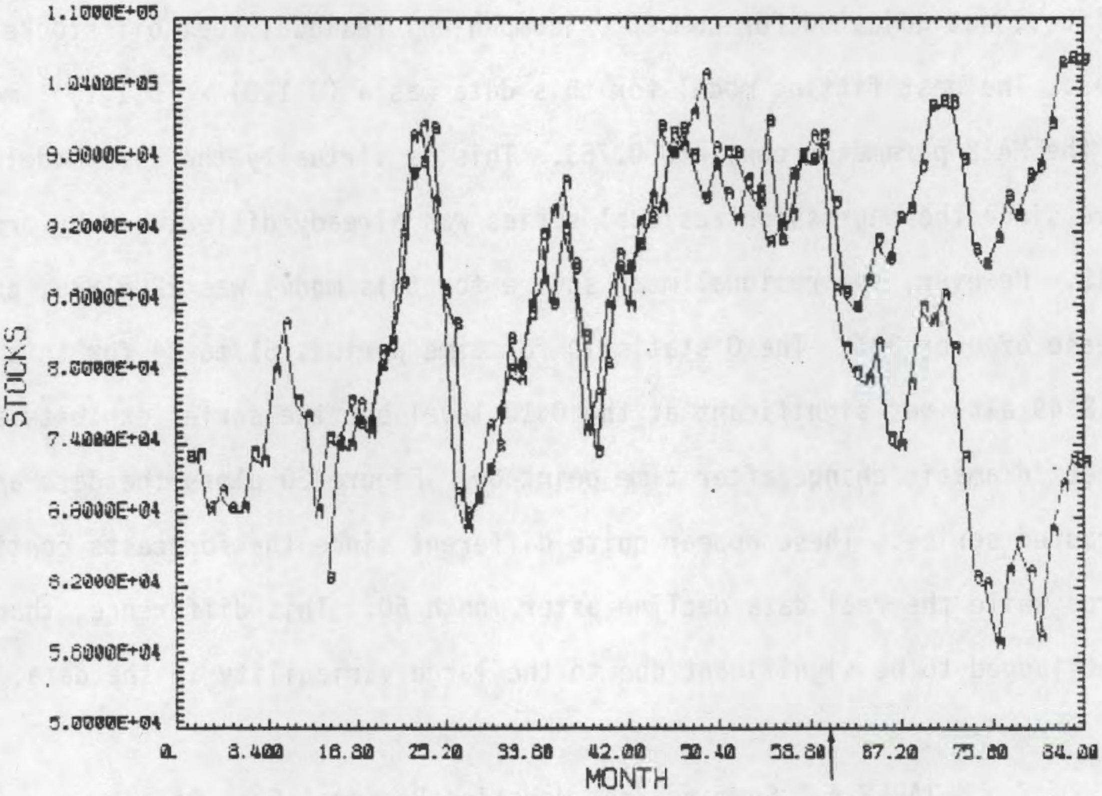

FIGURE 20. Plot of Observed (A) and Forecast Data (B) for Residual Fuel Oil (arrow denotes intervention)

the decrease in the number of companies to the general economic conditions at the time. Many industries that used residual fuel oil were closing or switching to other forms of energy.

## 5.3 MOTOR GASOLINE

Total monthly motor gasoline stocks versus month is given in Figure 21, with month 61 indicated by an arrow. Note that there tended to be twelve months between the major peaks and that there was little evidence to suggest a change at month 49. However, the early part of the series, months 1 to 25, appears different from the rest in that there was virtually no seasonal pattern. This aspect was not examined in detail at this time, but may deserve some consideration. The series analyzed was the pre-intervention data, times 1 to 60. Figure 22 shows the number of companies used to compute total gasoline stocks. Note that the intervention at month 61 was very obvious in this plot.

Figures 23 to 25 are plots of the ACF, PACF and IACF for the differenced series. As before, the ACF, PACF and IACF suggested the need for differences of order 1 and 12. Since the Bell and Hilmer procedure resulted in no significant regression of the differenced gasoline stocks on the differenced number of companies, several other methods were attempted to uncover some relationship between these two series. First, time series models (to be discussed later) were fit to the gasoline stocks data for months 1 to 60 and months 1 to 84. The residuals from these models were then compared to the number of companies reporting. In neither case did the plots show any structure nor was the regression of the residuals on the number of companies significant. Finally, the time series of the gasoline stocks divided by the number of companies reporting exhibits the same structure as the gasoline stocks series. Thus, any "company" effect on the gasoline stocks is more complex than just the number of companies reporting.

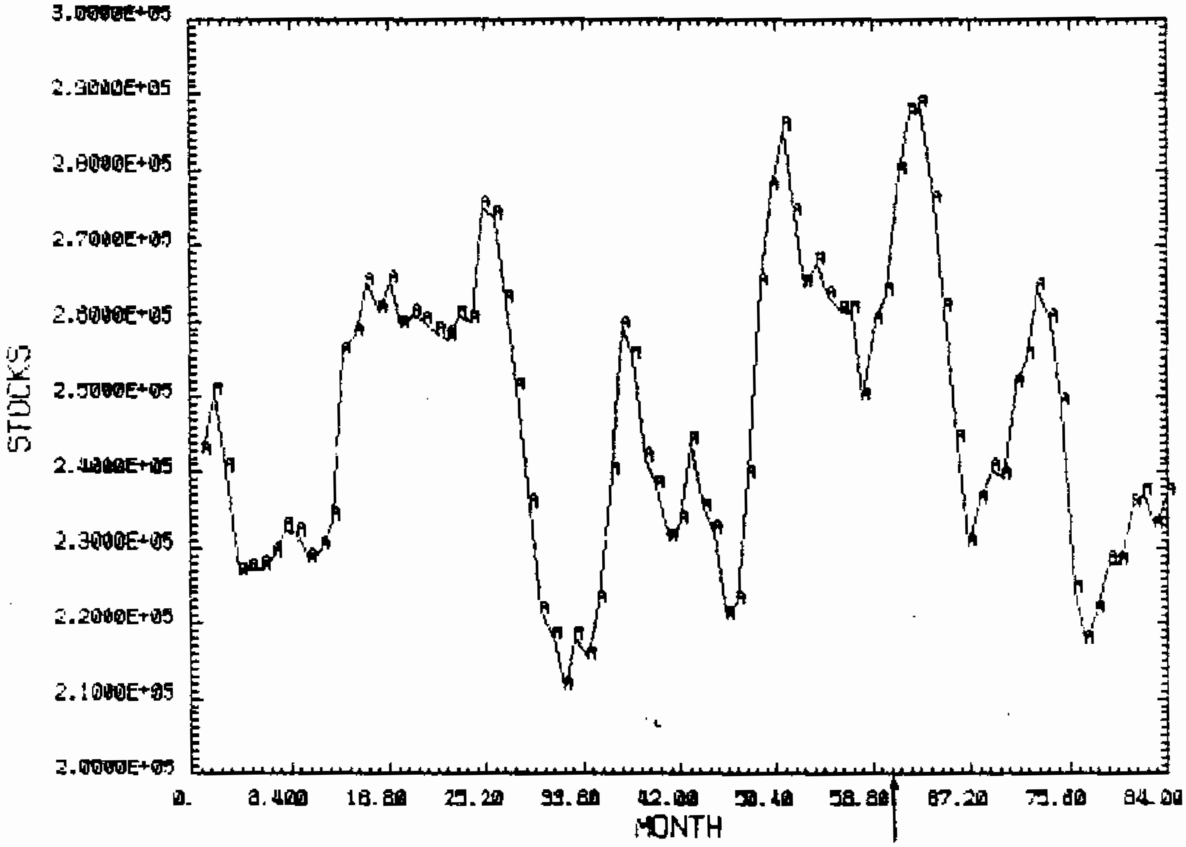

FIGURE 21. Total Gasoline Stocks by Month (arrow denotes intervention)

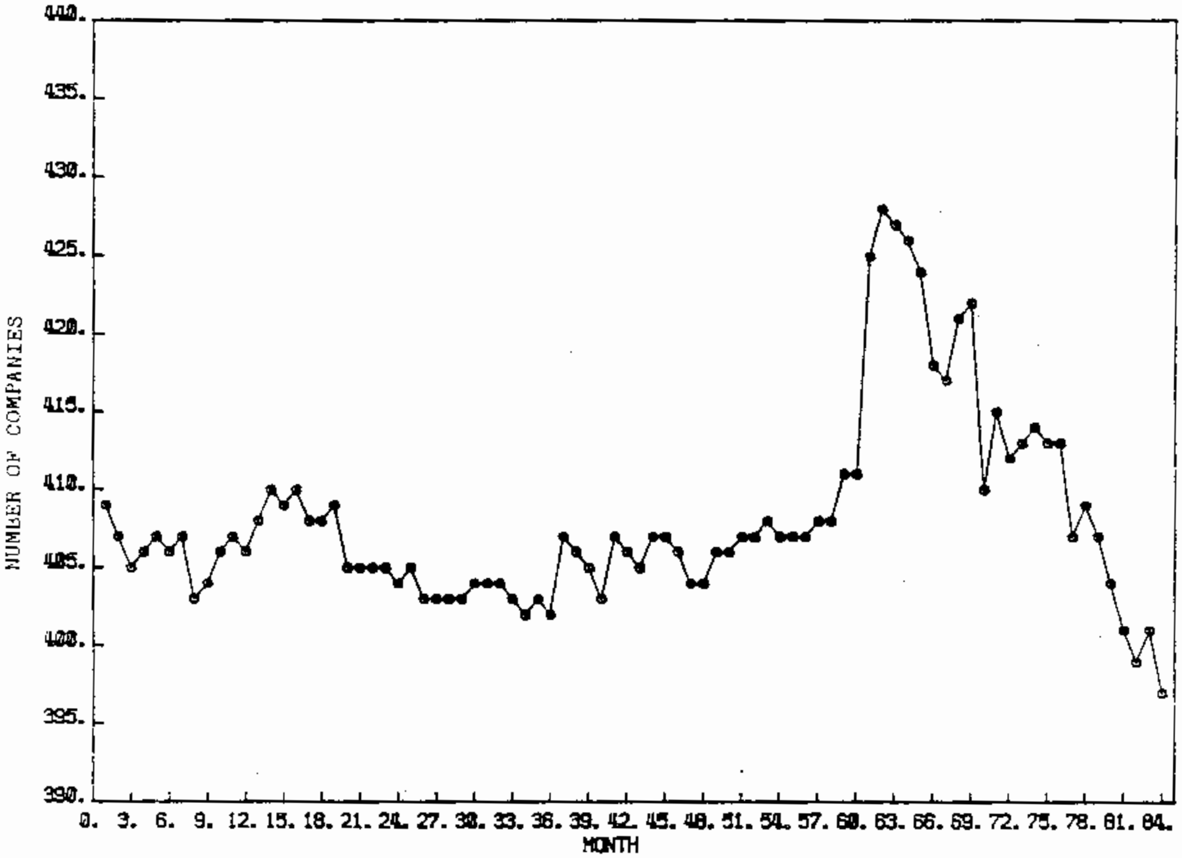

FIGURE 22. Number of Companies Used to Compute Total Gasoline Stocks

```
AUTOCORRELATIONS
LAG COVARIANCE CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1         STD
  0   86523853     1.00000 |                    |********************|         0
  1   30364498     0.35094 |              .     |*******             |  0.145865
  2   13557919     0.15670 |             .      |***   .             |  0.162841
  3    1206954     0.01395 |             .      |      .             |  0.166019
  4  -23009249    -0.26593 |             . *****|      .             |  0.166043
  5   -6024656    -0.06963 |             .     *|      .             |  0.174971
  6     384354     0.00444 |             .      |      .             |   0.17546
  7   12134806     0.14025 |             .      |***   .             |  0.175462
  8   15516037     0.17933 |             .      |****  .             |  0.177831
  9   -2299777    -0.02658 |             .     *|      .             |  0.181638
 10   -9319899    -0.10771 |             .    **|      .             |  0.181721
 11  -16869403    -0.19497 |             .  ****|      .             |  0.183074
 12  -28653941    -0.33117 |             *******|      .             |   0.18744
 13  -18765374    -0.21688 |            .   ****|       .            |  0.199501
 14   -5961775    -0.06890 |            .      *|       .            |  0.204456
 15     639414     0.00739 |            .       |       .            |  0.204949
 16    -928416    -0.01073 |            .       |       .            |  0.204955
 17    1288515     0.01489 |            .       |       .            |  0.204967
 18    2606265     0.03012 |            .       |*      .            |   0.20499
 19   -7822451    -0.09041 |            .     **|       .            |  0.205084
 20    2601634     0.03007 |            .       |*      .            |   0.20593
 21    1535481     0.01775 |            .       |       .            |  0.206024
 22  -17406051    -0.20117 |            .   ****|       .            |  0.206056
 23   -5555155    -0.06420 |            .      *|       .            |  0.210193
 24  -11947943    -0.13809 |            .    ***|       .            |   0.21061
                           '.' MARKS TWO STANDARD ERRORS
```

FIGURE 23. Plot of Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Motor Gasoline Series

```
PARTIAL AUTOCORRELATIONS
           LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
             1    0.35094  |              .     |*******             |
             2    0.03825  |              .     |*    .              |
             3   -0.05980  |              .    *|     .              |
             4   -0.29669  |              ******|     .              |
             5    0.14159  |              .     |***  .              |
             6    0.06659  |              .     |*    .              |
             7    0.14909  |              .     |***  .              |
             8   -0.31724  |              .     |     .              |
             9   -0.15188  |              .  ***|     .              |
            10   -0.09491  |              .   **|     .              |
            11   -0.04073  |              .    *|     .              |
            12   -0.20416  |              . ****|     .              |
            13   -0.09947  |              .   **|     .              |
            14    0.02285  |              .     |     .              |
            15    0.00677  |              .     |     .              |
            16   -0.15088  |              .  ***|     .              |
            17    0.02975  |              .     |*    .              |
            18    0.10450  |              .     |**   .              |
            19   -0.05788  |              .    *|     .              |
            20    0.09884  |              .     |**   .              |
            21   -0.04573  |              .    *|     .              |
            22   -0.36545  |             *******|     .              |
            23   -0.02356  |              .     |     .              |
            24   -0.09817  |              .   **|     .              |
```

FIGURE 24. Plot of Partial Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Motor Gasoline Series

```
INVERSE AUTOCORRELATIONS
          LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
            1    -0.19008  |             . ****|     .             |
            2    -0.16479  |             .  ***|     .             |
            3    -0.03196  |             .    *|     .             |
            4     0.20912  |             .     |**** .             |
            5    -0.09390  |             .   **|     .             |
            6     0.12004  |             .     |**   .             |
            7    -0.11682  |             .   **|     .             |
            8    -0.08303  |             .   **|     .             |
            9     0.12684  |             .     |***  .             |
           10     0.11979  |             .     |**   .             |
           11    -0.14016  |             .  ***|     .             |
           12     0.17998  |             .     |**** .             |
           13     0.16528  |             .     |***  .             |
           14    -0.05737  |             .    *|     .             |
           15    -0.07550  |             .   **|     .             |
           16     0.05914  |             .     |*    .             |
           17    -0.05831  |             .    *|     .             |
           18     0.07255  |             .     |*    .             |
           19     0.08345  |             .     |**   .             |
           20    -0.11286  |             .   **|     .             |
           21    -0.11870  |             .   **|     .             |
           22     0.22918  |             .     |*****.             |
           23     0.01536  |             .     |     .             |
```

FIGURE 25. Plot of Inverse Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Motor Gasoline Series

Several multiplicative models were suggested by Figures 23 through 25. Many different models were fit in the identification stage of the analysis including some non-seasonal models. However, the residuals from all non-multiplicative models exhibited seasonality. Two models appeared promising: a $(0,1,1) \times (0,1,1)^{12}$ model and a $(0,1,0) \times (0,1,2)^{12}$ model. Both contained a nonsignificant parameter (the estimation routine would not permit estimation of either model with the nonsignificant parameter removed) and the residual mean square of the second model was slightly less than that of the first. The residual ACF's from the two models were about the same, and neither showed a substantial lack-of-fit. The first model was chosen for analysis due to its comparative simplicity since it lacks the term of order 24 included in the second model.

The results of the Akaike procedure were ambiguous. The order chosen was 4 and the plot showed a slight dip at order 19. While this did not really support the model, it was not strong enough evidence to reject the model. More work on modeling this series might produce a better understanding of its behavior.

The results of the estimation for the model are given in Table 5.

TABLE 5. Summary from Modeling Motor Gasoline Series 1 to 60

Differences 1,12

| Type | Value | S.D. | t |
|---|---|---|---|
| MA1 | -0.237 | 0.149 | -1.58 |
| SMA(12) | 0.755 | 0.174 | 4.34 |

Residual Mean Square = 47901116

When calculated for the months 61 to 84, the Q statistic discussed previously had the value 30.59 which was almost significant at the 0.15 level. The residual plots, Figures 26 to 28, suggested that the intervention may be reflected in a change in the MA1 parameter. The post/pre change in this parameter was estimated to be -0.318 and was significant at about the 0.15 level. Thus, an estimate of the post-intervention MA1 parameter was -0.555. Neither the level change nor a change in the SMA12 parameter were estimated to be significant.

## 5.4 CRUDE OIL

The series chosen for analysis was total monthly stocks at refineries. A plot of monthly stocks versus time is given in Figure 29 where January 1981 is marked by an arrow. As can be seen, the data are quite variable with a large increase between the months 48 to 66, after which time the series seems to have returned to its behavior before time point 48. This aspect made modeling this series more difficult. Figure 30 plots the number of companies used in computing the total crude stocks. Figure 29 showed little evidence of seasonality and no strong indication of the 1981 intervention.

The ACF, PACF and IACF of the differenced series from months 1 to 60 are given in Figures 31 to 33. The ACF of the original series clearly showed that a difference of order 1 was required, but that no difference of order 12 was needed. The same methods as discussed for the gasoline stocks series were also used to examine any possible effect of the number of companies reporting on the refinery crude oil inventory series. The results of these analyses were similar to those for gasoline. That is, no significant effect was found. Thus modeling proceeded on the raw data series.

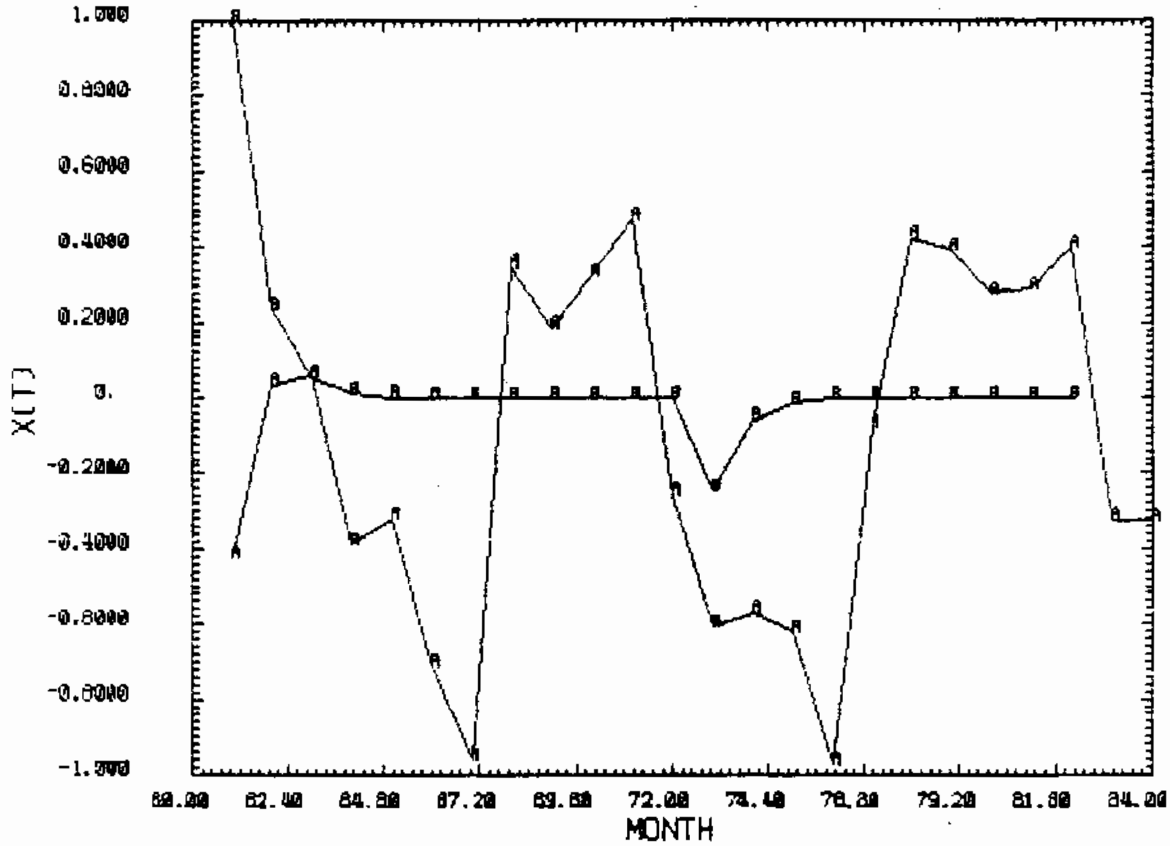

FIGURE 26. Plot of One-Step Forecast Errors (A) and Predicted Errors (B) Used to Test for Level Change at Time Point 61 for Gasoline

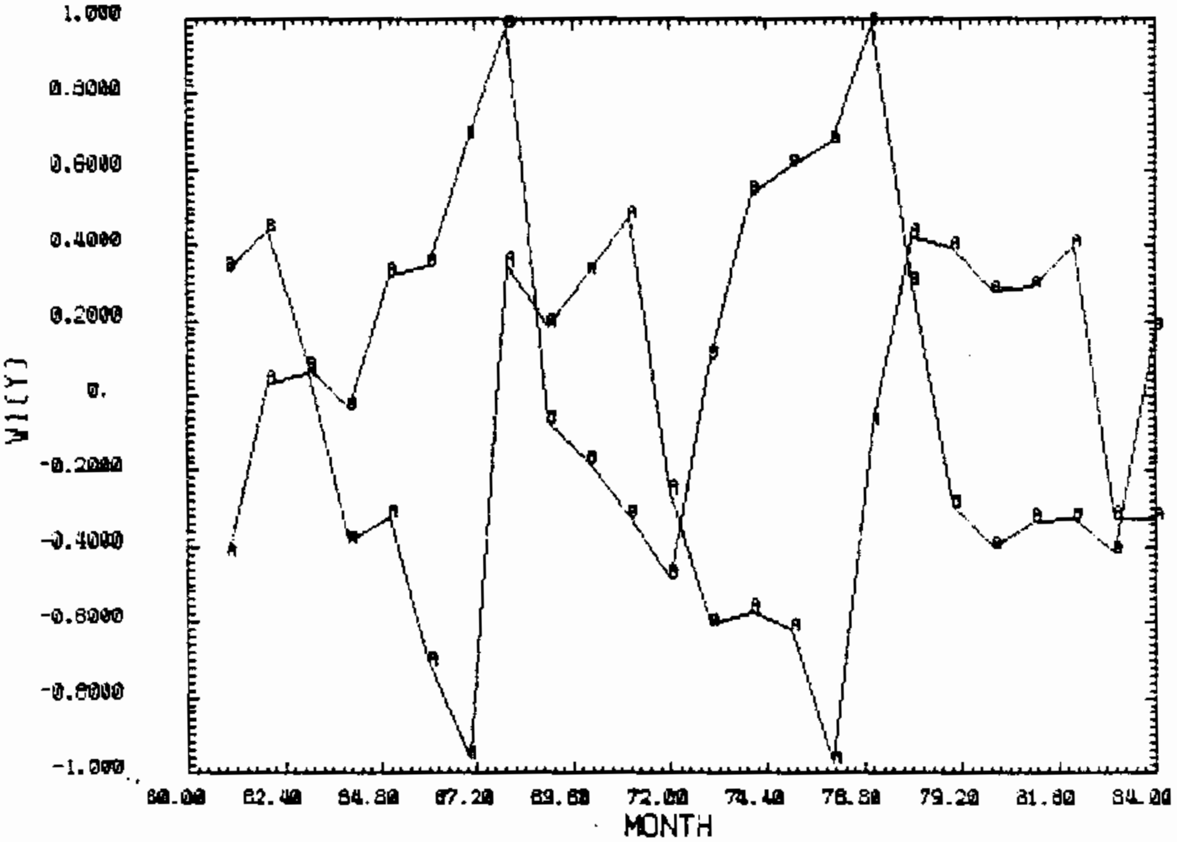

FIGURE 27. Plot of One-Step Forecast Errors (A) and Predicted Errors (B) Used to Test for MA1 Parameter Change at Time Point 61 for Gasoline

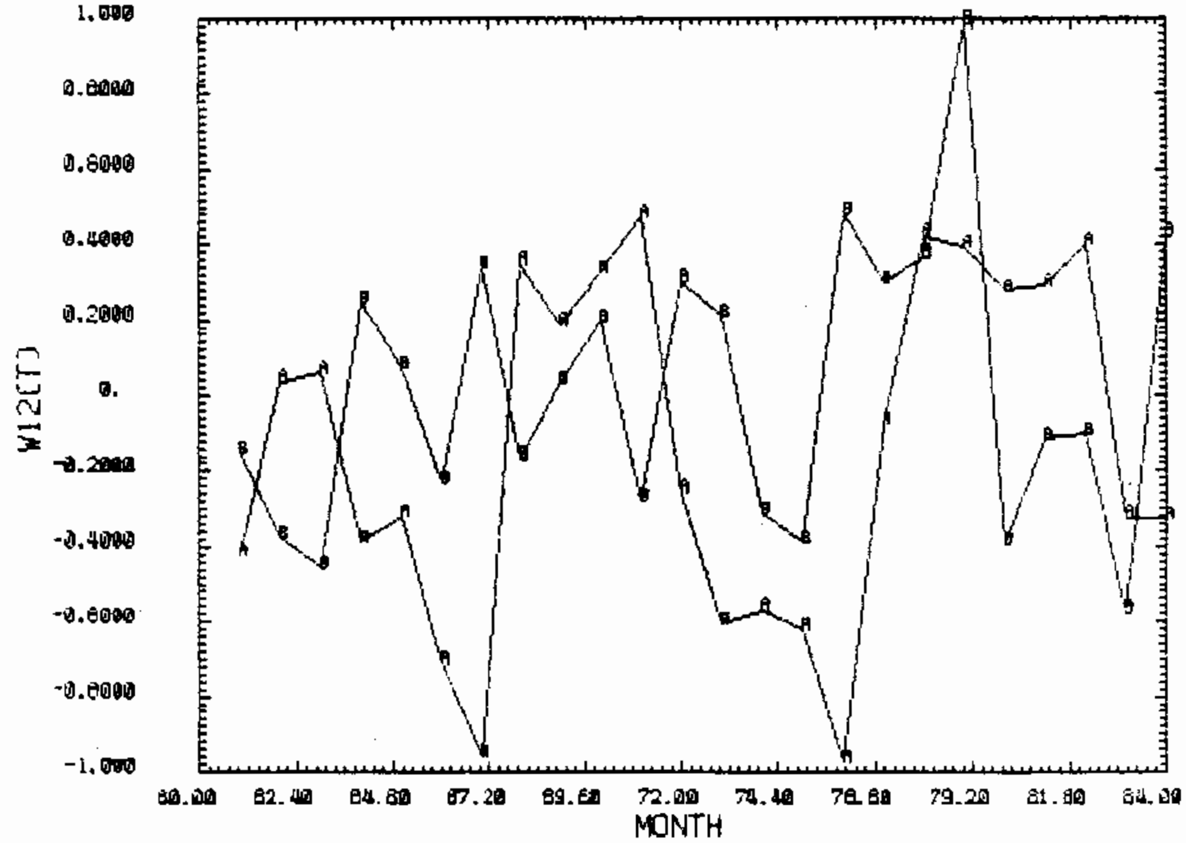

FIGURE 28. Plot of One-Step Forecast Errors (A) and Predicted Errors (B) Used to Test for MA12 Parameter Change at Time Point 61 for Gasoline

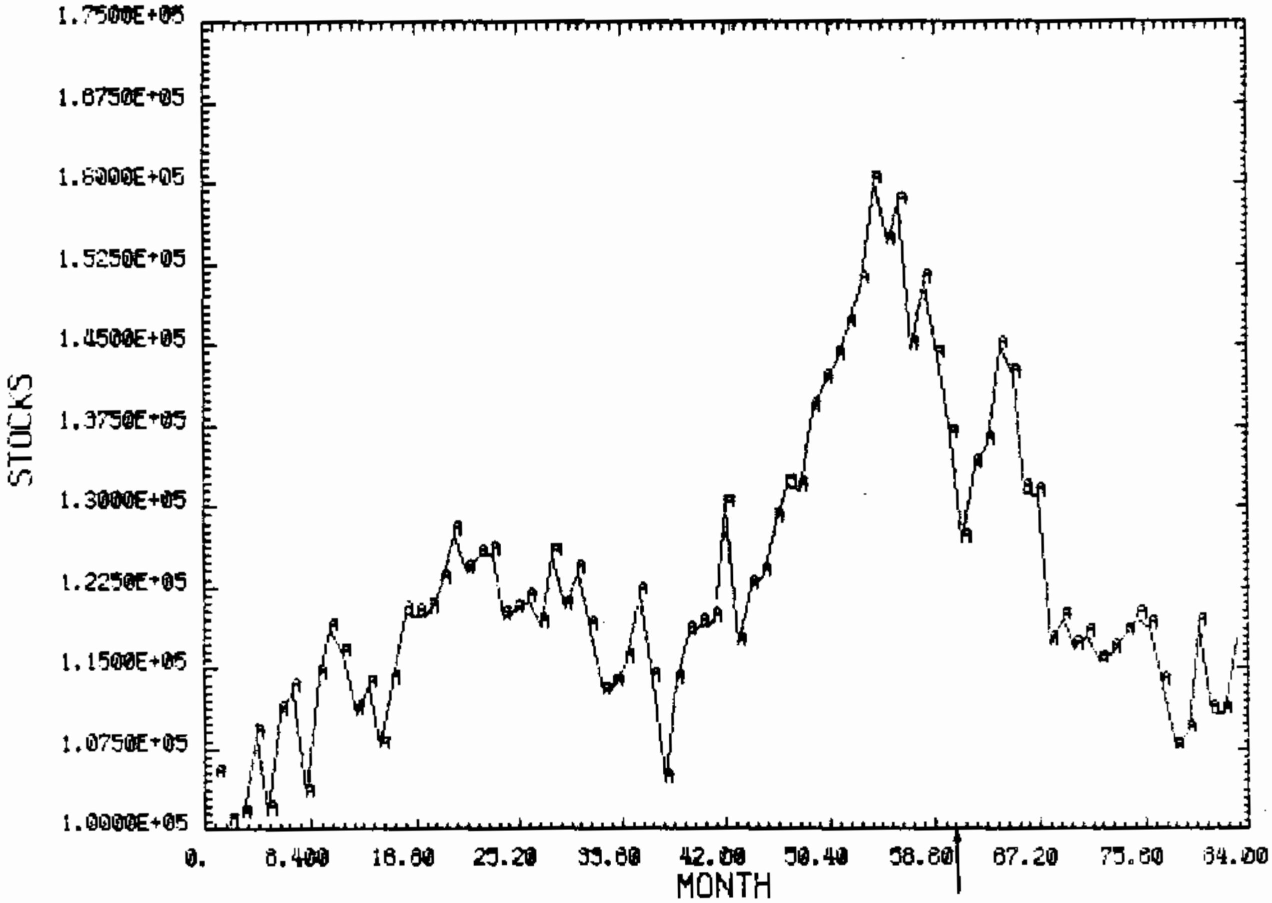

FIGURE 29. Total Crude Oil Stocks by Month (arrow denotes intervention)

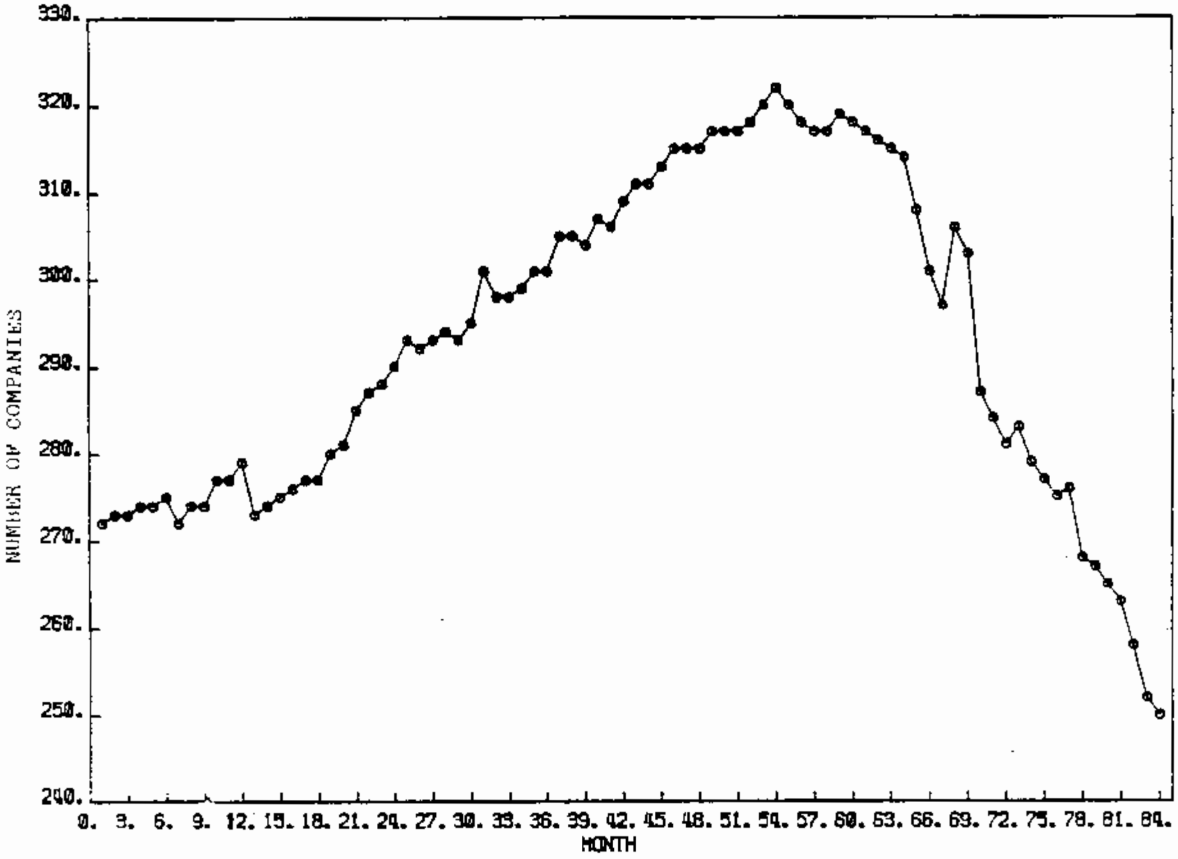

FIGURE 30. Number of Companies Used to Compute Total Crude Oil Stocks

```
AUTOCORRELATIONS
LAG COVARIANCE CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1         STD
  0   34148493     1.00000 |                    |********************|         0
  1  -10059760    -0.29459 |              ******|    .                |  0.130189
  2    -288032    -0.00843 |              .     |     .              |  0.141035
  3    3201879     0.09376 |              .     |**   .              |  0.141044
  4    -521310    -0.01527 |              .     |     .              |  0.142096
  5    -293896    -0.00861 |              .     |     .              |  0.142124
  6    -110626    -0.00324 |              .     |     .              |  0.142133
  7    3010951     0.08817 |              .     |**   .              |  0.142134
  8   -5756092    -0.16856 |              .  ***|     .              |  0.143058
  9     415240     0.01216 |              .     |     .              |  0.146386
 10    -652382    -0.01910 |              .     |     .              |  0.146403
 11   -3046103    -0.08920 |              .   **|     .              |  0.146445
 12    5453712     0.15971 |              .     |***  .              |  0.147363
 13   -4611451    -0.13504 |              .  ***|     .              |  0.150268
 14    2985978     0.08744 |              .     |**   .              |  0.152311
 15   -7191024    -0.21058 |              . ****|     .              |   0.15316
 16    4047530     0.11653 |              .     |**   .              |  0.157991
 17   -5514342    -0.16148 |              .  ***|     .              |  0.159491
 18   -4184451    -0.12254 |              .   **|     .              |  0.162238
 19    5123356     0.15003 |             .      |***   .             |  0.163799
 20   -1498924    -0.04389 |             .     *|      .             |  0.166112
 21   -2832903    -0.08296 |             .    **|      .             |  0.166309
 22   -1740572    -0.05097 |             .     *|      .             |  0.167009
 23     623311     0.01825 |             .      |      .             |  0.167272
 24    3689420     0.10804 |             .      |**    .             |  0.167306
                           '.' MARKS TWO STANDARD ERRORS
```

FIGURE 31. Plot of Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Crude Oil Series

```
PARTIAL AUTOCORRELATIONS
           LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
             1   -0.29459  |             ******|    .              |
             2   -0.10427  |              .  **|    .              |
             3    0.06676  |              .    |*   .              |
             4    0.03731  |              .    |*   .              |
             5    0.00624  |              .    |    .              |
             6   -0.01389  |              .    |    .              |
             7    0.08912  |              .    |**  .              |
             8   -0.12793  |              . ***|    .              |
             9   -0.07971  |              .  **|    .              |
            10   -0.07217  |              .   *|    .              |
            11   -0.10433  |              .  **|    .              |
            12    0.12785  |              .    |*** .              |
            13   -0.05460  |              .   *|    .              |
            14    0.06331  |              .    |*   .              |
            15   -0.21104  |              .****|    .              |
            16    0.00117  |              .    |    .              |
            17   -0.20007  |              .****|    .              |
            18   -0.23285  |              *****|    .              |
            19   -0.02093  |              .    |    .              |
            20    0.05931  |              .    |*   .              |
            21   -0.06639  |              .   *|    .              |
            22   -0.10943  |              .  **|    .              |
            23   -0.13176  |              . ***|    .              |
            24    0.07972  |              .    |**  .              |
```

FIGURE 32. Plot of Partial Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Crude Oil Series

```
INVERSE AUTOCORRELATIONS
           LAG CORRELATION -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
             1     0.38110  |              .    |********           |
             2     0.16307  |              .    |*** .              |
             3     0.00782  |              .    |    .              |
             4     0.07872  |              .    |**  .              |
             5     0.09889  |              .    |**  .              |
             6     0.09347  |              .    |**  .              |
             7     0.15532  |              .    |*** .              |
             8     0.27307  |              .    |*****              |
             9     0.19163  |              .    |****.              |
            10     0.11948  |              .    |**  .              |
            11     0.02644  |              .    |*   .              |
            12    -0.01536  |              .    |    .              |
            13     0.11967  |              .    |**  .              |
            14     0.08245  |              .    |**  .              |
            15     0.19389  |              .    |****.              |
            16     0.13191  |              .    |*** .              |
            17     0.23894  |              .    |*****              |
            18     0.14487  |              .    |*** .              |
            19     0.00005  |              .    |    .              |
            20     0.02139  |              .    |    .              |
            21     0.13251  |              .    |*** .              |
            22     0.11315  |              .    |**  .              |
            23     0.05058  |              .    |*   .              |
            24    -0.05386  |              .   *|    .              |
```

FIGURE 33. Plot of Inverse Autocorrelation Function for Months 1 to 60 for Differenced (1 and 12) Crude Oil Series

Figures 31 to 33 suggested an autoregressive model for the differenced data. The best fitting model was a (1,1,0) model. The results of the fit of this model are given in Table 6.

TABLE 6. Summary from Modeling Crude Oil

Series 1 to 60

Differences 1,12

| Type | Value | S.D. | t |
|---|---|---|---|
| AR1 | -0.304 | 0.128 | -2.37 |

Residual Mean Square = 32176502

The residuals from this model gave no indication of lack of fit. The Q statistic calculated for the time periods 49 to 84 was 29.11. Thus, this furnished little evidence of an effect of the revised survey forms. However, the large variability of this series would tend to mask any but the largest interventions, and further analyses could be directed at reducing this variability. One place to start would be to try to account for the behavior between months 48 and 66 and thus to remove that source of variability.

## 5.5 INTERVENTION ADJUSTMENT

### 5.5.1 Methods of Intervention Adjustment

In intervention adjustment the question to be answered is, "If the present model (post-intervention) had always been in effect and all other conditions had been as before, what should the pre-intervention data have been?" For some types of interventions, methods of adjusting for the change are rather straightforward. For instance, if the mean level of a series is known to have changed, and an estimate of the new level is available, it is obvious how to adjust the past series to obtain data comparable to the contemporary data. Furthermore, since some measure of the precision of the estimated new level will most likely also be available, error estimates on the adjusted values also can be produced.

When an intervention affects the stochastic parameters of the model, however, the techniques for adjustment are more difficult. Two methods that are applicable can be summarized as follows:

1.) Use the estimated post-intervention model to forecast the observations in the pre-intervention time period.

2.) Use the post-intervention model and the estimated pre-intervention shocks ($a_t$'s) to estimate the pre-intervention series.

These two are similar in that both use the estimated post-intervention model. However, the second method takes advantage of the past record also to obtain an estimate of the random component. Thus, the adjustments should reflect the past random shocks and should be better adjustments. The first method assumes total ignorance of the pre-intervention series and produces deterministic adjustments based completely on the form of the model. Since this method is the standard technique used for forecasting time series, there is a well-accepted method for assessing the variability of the forecasts and several computer programs for performing the forecasting and estimating confidence bands for the forecasts are available. Further comparisons of the two techniques are contained in Table 7.

TABLE 7. Forecasting Versus Adjustment

| Forecasting (Box-Jenkins) | Adjustment |
|---|---|
| 1.) Assumes parameter estimation errors do not contribute significantly to forecasting errors. | 1.) Has estimates of the shocks ($a_t$) available for the forecast period. |
| 2.) Assumes forecasts depend appreciably only on recent values of the series. | 2.) These estimated shocks have (most likely) non-zero expected values. |
| 3.) Assumes variance of forecasts comes only from lack of knowledge of future shocks ($a_t$'s). | 3.) Errors in adjustment come only from parameter estimation errors and the dependence on distant past (unknown) values of the series. |
| 4.) The forecast is $E(z_{t+\ell} \mid z_1, z_2, \ldots, z_t)$ and $E(a_t)$ is taken to be zero for the forecast period. | |

It should be noted that although commonly ignored in practice, both estimation and "start-up" errors are present in standard Box-Jenkins type forecasts also. They are the only sources of error in the adjustment procedure described above. However, errors in the parameter estimates manifest themselves not only in the time series model but also in the estimated shocks.

These two techniques will be illustrated on the motor gasoline series. As discussed in Section 5.3 the post-intervention model for motor gasoline was

$$(1 - B)\ (1 - B^{12})\ z_t = (1 + 0.555\ B)\ (1 - 0.755\ B^{12})\ a_t\ . \qquad (7)$$

Expanding and adjusting the coefficients results in

$$z_t = z_{t+1} + z_{t+12} - z_{t+13} + a_{t+13} + 0.555\ a_{t+12} - 0.755\ a_{t+1} - 0.419\ a_t\ . \qquad (8)$$

This expression is used in both methods of intervention adjustment. They differ in the values used in the $a_t$ series. For the first method, the $a_t$ are taken to be zero for all pre-intervention periods, that is for months 1 to 60, and are estimated, by solving for $a_t$ in formula (7) above, for months 61 to 84. Substituting the respective values in the above expression produces a recursive formula for estimating the adjusted series for months 1 to 60. The original and the series adjusted by the deterministic method are plotted in Figure 34. The deterministic nature of these adjusted values was evident in the regular shape of the graph, as was the seasonal effect. It should be noted that this method results in the minimum mean-square error forecast of the pre-intervention series given only the series from 61 to 84 and expressions for the variance of any forecast can be derived.

However, in an intervention analysis more is known than just the post-intervention series. The second method of adjustment that incorporates a random component attempts to use this increased information. Specifically, the pre-intervention model was used to recursively estimate $a_t$ for t = 1 to 60. These were combined with the estimated $a_t$ for t from 61 to 84 and used in the recursion formula. The adjusted series with this method and the original data are plotted in Figure 35. As can be seen, this method tends to track the data better than the deterministic method and is, in general, expected to be a better method of adjusting for an intervention. From the graph it was evident

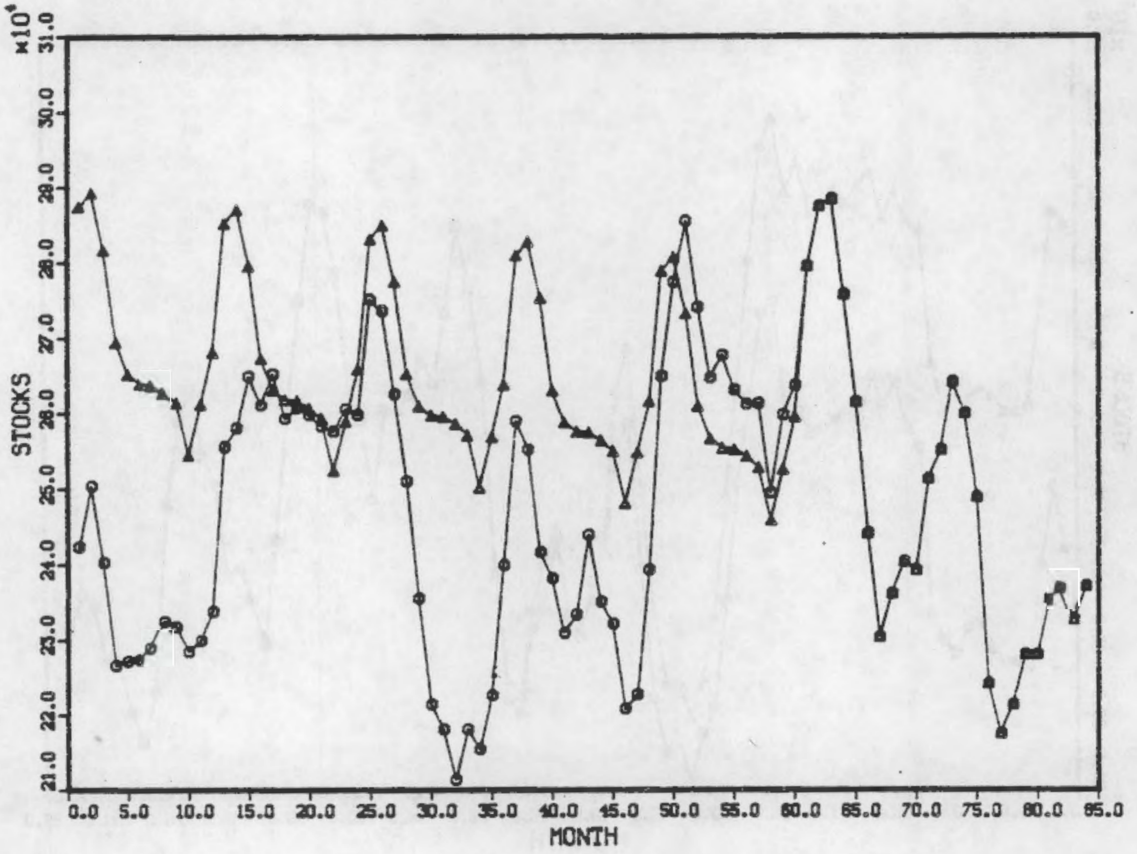

FIGURE 34. Gasoline Stocks (O) and Deterministic Intervention Adjustment (Δ)

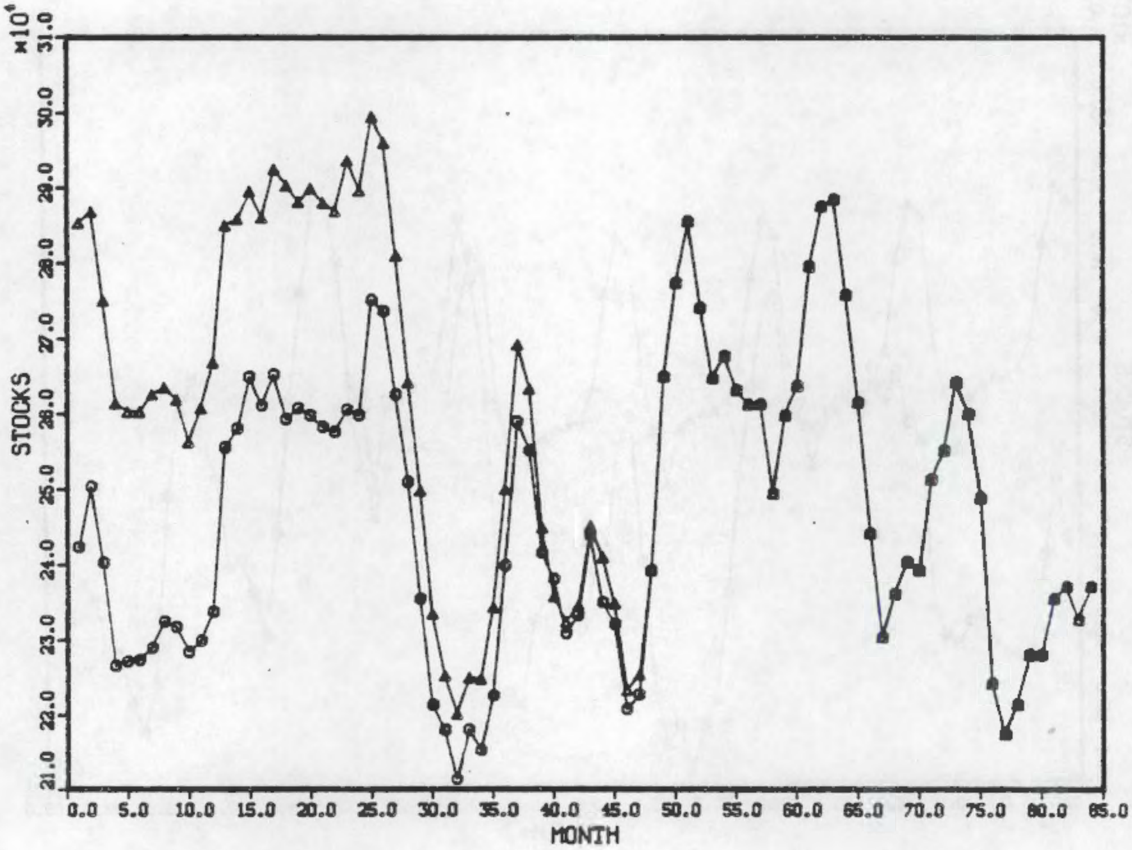

FIGURE 35. Gasoline Stocks (O) and Intervention Adjustment Incorporating Estimated Random Component (Δ)

that the major differences between the past data and the adjusted data are in the first 28 months, where the adjusted data were quite a bit greater.

### 5.5.2 Adjustment Errors

As noted above, the errors of a Box-Jenkins type forecast for intervention adjustment are well understood up to the assumptions outlined in Table 7. However for the type of adjustment that uses the estimated shocks, the error structure is of necessity more complicated since the sources of error are exactly those taken to be unimportant by the Box-Jenkins procedure. To better explain the errors a different form of the models will be of use.

As in Section 4.1, let $w_t$ denote the differenced data and suppose the $w_t$'s are represented by the stationary invertible model:

$$\phi(B)\ w_t = \theta(B)\ a_t \tag{9}$$

where as before B is the backshift operator and the $(a_t)$ are a series of independent identically distributed random variables with mean zero and common variance $\sigma_a^2$. A surprisingly useful dual form of the model can be obtained by replacing B with $B^{-1} = F$, that is $F(z_t) = z_{t+1}$ and by replacing the $a_t$'s with $e_t$'s where $(e_t)$ are also a series of independent identically distributed random variables with mean zero and variance $\sigma_a^2$. This "backward" form of the model is then

$$\phi(F)\ w_t = \theta(F)\ e_t \tag{10}$$

and expresses current values of the series in terms of future values of the series and future shocks. It has the same correlation structure as the "forward" model and is thus appropriate for "forecasting" or adjusting in the past. Psi-weight and pi-weight forms of the model (10) can be obtained in the same way as for the standard model.

In the derivation of formulas and calculation methods for the adjusted values of the pre-intervention series, it is useful to reiterate the steps by which the model and adjustments are obtained. As above, let $w_t$ and $\tilde{w}_t$ denote the data and the adjusted values, respectively, after differencing to achieve stationarity. The pre-intervention data are used to fit a model which can be written in the form:

$$\hat{\Pi}(B)\ w_t = \hat{a}_t \tag{11}$$

or in backward form as

$$\hat{\Pi}(F)\ w_t = \hat{e}_t\ . \tag{12}$$

The "hat" is used here to emphasize the fact that the quantities are estimated. When the model (12) is used to produce the estimated "shocks", $\hat{e}_t$, note also that there is a "start-up" effect. That is; the pi-weights are an infinite sequence whereas there are only finitely many $w_t$'s. Thus for $t < 1$ and $t > N$, $z_t$ and hence $w_t$ are taken to be zero which results in the $\hat{e}_t$ being zero in that region also.

Once the $\hat{e}_t$ have been obtained, the estimated <u>post-intervention model</u> (in psi-form) is used as follows to estimate adjusted data:

$$\tilde{w}_t = \hat{\tilde{\psi}}(F)\ \hat{e}_t \qquad t = t_o,\ t_o - 1,\ \ldots,\ 1\ . \tag{13}$$

Thus from (12) and (13):

$$\begin{aligned}\tilde{w}_t &= \hat{\pi}(\tilde{f})\ \hat{\tilde{\psi}}(F)\ w_t \\ &= \sum_{k=0}^{\infty} \hat{\gamma}_k F^k \ w_t \\ &= \sum_{k=0}^{N-t} \hat{\gamma}_k w_{t+k} \ .\end{aligned} \tag{14}$$

As previously noted the adjusted values so computed are estimated values since there are only finitely many $w_t$'s and since there are estimation errors in the $\hat{\gamma}_k$'s. Let $\tilde{\omega}_k$ stand for the "true" differenced adjusted values and assume without loss of generality that $E\tilde{w}_t = 0$ and $E\tilde{\omega}_t = 0$. In analogy to (14):

$$\tilde{\omega}_t = \sum_{k=0}^{\infty} \gamma_k w_{t+k} \tag{15}$$

Thus from (14) and (15);

$$\begin{aligned}\tilde{\omega}_t - \tilde{w}_t &= \sum_{k=0}^{N-t} (\gamma_k w_{t+k} - \hat{\gamma}_k w_{t+k}) + \sum_{k=N-t+1}^{\infty} \gamma_k w_{t+k} \\ &= \sum_{k=0}^{N-t} (\gamma_k - \hat{\gamma}_k)\ w_{t+k} + \sum_{k=N+1}^{\infty} \gamma_{k-t}\ w_k \ .\end{aligned} \tag{16}$$

The mean-square error in the adjusted values can thus be written

$$\begin{aligned}E(\tilde{w}_t - \tilde{\omega}_t)^2 &= E\left[\sum_{k=0}^{N-t} (\gamma_k - \hat{\gamma}_k)\ w_{t+k}\right]^2 \\ &+ E\left[\sum_{k=N+1}^{\infty} \gamma_{k-t}\ w_k\right]^2 + 2E\left[\sum_{k=0}^{N-t} (\gamma_k - \gamma_k) w_{t+k} \sum_{k=N+1}^{\infty} \gamma_{k-t} w_k\right] .\end{aligned} \tag{17}$$

The expression (17) is generally intractable computationally. To derive a more useful approximation, note that for $\ell$ periods ($\ell = 0, 1, \ldots, t_0-1$) before the intervention point $t_0$, the true adjusted value of the original, undifferenced series can be written as:

$$\tilde{z}_{t_o-\ell} = \sum_{k=0}^{\infty} \Psi_k e_{t_o-\ell+k} \tag{18}$$

where the $\Psi_k$'s correspond to the post-intervention series. Assuming that the estimation errors in the psi-weights contribute little to the error and that the dependence on far future values is slight, an "estimated" adjusted value is then

$$\hat{\tilde{z}}_{t_o-\ell} = \hat{e}_{t_o-\ell} + \Psi_1 \hat{e}_{t_o-\ell+1} + \dots + \Psi_\ell \hat{e}_{t_o} + \sum_{k=\ell+1}^{\infty} \Psi_k e_{t_o-\ell+k} \; . \tag{19}$$

"Estimated" is in quotes since the estimated adjusted value is not and could not be calculated in this way. This form, however, permits the calculation of the mean square error of estimation. The error of adjustment is then

$$\epsilon_{t_o-\ell} = \tilde{z}_{t_o-\ell} - \hat{\tilde{z}}_{t_o-\ell} = \sum_{k=0}^{\ell} \Psi_k \left(e_{t_o-\ell+k} - \hat{e}_{t_o-\ell+k}\right) \; , \tag{20}$$

and therefore, the mean-square error of adjustment will thus be approximated by

$$E\, \epsilon^2_{t_o-\ell} = \sum_{k=0}^{\ell} \Psi_k^2 \, E\left(e_{t_o-\ell+k} - \hat{e}_{t_o-\ell+k}\right)^2 \; . \tag{21}$$

It is interesting to compare this error of adjustment to the usual forecast error variance in the Box-Jenkins procedure. Recall that for the Box-Jenkins procedure the forecast error for a $\ell + 1$ lead forecast would be

$$\sum_{k=0}^{\ell} \Psi_k \, a_k \; . \tag{22}$$

Thus the forecast variance is

$$\sum_{k=0}^{\ell} \Psi_k^2 \, \sigma_a^2 \; . \tag{23}$$

Note that the expression (21) would reduce to this if all of the $\hat{e}_k$'s were taken to be zero. Furthermore, since it is reasonable to expect that the $\hat{e}_k$'s will be positively correlated with the $e_k$'s, the mean-square error of adjustment should be smaller than the Box-Jenkins forecast variance.

It is now necessary to assess the magnitude of $E(e_k - \hat{e}_k)^2$. Since the $\hat{e}_t$'s are computed from the pre-intervention model, (12) is used to write:

$$\hat{e}_k = 1 - \sum_{i=1}^{t_o-t} \hat{\pi}_k F^i w_t \qquad (t = 1, 2, \ldots, t_o-1) \ . \tag{24}$$

Thus, ignoring the dependence of the $e_t$'s on far future values of the $w_t$'s

$$e_t - \hat{e}_t = \sum_{i=1}^{t_o-t} (\pi_i - \hat{\pi}_i) w_{t+i} \ , \tag{25}$$

and the mean-square error is then for $t = 1,2,\ldots,t_o-1$.

$$E(e_t - \hat{e}_t)^2 = E \quad \sum_{i=1}^{t_o-t} (\pi_i - \hat{\pi}_i) w_{t+i} \quad {}^{2} \ , \tag{26}$$

which will be approximated by

$$\begin{aligned} E(e_t - \hat{e}_t)^2 &\cong \sum_{i=1}^{t_o-t} E(\pi_i - \hat{\pi}_i)^2 \, \sigma_w^2 + 2 \sum_{i=1}^{t_o-t} \sum_{j=i+1}^{t_o-t} E(\pi_i - \hat{\pi}_i)(\pi_j - \hat{\pi}_j) \operatorname{cov}(w_i, w_j) \\ &= \sum_{i=1}^{t_o-t} \Pi_{ii} W_{ii} + 2 \sum_{i=1}^{t_o-t} \sum_{j=i+1}^{t_o-t} \Pi_{ij} W_{ij} \ , \end{aligned} \tag{27}$$

where $\underset{\sim}{\Pi}$ and $\underset{\sim}{W}$ are the covariance matrices of the pi's and the w's, respectively and the ij subscript indicates the ij^th term. For $t = t_o$ we will take $(e_{t_o} - \hat{e}_{t_o}) = \sigma_a^2$. Note that for the above approximation it was assumed that the $\hat{\Pi}_i$ and $w_j$ were independent, which is not precisely true since the $\hat{\Pi}$'s are estimated from the w's. Furthermore, since the w's form a stationary series, we could write $W_{ij} = R(i-j)$.

$\underset{\sim}{W}$ is determined completely by the form of the pre-intervention model, whereas

$$\pi(B) = 1 + \pi_1 B + \ldots = \frac{\phi(B)\ \Phi(B)}{\theta(B)\ \Theta(B)} \ . \qquad (28)$$

Thus $\Pi$ can be obtained by propogating the parameter estimation errors. That is, any $\pi_j$ is some function of the vector of parameters, say $\underset{\sim}{\beta}' = (\beta_1, \beta_2, \ldots, \beta_p)$. Writing

$$\pi_j = f(\underset{\sim}{\beta})$$

and expanding with a Taylor series about the estimated parameters $\hat{\beta}' = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_2)$ gives

$$\hat{\pi}_j = \hat{\pi}_j + \sum_{i=1}^{p} \frac{\partial f}{\partial \beta_i}\bigg|_{\underset{\sim}{\hat{\beta}}} (\beta_i - \hat{\beta}_i) + \frac{1}{2} \sum_{i=1}^{p} \frac{\partial^2 f}{(\partial \beta_i)^2}\bigg|_{\underset{\sim}{\hat{\beta}}} (\beta_i - \hat{\beta}_i)^2$$

$$+ \sum_{i=1}^{p} \sum_{k=i+1}^{p} \frac{\partial^2 f}{\partial \beta_i \partial \beta_k}\bigg|_{\underset{\sim}{\hat{\beta}}} (\beta_i - \hat{\beta}_i)(\beta_k - \hat{\beta}_k) + \text{higher order terms.}$$

Assuming that the distribution of $\hat{\beta}$ is symmetric about $\hat{\beta}$ and that the higher order terms can be ignored, the mean square error of $\hat{\pi}_j$ can be estimated as

$$E(\pi_j - \hat{\pi}_j)^2 \cong \sum_{i=1}^{p} \left[\frac{\partial f}{\partial \beta_i}\bigg|_{\underset{\sim}{\hat{\beta}}}\right]^2 \text{Var } \hat{\beta}_i$$

$$+ 2 \sum_{i=1}^{p} \sum_{k=i+1}^{p} \left[\frac{\partial f}{\partial \beta_i}\bigg|_{\underset{\sim}{\hat{\beta}}}\right] \left[\frac{\partial f}{\partial \beta_k}\bigg|_{\underset{\sim}{\hat{\beta}}}\right] \text{Cov}(\hat{\beta}_i, \hat{\beta}_k)$$

$$+ \frac{1}{4} \sum_{i=1}^{p} \left[\frac{\partial^2 f}{(\partial \beta_i)^2}\bigg|_{\underset{\sim}{\hat{\beta}}}\right]^2 E(\beta_i - \hat{\beta}_i)^4$$

and the covariance of $\hat{\pi}_j$ and $\hat{\pi}_k$ as

$$E(\pi_j-\hat{\pi}_j)(\pi_k-\hat{\pi}_k) = \sum_{i=1}^{p} \sum_{\ell=1}^{p} \left.\frac{\partial f}{\partial \beta_i}\right|_{\hat{\beta}_i} \left.\frac{\partial g}{\partial \beta_\ell}\right|_{\hat{\beta}_i} \mathrm{Cov}(\hat{\beta}_i,\hat{\beta}_\ell),$$

where $\pi_k = g(\underset{\sim}{\beta})$.

It is likely that in any application only a few of the $\Pi$ terms will be required.

This procedure for calculating the adjustment errors will be illustrated on the motor gasoline series. Since the model for the pre-intervention motor gasoline series is a (0, 1, 1) x $(0, 1, 1,)^{12}$ series with estimated parameters

$\hat{\theta} = -0.237$ and $\hat{\Theta} = 0.755$ and $\sigma_a^2 = 49701116$ the covariance structure of the $w_t$'s is

$$R(0) = W_{ii} = (1 + \hat{\theta}^2)(1 + \hat{\Theta}^2)\sigma_a^2 = 1.65821\ \sigma_a^2$$

$$R(1) = W_{i,i+1} = -\hat{\theta}(1 + \hat{\Theta}^2)\sigma_a^2 = 0.37210\ \sigma_a^2$$

$$R(11) = W_{i,i+11} = \hat{\theta}\hat{\Theta}\sigma_a^2 = -0.17894\ \sigma_a^2$$

$$R(12) = W_{i,i12} = -\hat{\Theta}(1 + \hat{\theta}^2)\sigma_a^2 = -0.79741\ \sigma_a^2$$

$$R(13) = W_{i,i+13} = \hat{\theta}\hat{\Theta}\sigma_a^2 = -0.17894\ \sigma_a^2$$

All other $W_{ij} = 0$ .

Since

$$\pi(B) = \frac{1}{(1 + 0.237B)(1 - 0.755B^{12})}$$

the pi-weights are

$$\pi_0 = 1$$

$$\pi_j = -\theta^j \qquad 1 \le j \le 11$$

$$\pi_{12} = -(\theta^{12} + \Theta)$$

$$\pi_j = -\theta^{j-12}\,(\theta^{12} + \Theta), \qquad 13 \le j \le 23$$

$$\pi_{24} = -(\theta^{24} + \Theta^2)$$

$$\pi_j = -\theta^{j-24}\,(\theta^{24} + \Theta^2), \qquad 25 \le j \le 35$$

$$\vdots$$

The standard deviations of the estimates of $\theta$ and $\Theta$ (0.149 and 0.174, respectively) and their correlation, -.027, can be propagated to produce the $\Pi$ matrix as follows. Writing $i = 12k + \ell$ for $k = 0,1,2,\ldots$ and $0 \le \ell \le 11$ gives

$$\pi_i = -\theta^{\ell}(\theta^{12k} + \Theta^k)$$

for $i \ge 12$ and $\pi_i = -\theta^i$ for $0 \le i < 12$. Thus if $i = 12k+\rho \ne j = 12k'+\ell'$,

$$\begin{aligned}
\Pi_{ij} &= E(\pi_i - \hat{\pi}_i)(\pi_j - \hat{\pi}_j) \\
&\cong [(12k+\ell)\hat{\theta}^{12k+\ell-1} + X(k)\ell\hat{\theta}^{\ell-1}\hat{\Theta}^{k}][(12k'+\ell')\hat{\theta}^{12k'+\ell'-1} \\
&+ X(k')\ell'\hat{\theta}^{\ell'-1}\hat{\Theta}^{k'}]\mathrm{Var}(\hat{\theta}) \;+ kk'\hat{\theta}^{\ell+\ell'}\hat{\Theta}^{k+k'-2}\mathrm{Var}(\hat{\Theta}) \\
&+ [((12k+\ell)\hat{\theta}^{12k+\ell-1} + X(k)\ell\hat{\theta}^{\ell-1}\hat{\Theta}^{k})(k'\hat{\theta}^{\ell}\hat{\Theta}^{k'-1}) + ((12k'+\ell')\hat{\theta}^{12k'+\ell'-1} \\
&+ X(k')\ell'\hat{\theta}^{\ell'-1}\hat{\Theta}^{k'})(k\hat{\theta}^{\ell}\hat{\Theta}^{k-1})]\;\rho(\hat{\theta},\hat{\Theta})\sqrt{\mathrm{Var}\hat{\theta}\;\mathrm{Var}\hat{\Theta}}\;,
\end{aligned}$$

where $X(k) = \begin{cases} 0 \text{ if } k = 0 \\ 1 \text{ if } k > 0 \end{cases}$ , and

$$\Pi_{11} = \mathrm{Var}\;\hat{\theta};$$

$$\Pi_{ii} \cong i^2\hat{\theta}^{2i-2}\,\mathrm{Var}\;\hat{\theta} + \frac{3i^2(i-1)^2}{4}\hat{\theta}^{\,2i-4}\,(\mathrm{Var}\;\hat{\theta})^2, \quad 2 \le i \le 11\;;$$

$$\Pi_{ii} = [(12k+\ell)\hat{\theta}^{12k+\ell-1} + \ell\hat{\theta}^{\ell-1}\hat{\Theta}^{k}]^2 \quad \mathrm{Var}(\hat{\theta})$$

$$+ k^2\hat{\theta}^{2\ell}\hat{\Theta}^{2k-2}\ \mathrm{Var}(\Theta)$$

$$+ 2k\hat{\theta}^{\ell}\hat{\Theta}^{k-1}(12k+\ell)\hat{\theta}^{12k+\ell-1} + \ell\hat{\theta}^{\ell}\hat{\Theta}^{k})\rho(\hat{\theta},\hat{\Theta})\ \sqrt{\mathrm{Var}\ \hat{\theta}\ \mathrm{Var}\ \hat{\Theta}}$$

$$+ \frac{3}{4}\left[(12k+\ell)(12k+\ell-1)\hat{\theta}^{12k+\ell-2} + \ell(\ell-1)\theta^{\ell-2}\Theta^{k}\right]^2 (\mathrm{Var}\ \hat{\theta})^2$$

$$+ \frac{3}{4}\, k^2\ (k-1)^2\ \theta^{2\ell}\Theta^{2k-4}\ (\mathrm{Var}\ \Theta)^2 \quad ;$$

for $i = 12k+\ell \geq 12$ .

The Ψ-weights are calculated from

$$\Psi(B) = \sum_{i=0}^{\infty} B^{i} \sum_{i=0}^{\infty} B^{12i}\ (1 + 0.555B)(1 - 0.755B^{12}) \qquad (29)$$

and thus are

$\psi_1 = \psi_2 = \ldots = \psi_{11} = 1 - \theta = 1.555$ $\qquad \psi_{12} = 2 - \theta - \Theta = 1.800$

$\psi_{13} = \psi_{14} = \ldots = \psi_{23} = (1-\theta)(2-\Theta) = 1.936$ $\qquad \psi_{24} = (1-\theta)(2-\Theta) + (1-\Theta) = 2.181$

$\psi_{25} = \psi_{26} = \ldots = \psi_{35} = (1-\theta)(3-2\Theta) = 2.317$ $\qquad \psi_{36} = (1-\theta)(3-2\Theta) + (1-\Theta) = 2.562$

.
.
.

Substituting these values into the above formulas results in the mean-square errors of the adjusted values given in Table 8. Figure 36 is a plot of the data and the Box-Jenkins type adjusted values showing a one-sigma error band for the adjusted value. Figure 37 is a plot of the data and second method of intervention adjustment with an error band equal to the square root of the mean-square error of adjustment. As can be seen, the second method of intervention adjustment produces estimates that not only track the form of the original data better, but also have a much smaller error.

TABLE 8. Mean-Square Errors of Adjusted Values

| Time | Adjusted Value | SQRT (Mean-Square Error) |
|---|---|---|
| 1 | 285362 | 45605.2 |
| 2 | 286866 | 44941.3 |
| 3 | 274976 | 44267.5 |
| 4 | 261434 | 43583.3 |
| 5 | 260256 | 42888.1 |
| 6 | 260327 | 42181.5 |
| 7 | 262583 | 41462.9 |
| 8 | 263473 | 40731.6 |
| 9 | 261867 | 39987.2 |
| 10 | 256286 | 39229.8 |
| 11 | 260771 | 38458.8 |
| 12 | 266801 | 37154.9 |
| 13 | 284963 | 35570.1 |
| 14 | 285774 | 35026.0 |
| 15 | 289494 | 34473.3 |
| 16 | 285993 | 33911.8 |
| 17 | 292455 | 33340.7 |
| 18 | 290255 | 32759.6 |
| 19 | 288126 | 32168.0 |
| 20 | 289932 | 31565.5 |
| 21 | 287998 | 30951.5 |
| 22 | 286934 | 30326.3 |
| 23 | 293552 | 29689.2 |
| 24 | 289591 | 28450.3 |
| 25 | 299483 | 26867.9 |
| 26 | 295955 | 26430.2 |
| 27 | 280960 | 25985.2 |
| 28 | 264213 | 25532.3 |
| 29 | 249897 | 25071.4 |
| 30 | 233524 | 24601.7 |
| 31 | 225345 | 24123.0 |
| 32 | 220183 | 23634.6 |
| 33 | 225011 | 23136.3 |
| 34 | 224822 | 22628.5 |
| 35 | 234395 | 22110.4 |
| 36 | 250122 | 20901.6 |
| 37 | 269202 | 19279.5 |
| 38 | 263318 | 18953.4 |
| 39 | 245018 | 18621.5 |
| 40 | 235771 | 18283.7 |
| 41 | 232439 | 17939.5 |
| 42 | 234634 | 17588.5 |
| 43 | 245059 | 17230.4 |
| 44 | 241037 | 16864.8 |
| 45 | 234789 | 16491.5 |
| 46 | 223288 | 16111.5 |
| 47 | 225429 | 15723.6 |
| 48 | 239322 | 14530.1 |
| 49 | 264882 | 12829.1 |

TABLE 8. Mean-Square Errors of Adjusted Values (Continued)

| Time | Adjusted Value | SQRT (Mean-Square Error) |
|---|---|---|
| 50 | 277406 | 12642.8 |
| 51 | 285505 | 12453.6 |
| 52 | 274069 | 12261.6 |
| 53 | 264723 | 12066.5 |
| 54 | 267691 | 11868.2 |
| 55 | 263123 | 11666.6 |
| 56 | 261194 | 11461.5 |
| 57 | 261323 | 11253.5 |
| 58 | 249510 | 11045.0 |
| 59 | 259782 | 10843.8 |
| 60 | 263707 | 6921.1 |

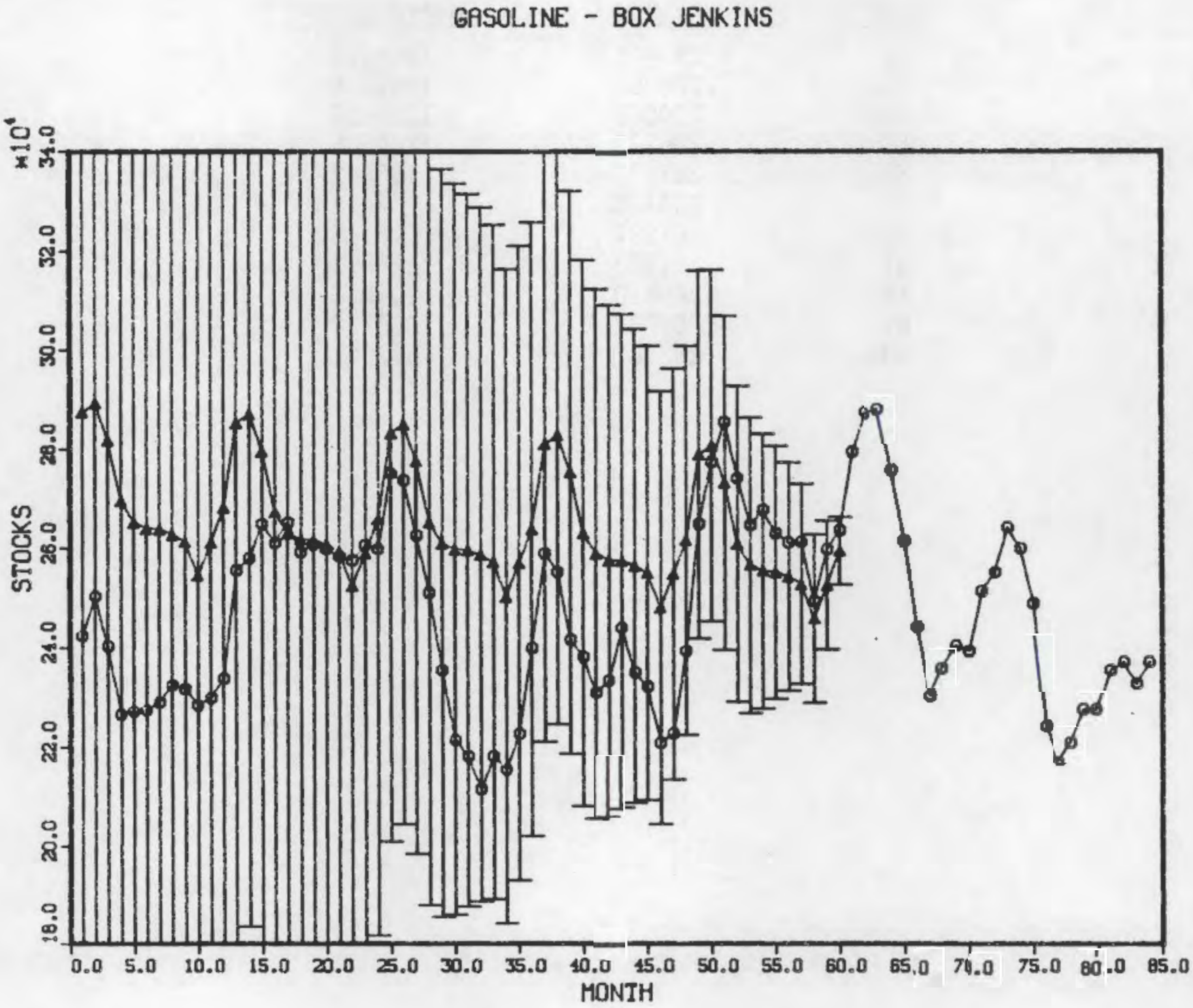

FIGURE 36. Box-Jenkins Type Adjustment with One-Sigma Error Bar

GASOLINE

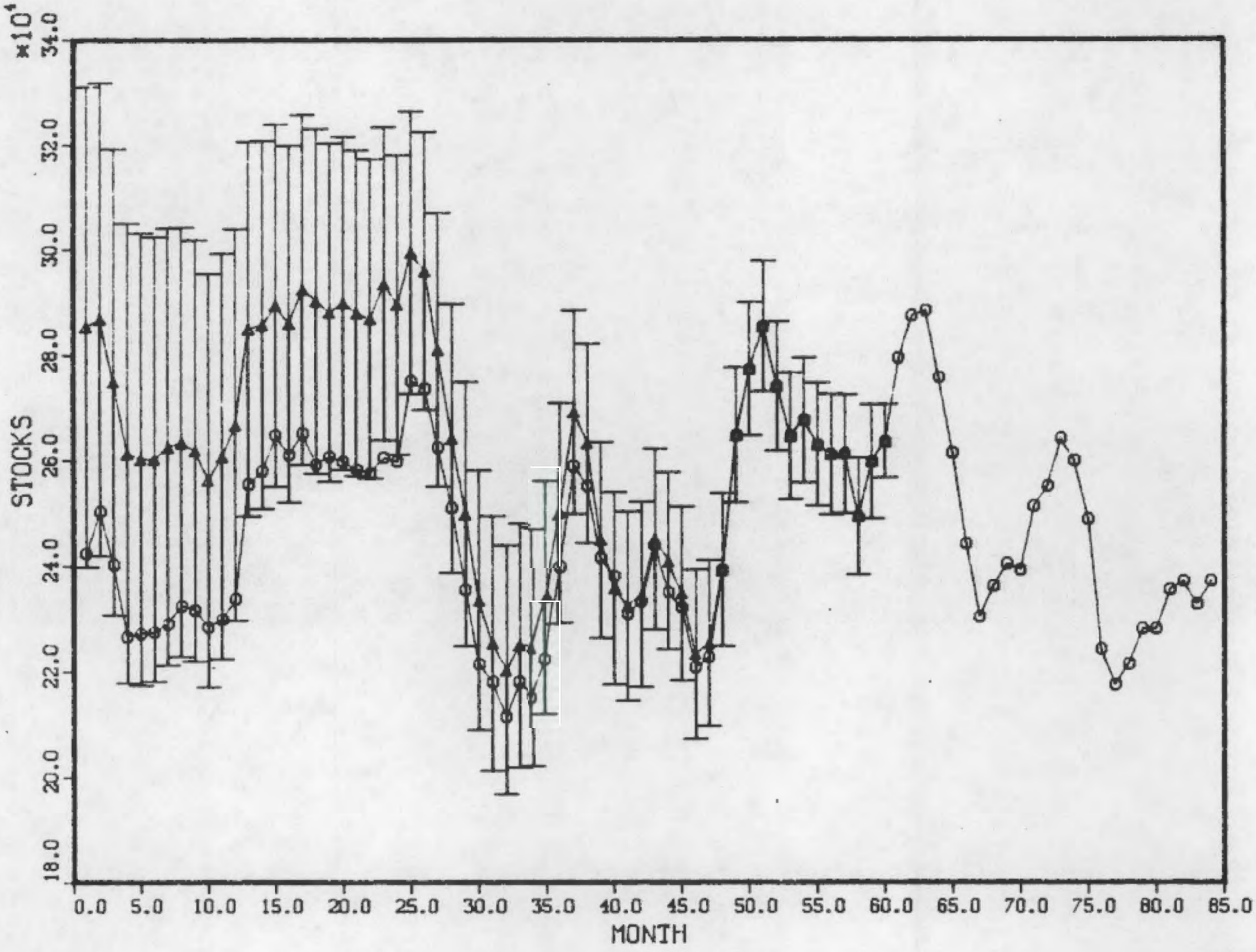

FIGURE 37. Adjustment Using Estimated Shocks with One-Sigma Error Bar

# 6.0 MODELING AND ADJUSTMENT FOR INTERVENTION USING CUMULATIVE TRANSITIONAL STATE SCORE AND CUMULATIVE SUMS

## 6.1 FRAMEWORK FOR CTSS AND CUSUM INDICES

In this section of the report we outline the basic principles underlying two types of cumulative scores which may be applied to monitoring changes in time-series data. The first is the cumulative transitional state score (CTSS) developed by Kumbaraci (1974) and Gardenier (1979); the second are cumulative sums (CUSUMs), as described in Barnard (1959), Johnson (1961), Johnson and Leone (1962), and DeBruyn (1961). CUSUM indices are based on Wald's sequential probability ratio test (SPRT) developed in the early 1940s, which Page (1961) applied to the testing of sequential hypotheses. Because CTSS is a newer index, there are fewer published references relating to its derivation and application. However, in depicting the short-term pattern of fluctuations in time-series data, CTSS appears to have advantages over CUSUMs.

Both indices are oriented toward identifying whether statistically significant changes are apparent in the data prior to and post intervention. Application of CUSUMs also may result in information sufficient to incorporate an adjustment to the long-term process mean value when a significant shift is observed.

### 6.1.1 Methodological Background

The CUSUM index aggregates deviations from a long-term process average; thus one cumulates the sum of $(X_i - m)$, where m corresponds to a reference value or a long-term process average. The reference value is denoted m because a mean is usually used. In applying CUSUMs, one uses two parameters, h and k in deciding whether the overall process has changed. The k is usually a multiplier on standard deviation units that is added to or subtracted from the reference value m prior to cumulating deviations. Let a long-term process average be denoted by $\bar{X}$. Then an interval $\bar{X} \pm k\sigma$ is established as an acceptable region beyond which deviations are cumulated. If the cumulative deviations reach the value of $\pm h$, the decision is made that the mean of the process has shifted. The choice of k is usually done by the user, depending upon the long-term reliability of the data and the consequences of not detecting a shift. The choice of h is dependent upon the basic principles of hypothesis testing, as outlined below.

Define $H_0$ as the hypothesis that the process is in control and $a_0$ as the probability of a false alarm. The $a_0$ is usually denoted by $\alpha$ in hypothesis-testing terminology. Also define $H_1$ as the probability that the new data are biased by g standard deviation units and $a_1$ as the probability of $H_0$ when $H_1$ is true (usually denoted as $\beta$ in hypothesis testing). For successive observations the ratio of the likelihood of observed values under $H_0$ and $H_1$ is calculated and used to decide whether to accept $H_0$, $H_1$ or to continue sampling.

The process mean, $\bar{X}$ is accepted if:

$$\frac{1}{\sigma}\sum_{i=1}^{n}(X_i - \overline{X}) \le \frac{1}{g}\ln\frac{a_1}{1 - a_0} + \frac{1}{2}Ng .$$

An alternative $\bar{X} + g\sigma$ is accepted as the new mean if:

$$\frac{1}{\sigma}\sum_{i=1}^{n}(X_i - \overline{X}) \ge \frac{1}{g}\ln\frac{1 - a_1}{a_0} + \frac{1}{2}Ng .$$

One continues to observe the process until one of the two decisions is reached.

In the literature, the run length of the CUSUM distribution is defined as the number of observations until $H_0$ is rejected. A small value for run length implies timely detection of shifts in the process mean.

Woods and Pike (1981) applied these concepts to detecting cumulative inventory differences and pointed out that reducing h for fixed values of k increases the false alarm rate. Other applications of the CUSUM technique to evaluating time-series data in materials inventory have been demonstrated by Cobb (1981), Markin and Shipley (1982) and Wincek et al. (1979).

In contrast to CUSUMs which are based upon the sum of deviations from a reference value m, the CTSS index is based upon the transitions between successive observations. States are defined in the CTSS procedure as regions or partitions, similar to zones in control charts used in quality control, using the statistical characteristics of previously observed values. An illustration of the definition of states is given in Figure 38; the summation scheme based upon the transitions in successive states is shown in Table 9.

In Figure 38 we may trace a sequence of observations over time and categorize each transition as to whether it is in State +2, +1, 0, -1 or -2. In the present analyses we have defined the range +2 to -2 as the upper and lower limits of stock bands for the year 1979-1980 of Weekly Petroleum Status Report (EIA 1983). The Weekly Petroleum Status Report updates the bands every six months. Thus, both the level and width of the bands may change. Because we wanted to establish fairly uniform limits against which fluctuations prior to and past intervention could be traced, we chose a specific year as a basis to set the b and $\alpha$. This established a "background" for sequence-related analyses.

Stock bands were computed using the mean and standard error of past observations and a seasonal adjustment factor using the Census X-11 procedure (Bureau of Census 1967). This range was then interpolated to yield the limits for +1 to -1. The horizontal line y on the diagram represents a hypothetical

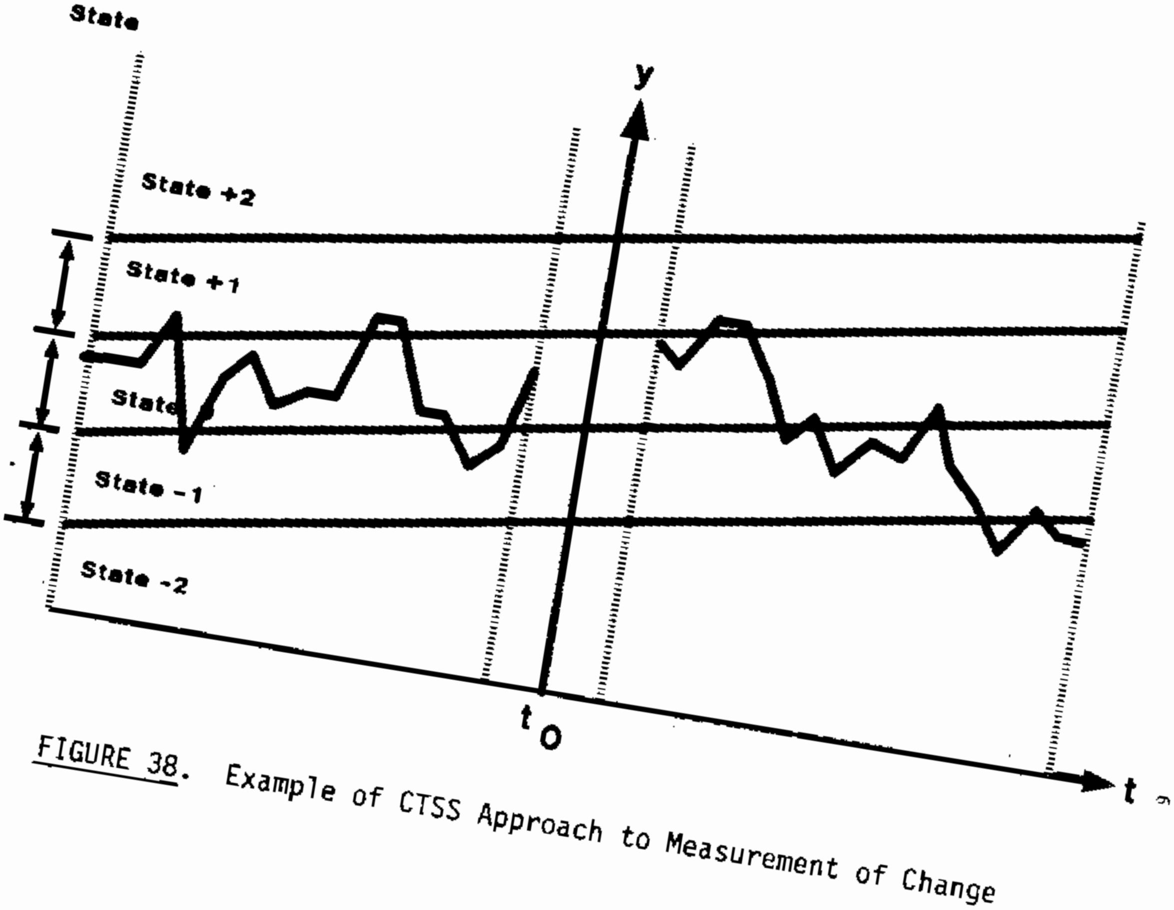

FIGURE 38. Example of CTSS Approach to Measurement of Change

TABLE 9. CTSS Transition Matrix with 5-State System

| Initial | | | Final State | | |
|---|---|---|---|---|---|
| State | -2 | -1 | 0 | +1 | +2 |
| -2 | -1 | +1 | +2 | +3 | +4 |
| -1 | -1 | 0 | +1 | +2 | +3 |
| 0 | -2 | -1 | 0 | +1 | +2 |
| +1 | -3 | -2 | -1 | 0 | +1 |
| +2 | -4 | -3 | -2 | -1 | +1 |

intervention point before and after which the state of the observations was to be evaluated.

Table 9 is a 5 x 5 matrix of the possible fluctuations from one time period and the next, $t_i$ to $t_{i+1}$. That is, if an observation is at State 0 at $t_i$, it may remain within the boundaries of the same state or it may fluctuate upward or downward to any of the other possible states. The same is true for fluctuations from any other row of the matrix depicting initial states.

CTSS incorporates a scoring scheme to this matrix of transitions. The cumulative sum increases or decreases depending upon the number of boundary crossings in adjoining time periods. Different evaluation scenarios may dictate different scoring schemes depending upon the pattern of past observations and their short-term transition patterns. In the present analysis, the cumulative sum was not incremented if the observations were in the -1, 0 or +1 range and remained therein. If at any time period observations were found to be outside the -2 to +2 range and remained in these states (i.e., no significant reversal toward the overall process characteristics was observed), an adjustment to the summation procedure was applied. The cumulative sum was incremented by +1 if the previous observation was in the +2 state; it was decreased by -1 if the previous observation was in the -2 state.

The framework for this methodology is reminiscent of CUSUMs in that it cumulates deviations. It differs from CUSUMs in that it evaluates the transition matrix for successive observations. In cases where the effect of intervention is to be evaluated, the pattern of observations in the proximity of the intervention is of interest. CTSS also allows for an "uncertainty zone" in each evaluation so that outliers do not significantly affect the results.

Statistical Markovian properties of time-series data and methods for evaluating transitions have been discussed by Billingsley (1961) and Whittle (1955). A nonparametric test for randomness in multinomial observations, which may be relevant to the partitioning of ranges presented in the CTSS methodology, is given by Bennett (1964). Sobel et al. (1973) used the Rao-Blackwell theorem to find the conditional expectation of CTSS up to Time T for the run of zero- or one-valued successive evaluations, where ones occur with probability p and q = 1 - p. The minimum unbiased estimator for the expectation of CTSS up to Time T and its variance are:

$$E(\mathrm{CTSS}\ T) = \frac{T}{n} + \frac{T}{n}(T - 1)\frac{(T - 1)}{n - 1} + (1 - \frac{T}{n})\ T(\frac{T - 1}{n - 1})$$

$$= \frac{T}{n} + \frac{T\ (T - 1)}{n\ (n - 1)}\ (T - 1 + n - T)$$

$$= \frac{T^2}{n}$$

$$\mathrm{Var}\ (T^2/n) = \frac{E\ T\ - (ET^2)^2}{n^2}$$

$$= \frac{pq}{n}\ [1 + 6(n - 1)p + 2(n - 1)(2n - 3)p^2]$$

### 6.1.2 Adjustment to Methods in Present Application

Initial exploratory application of the CUSUM technique to the four data series showed that $H_0$ was rejected often due to the seasonality exhibited in the data series. Thus, an adjustment to the CUSUM computations was made, correcting for seasonality in the data using the X-11 procedure. The CUSUM index in the present analyses cumulated $\bar{X} + S \pm k\sigma$, where S corresponded to the seasonal factor for each month. $\bar{X}$ was calculated for the years 1975-1978; k was chosen as ½-standard deviations of the variation in the series during the same time period; the rejection level h corresponded to 4k.

When $H_0$ was rejected at Time T for a particular series, an adjustment to the long-term process average was made using the following formula:

$$\bar{X}_N = \bar{X} \pm [N\ k + (CUSUM/N)]$$

where $\bar{X}_N$ corresponds to the new process average and N refers to the number of successive observations where positive deviations from $\bar{X}$ were found to occur. The process for monitoring CUSUMs was reinitiated after each such modification of the process mean.

In using CTSS, the transitional state score (TS) was recorded as well as the value for CTSS for each month from 1976-1982. CTSS was tested for significance using 6-month and yearly time intervals prior to and post intervention. The cut-off points for significance at $\alpha = 0.05$ and $\alpha = 0.01$ were established by reference to Table 10 for $T = 6$ and $T = 12$. Table 10 has been generated by simulating the possible values of CTSS for a range of time intervals. In addition, contingency tables were formed for each data series, tabulating by year, the number of observations where TS was observed as 0, $\geq 1$ or $\leq 1$. Reclassifying TS values which were greater than or less than 1 as +1 was done in order to minimize the number of zero entries in the contingency tables. A Chi-square test was applied to each such table to test for time-related effects.

## 6.2 APPLICATION OF CTSS ANO CUSUM TO DATA SERIES

This section includes the results, along with schematic displays, showing the application of CTSS and CUSUM indices to stocks of motor gasoline, distillate fuel oil, residual fuel oil and crude oil during the years 1976-1982. The data used are published estimates residing in EIA computer files which generate the Weekly Petroleum Status Report. The analysis used the single monthly figure in the computer file, not the composite generated for the

TABLE 10. Tolerance Limits for CTSS for $\alpha = 0.05$ and $\alpha = 0.01$

| T | Acceptable Value for S<br>$\alpha = 0.05$ | <br>$\alpha = 0.01$ |
|---|---|---|
| 4 | 4 | -- |
| 5 | 4 | 5 |
| 6 | 5 | 6 |
| 7 | 5 | 6 |
| 8 | 6 | 7 |
| 9 | 6 | 7 |
| 10 | 7 | 8 |
| 11 | 7 | 8 |
| 12 | 7 | 9 |
| 13 | 8 | 9 |
| 14 | 8 | 9 |
| 15 | 8 | 10 |
| 16 | 9 | 10 |
| 17 | 9 | 11 |
| 18 | 9 | 11 |
| 19 | 10 | 12 |
| 20 | 10 | 12 |

time-series analysis discussed in the previous chapter. The two estimates are quite close but not identical in some cases. Data are presented in monthly intervals; evaluations of CTSS are applied to transitions in adjacent months.

Figures 39 through 42 present the data superimposed upon the zones, from which the TS and the CTSS are calculated. Figure 39 shows motor gasoline, Figure 40 shows distillate fuel oil, Figure 41 shows residual oil, and Figure 42 shows crude oil. The first and second lines above the figures show the values of CTSS and TS at each month, using the rules presented in the preceeding methodology section.

Significance testing of CTSS was done using the cut-off points shown in Table 10 for yearly intervals as well as for 6-month intervals. Table 11 gives a summary presentation.

For motor gasoline, a significant increase was observed during 1977 and 1980, prior to the intervention point of January 1981. The same trend existed for distillate--a significant increase during 1977 and 1980. Distillate stocks showed a decrease in CTSS after mid-1981 and continued to decrease throughout 1982. CTSS differences for residual fuel oil started increasing in late 1978 and showed a significant increase until the intervention point (early 1981). A decrease which becomes significant by mid-1982 was observed thereafter.

Tables 12 through 15 show the results of an analysis of the matrix of transitions (TS scores) at annual intervals during 1976-1982. Results of the application of a Chi-square analysis to the total matrix, to the pre- versus post-intervention time periods, and each consecutive pair of two years is also shown. The number of columns in the Chi-square analyses has been consistently taken as three, collapsing TS scores more extreme than ±1 into ±1 in order to avoid difficulties in applying the analysis to a sparse matrix. In the results of Chi-square analyses presented at the foot of the tables, a p-value of 0.10 was used.

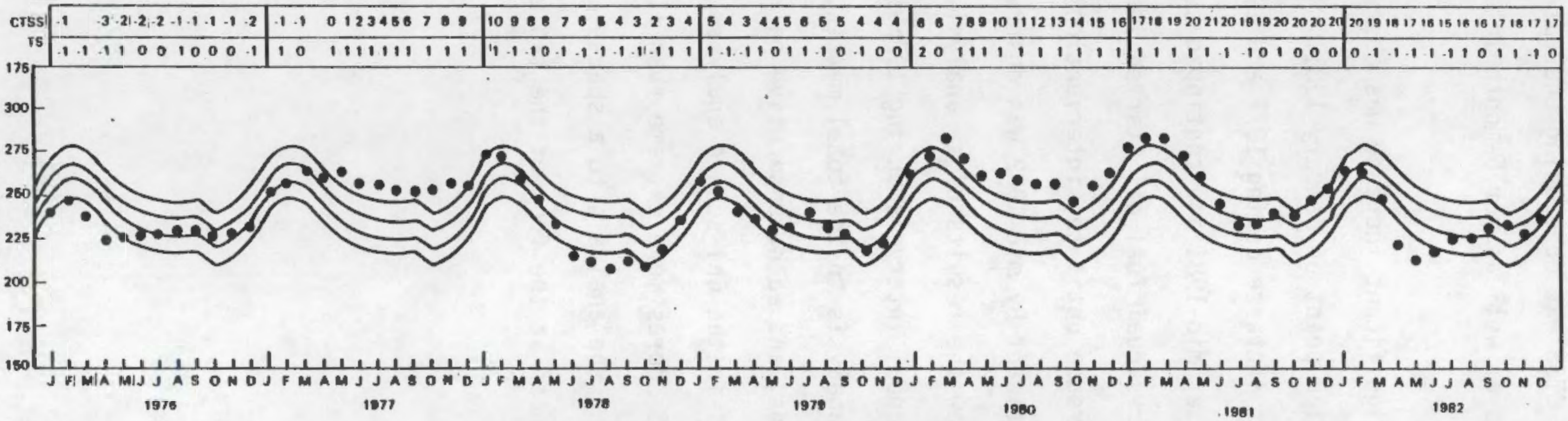

FIGURE 39. Evaluations for CTSS and TS for Motor Gasoline Stocks During 1976-1982

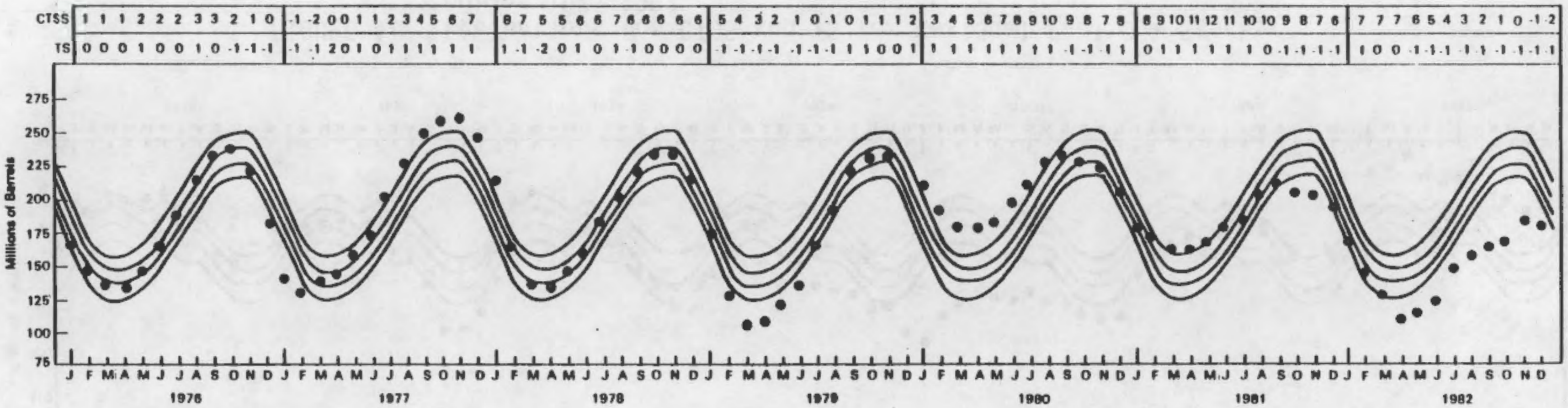

FIGURE 40. Evaluations for CTSS and TS for Distillate Stocks During 1976-1982

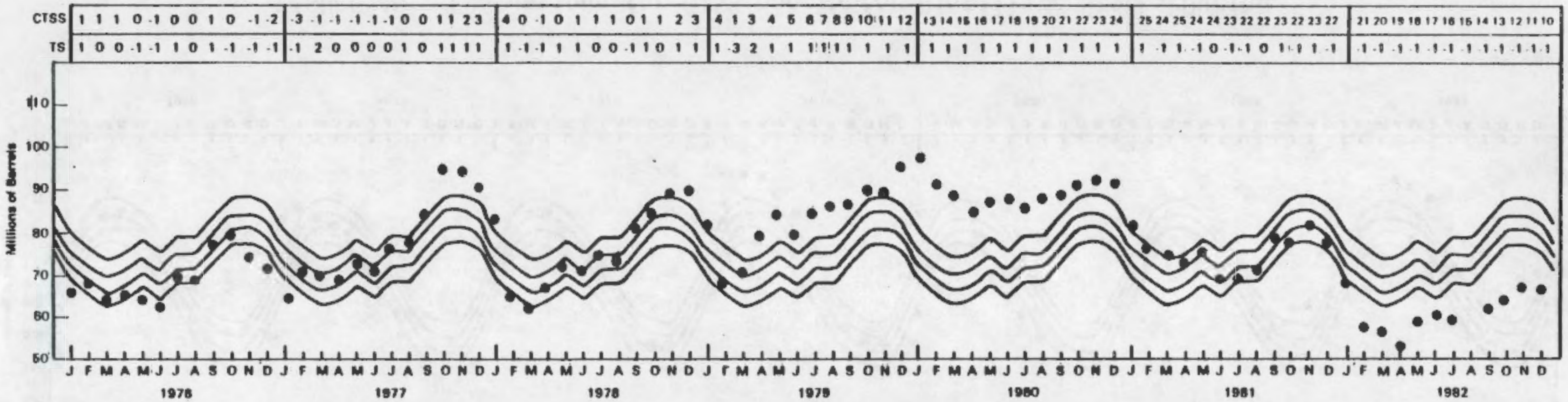

FIGURE 41. Evaluations for CTSS and TS for Residual Fuel Oil Stocks During 1976-1982.

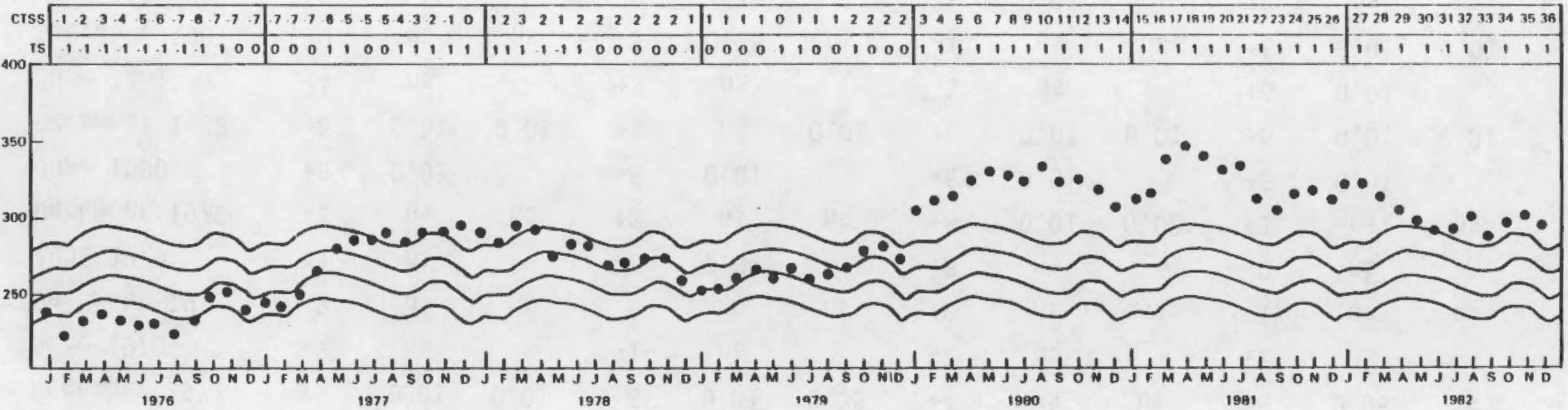

FIGURE 42. Evaluations for CTSS and TS for Crude Oil Stocks During 1976-1982

TABLE 11. CTSS Comparisons for Time Series at Yearly and Six-Month Intervals

| | Motor Gasoline | | | Distillate | | | Residual | | | Crude Oil | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Time | CTSS | $p_m$* | $p_a$* | CTSS | $p_m$* | $p_a$* | CTSS | $p_m$* | $p_a$* | CTSS | $p_m$* | $p_a$* |
| June 1976 | -1 | ns | | +1 | ns | | -1 | ns | | -4 | ns | |
| December 1976 | 0 | ns | ns | -2 | ns | ns | -2 | ns | ns | -2 | ns | ns |
| June 1977 | +5 | 0.05 | | +1 | ns | | +1 | ns | | -2 | ns | |
| December 1977 | +6 | 0.01 | 0.01 | +6 | 0.01 | 0.05 | +4 | ns | ns | +5 | 0.05 | ns |
| June 1978 | -3 | ns | | -1 | ns | | -2 | ns | | +2 | ns | |
| December 1978 | -2 | ns | ns | 0 | ns | ns | +2 | ns | ns | -1 | ns | ns |
| June 1979 | +1 | ns | | -6 | 0.01 | | +3 | ns | | 0 | ns | |
| December 1979 | -1 | ns | ns | +2 | ns | ns | +6 | 0.01 | 0.02 | +1 | ns | ns |
| June 1980 | +6 | 0.01 | | +6 | 0.01 | | +6 | 0.01 | | +6 | 0.01 | |
| December 1980 | +6 | 0.01 | 0.01 | +1 | ns | 0.05 | +6 | 0.01 | 0.01 | +6 | 0.01 | 0.01 |
| June 1981 | +4 | ns | | +3 | ns | | -1 | ns | | +6 | 0.01 | |
| December 1981 | 0 | ns | ns | -5 | 0.05 | ns | -1 | ns | ns | +6 | 0.01 | 0.01 |
| June 1982 | -5 | 0.05 | | -2 | ns | | -6 | 0.01 | | +4 | ns | |
| December 1982 | +2 | ns | ns | -6 | 0.01 | 0.05 | -6 | 0.01 | 0.01 | +6 | 0.01 | 0.01 |

---

NOTE: CTSS entries are differences between cumulative values of intervals being compared. Annual CTSS differences are obtained by the sum of 6-month interval differences.

*pm = p-value for 6-month basis of testing for time-series changes.
pa = p-value for annual basis of testing for time-series changes.

In summary, significant differences in the pattern of TS values were traced in all four data series during the 1976-1982 time interval. In the two years prior to the intervention, 1979-1981, significant differences were also observed in all four data series. For motor gasoline, distillate fuel oil and crude oil, the transition matrices stabilized by the end of 1981, about a year after the intervention. Wide fluctuations in the data were observed for residual fuel oil at a few times several years before the intervention combined with an upward trend. The 1981 intervention elicited a downward trend.

TABLE 12. Transition Matrix of TS Values in Adjacent Years for Motor Gasoline Stocks

| Year | -1 | 0 | +1 | +2 |
|---|---|---|---|---|
| 1976 | 4 | 5 | 2 | |
| 1977 | 0 | 1 | 11 | |
| 1978 | 8 | 1 | 3 | |
| 1979 | 4 | 4 | 4 | |
| 1980 | 0 | 1 | 10 | 1 |
| 1981 | 2 | 4 | 6 | |
| 1982 | 6 | 3 | 3 | |

Results of Statistical Comparisons

| | Chi-Square | d.f. | p |
|---|---|---|---|
| Full Matrix: 1976-1982 | 36.15 | 12 | <0.005 |
| Pre/Post Intervention | 1.45 | 2 | ns |
| 1976/1977 | 12.88 | 2 | <0.005 |
| 1977/1978 | 12.57 | 2 | <0.005 |
| 1978/1979 | 3.28 | 2 | ns |
| 1979/1980 | 8.34 | 2 | <0.02 |
| 1980/1981 | 4.76 | 2 | <0.10 |
| 1981/1982 | 3.14 | 2 | ns |

TABLE 13. Transition Matrix of TS Values in Adjacent Years for Distillate Fuel Oil Stocks

| Year | -2 | -1 | 0 | +1 | +2 |
|---|---|---|---|---|---|
| 1976 | | 3 | 5 | 3 | |
| 1977 | | 2 | 2 | 7 | 1 |
| 1978 | 1 | 2 | 6 | 3 | |
| 1979 | | 7 | 2 | 3 | |
| 1980 | | 3 | 0 | 9 | |
| 1981 | | 6 | 2 | 4 | |
| 1982 | | 9 | 2 | 1 | |

Results of Statistical Comparisons

| | Chi-Square | d.f. | p |
|---|---|---|---|
| Full Matrix: 1976-1982 | 29.67 | 12 | <0.005 |
| Pre/Post Intervention | 7.10 | 2 | <0.05 |
| 1976/1977 | 3.72 | 2 | ns |
| 1977/1978 | 4.47 | 2 | ns |
| 1978/1979 | 3.60 | 2 | ns |
| 1979/1980 | 6.60 | 2 | <0.05 |
| 1980/1981 | 4.92 | 2 | <0.10 |
| 1981/1982 | 2.40 | 2 | ns |

TABLE 14. Transition Matrix of TS Values in Adjacent Years for Residual Fuel Oil Stocks

| Year | -4 | -3 | -2 | -1 | 0 | +1 | +2 |
|---|---|---|---|---|---|---|---|
| 1976 | | | | 5 | 3 | 3 | |
| 1977 | | | | 1 | 5 | 5 | 1 |
| 1978 | 1 | | | 2 | 3 | 6 | |
| 1979 | | 1 | | | | 10 | 1 |
| 1980 | | | | | | 12 | |
| 1981 | | | | 6 | 2 | 4 | |
| 1982 | | | | 12 | | | |

Results of Statistical Comparisons

| | Chi-Square | d.f. | p |
|---|---|---|---|
| Full Matrix: 1976-1982 | 56.11 | 12 | <0.001 |
| Pre/Post Intervention | 9.61 | 2 | <0.01 |
| 1976/1977 | 4.13 | 2 | ns |
| 1977/1978 | 1.50 | 2 | ns |
| 1978/1979 | 5.47 | 2 | <0.10 |
| 1979/1980 | (a) | (a) | (a) |
| 1980/1981 | 12.00 | 2 | <0.005 |
| 1981/1982 | 8.00 | 2 | <0.02 |

---

(a) Not evaluable due to sparse matrix.

TABLE 15. Transition Matrix of TS Values in Adjacent Years for Crude Oil Stocks

| Year | -1 | 0 | +1 |
|---|---|---|---|
| 1976 | 8 | 2 | 1 |
| 1977 | | 5 | 7 |
| 1978 | 3 | 5 | 4 |
| 1979 | 1 | 9 | 2 |
| 1980 | | | 12 |
| 1981 | | | 12 |
| 1982 | | | 12 |

Results of Statistical Comparisons

| | Chi-Square | d.f. | p |
|---|---|---|---|
| Full Matrix: 1976-1982 | 77.05 | 12 | <0.001 |
| Pre/Post Intervention | 31.48 | 2 | <0.001 |
| 1976/1977 | 13.77 | 2 | <0.005 |
| 1977/1978 | 3.82 | 2 | ns |
| 1978/1979 | 2.81 | 2 | ns |
| 1979/1980 | 17.14 | 2 | <0.001 |
| 1980/1981 | (a) | (a) | (a) |
| 1981/1982 | (a) | (a) | (a) |

---

(a) Not evaluable due to sparse matrix.

As discussed previously, a parallel analysis using seasonally adjusted CUSUMs was applied to the four data series. Wherever $H_0$ was rejected, the long-term mean was adjusted upward or downward and the CUSUM computations were reinitiated. Table 16 summarizes the results of this analysis. It tracks the mean of each data series over time showing when it was updated as well as when adjustments made to the mean. A summary comparison of times for detecting change prior- and post-intervention is shown at the foot of the table. An initial scan of the results based upon CUSUMs shows no significant shift in the pattern of observations after intervention as compared to data prior to the intervention point. The one point of departure is in distillate stocks data where the only adjustment to the series was made a year after the intervention date.

TABLE 16. Changes Observed Prior and Post-Intervention and Adjustments to Mean Value Using CUSUMs

| Series | h | k | Initial $\bar{X}$ | ± | New $\bar{X}$ | Date |
|---|---|---|---|---|---|---|
| Motor Gasoline | 32 | ±8 | 239 | -20 | 219 | 04/1976 |
| | | | | +16 | 231 | 11/1976 |
| | | | | +16 | 247 | 02/1977 |
| | | | | +18 | 265 | 08/1977 |
| | | | | -16 | 249 | 04/1978 |
| | | | | -20 | 229 | 07/1978 |
| | | | | +29 | 258 | 03/1980 |
| | | | | -15 | 243 | 03/1982 |
| | | | | -17 | 226 | 05/1982 |
| Distillate | 80 | ±20 | 190 | -47 | 143 | 01/1982 |
| Residual | 16 | ±4 | 74 | +9 | 83 | 10/1977 |
| | | | | -14 | 69 | 03/1978 |
| | | | | +7 | 76 | 10/1978 |
| | | | | +10 | 86 | 07/1979 |
| | | | | +10 | 96 | 03/1980 |
| | | | | -12 | 84 | 10/1980 |
| | | | | -10 | 74 | 08/1981 |
| | | | | -9 | 65 | 04/1982 |
| | | | | -8 | 58 | 10/1982 |
| Crude Oil | 48 | ±12 | 300 | -12 | 288 | 07/1976 |
| | | | | +33 | 321 | 06/1977 |
| | | | | +21 | 342 | 12/1977 |
| | | | | -19 | 323 | 08/1978 |
| | | | | +26 | 349 | 01/1980 |
| | | | | +23 | 372 | 06/1980 |
| | | | | -29 | 343 | 01/1982 |

| | Time to Detect Change (months) | | |
|---|---|---|---|
| | Prior to Intervention Mean | Post Intervention Mean | Post Intervention to First |
| Motor Gasoline | 7.8 | 7.3 | 14 |
| Distillate | --- | --- | 12 |
| Residual | 7.2 | 7.3 | 8 |
| Crude | 9.6 | --- | 12 |

# 7.0 CONCLUSIONS

This section brings together the several conclusions scattered throughout the report. The conclusions are of two types: conclusions concerning the overall methodology of intervention analysis and adjustment, and conclusions concerning the specific time series analyzed in the report. The two main methods used in the report, Box-Tiao Intervention Analysis and Cumulative Transitional State Scores (CTSS), are compared and contrasted. Since CTSS and Cumulative Sums (CUSUMs) are quite similar in philosophy, the CUSUMs method will be discussed with CTSS. Some recommendations for further research are also made.

## 7.1 OVERALL CONCLUSIONS

The importance of intervention analysis in situations such as studied in this report is well enough established to require no further justification. Suffice it to say that it is the exception rather than the rule that a time series can be collected for any significant length of time without some important change occurring. The effect of adjustment of pre-intervention data to be consistent with post-intervention data is less well understood, but it seems obvious that adjustment of data must always be done with care. The assumptions necessary to use intervention-adjusted data are extensive and probably limit the usefulness of adjusted data to such things as inputs into models that require consistent data. The Box-Tiao-like method presented in this report is a straight-forward method to accomplish adjustment, making use of knowledge of the past and, more importantly, providing a method for assessing the variability of the adjustments.

Several approaches to studying an intervention are summarized in the report and detailed analyses are performed using three of the approaches. The first of these is not specifically an intervention analysis technique but may be applied appropriately in several situations.

If there are other variables available that could possibly affect the time series of interest, the first step that should be taken is to remove the effects of these variables. After the removal of the effects, the intervention may no longer be present, as evidenced in the case of residual fuel oil stocks. However, even if the intervention is still present, failure to account for these variables can complicate the analysis or possibly can cause incorrect modeling (Bell and Hillmer 1983).

Assuming that the effects of other independent variables have been removed, there are several approaches to modeling the resulting time series and studying the intervention. When the quantity of data permits, the most flexible method is to independently model the pre- and post-intervention data and to compare the resulting parameter estimates using their asymptotic Gaussian distributions. This method requires a minimum of additional assumptions. Unfortunately, in most applications, it appears that the quantity of post-intervention data is insufficient for this approach. A more common situation is that something has happened that has led people to suspect a change in a series at a known time. It is then usually of interest to determine if the intervention is significant and to identify its nature as soon as possible. One rarely has the luxury of waiting until enough data has been collected to apply this method.

For all of the series studied in this report, the number of post-intervention data points was only 24, not enough to permit independent pre- and post-intervention modeling. Therefore, two approaches were used that required more be known or assumed about the series, but that worked with limited data. The first method used was described by Box and Tiao (1976). As discussed in Section 4.1, this method requires good pre-intervention modeling and some notion of the expected effect of the intervention on the model. The time of the intervention must also be known. With these restrictions it is a quite useful technique since it can find changes in the level of a series and in the stochastic parameters. Since the parameter changes can be estimated, this technique also allows the adjustment of pre-intervention data. Philosophically, the method is in the mainstream of Box-Jenkins-Tiao time-series analysis. It centers around a "fit, forecast and test" approach. Though Box-Jenkins techniques were used for the fitting in the analyses described in this report, other methods could be used. In summary, the Box-Tiao method is generally useful and can be modified to incorporate any number of favored approaches. This method should be considered for the type of problem studied in this report.

The other techniques used to analyze the series in this report were CUSUMs and CTSS. These approaches are in the spirit of quality control and, as such, are designed more to monitor than to model a time series. Both methods accumulate a sum and raise a warning when the aggregate exceeds some predetermined limit. They do not necessarily require that the location of the intervention be known beforehand. Further, if the CUSUMs technique indicates that a significant shift in the mean of a process has occurred, an estimate of the adjustment to the mean is available.

As applied to the series analyzed in this report the standard techniques had to be modified to account for the 12-month seasonal trend inherent in the data series. Even when so modified, the methods are still designed primarily to detect a change in the level of the series. How well short-term level changes can be used to model changes in stochastic parameters of a series is not known.

Other possible problems with applying CTSS or CUSUMs methods to time series are the effect of the autocorrelation on the determination of the size of the acceptance regions or state boundaries and on statistical tests performed on transition scores. Whereas CUSUMs accumulates deviations from a mean, CTSS accumulates scored transitions of successive observations. Two effects of this scoring are a resistance to outliers and a reduction of the autocorrelations. However, more general aspects of the effect of the method used to score transitions on the properties of the test are not well understood. For detecting interventions these methods represent an alternative to the Box-Tiao procedure. For shifts in the mean, CTSS or CUSUMs identify the change fairly quickly. However, the limitations discussed above, especially the inability of these methods to provide estimates of changes in stochastic parameters, hinder the application of these techniques to general intervention analysis problems.

## 7.2 SPECIFIC CONCLUSIONS

For distillate fuel oil stocks, all approaches to intervention analysis identified changes in the series after January 1980 and after January 1981. The methods indicated that these changes were because of a decrease in the overall level of the series. As noted in Section 5.1, the changes in the survey forms should have had no effect on the distillate fuel oil stock series, and it was theorized that these observed changes were caused by general shifts in the use of distillate fuel oil. Attempts to explain the changes by changes in the number of companies reporting were unsuccessful. It should be stated again that care must be exercised in interpreting the change after January 1981, especially in light of the earlier observed change in the series.

As in the case of distillate fuel oil, the changes to the survey forms should have had no direct effect on the residual fuel oil stocks series. However, the residual series exhibits an apparent downturn in the neighborhood of the time when the forms were changed. Both the CTSS and CUSUMs techniques identified several changes in the series, probably caused by the nonstationarity. However, the change in the series after January 1981 was found to be a statistically significant change by the CTSS method, whereas the CUSUMs method did not clearly identify this as a significant change. Likewise, when applied naively to the raw data series, the Box-Tiao procedure did not find the January 1981 change to be statistically significant. This lack of statistical significance is probably because of the large variability in the data, since the downturn in the data is quite marked. However, when an attempt is made to reduce the variability by incorporating the number of companies reporting into the analysis, the apparent downturn disappeared. Thus the change in the residual series is related to the decrease in the number of companies after January 1981, which in turn may be related to the general economic climate during that period which caused many users to switch from residual fuel oil to other fuel sources.

For motor gasoline, the different methods produced different results. Neither CUSUMs not CTSS identified the difference between the pre and post intervention in January 1981 as even slightly statistically significant, although both methods identified several significant changes in the data series when compared year to year. It is possible that these changes are due to the inherent nonstationarity in the series. The Box-Tiao procedure concluded that the intervention was significant at between the 0.15 and 0.10 levels, not a very strong change in the series. It also indicated that the effect of the intervention was a change in the first-order, moving-average parameter. But, as expected, when the pre-intervention data were adjusted using the post-intervention model, the adjusted values were not significantly different from the original data even when based on the tighter confidence intervals derived in Section 5. Incorporating the number of companies reporting in the analysis yielded no useful results.

The crude oil stocks series presented different sorts of problems. Many different interpretations of the time-series plot and the effect of the January 1981 intervention were possible. The series exhibited large variability and the Box-Tiao procedure did not identify the intervention as significant. CTSS and CUSUMs identified several significant changes and CTSS found the January 1981 intervention to be very significant. However, these two procedures were performed on the series including tank farm and pipeline stocks, whereas the Box-Tiao analysis was performed on only the refinery stocks series. The large trend present in the data may be causing CTSS and CUSUMs to identify changes in the mean caused by nonstationarity as interventions. Moreover, the effect of seasonally adjusting data, such as the crude stocks data, that do not seem to exhibit seasonality is not clear. In summary, no simple explanation of the behavior of the crude stocks data appears possible. Clearly, more work, especially the consideration of other explanatory variables (number of companies reporting was not useful), is necessary to understand crude oil stocks data.

## 7.3 RECOMMENDATIONS

The methods discussed in this report represent several valuable techniques for identifying an intervention and for assessing its effect. Further work is possible, not only on these techniques but also on applying other techniques. For instance, if it were desired to automate the modeling phase of the Box-Tiao analysis somewhat, an application of the Akaike Information Criteria to certain classes of models might be appropriate along with different methods of modeling and fitting.

The expression for the error in the adjusted values was derived through the use of several simplifying assumptions. The importance of these assumptions needs to be better understood and the general question of the variability of the adjusted values needs to be studied in more detail.

Finally, for CUSUMs, more work is necessary to evaluate its behavior when the data are not independent, especially when this dependency can be modeled by a time-series model. The effect of the particular scoring function on the behavior of the CTSS procedure would appear to be an interesting and important area.

# 8.0 REFERENCES

Akaike, H. 1976. "Canonical Correlation Analysis of Time Series and the Use of an Information Criterion". In R. Mehra and D. G. Lainiotis (eds), Advances and Case Studies in System Identification, Academic Press, New York, New York.

Barnard, H. 1959. "Control Charts and Stochastic Processes." J. Royal Stat. Soc. 21(B):239-270.

Bell, W.R. and S.C. Hillmer. 1983. "Modeling Time Series with Calendar Variation." J. Amer. Statis. Assoc. 78:526-534.

Bennett, B.M. 1964. "A Nonparametric Test for Randomness in a Sequence of Multinomial Trials." Biometrics 20:182-190.

Billingsley, P. 1961. "Statistical Methods in Markov Chains." Annals of Math. Stat. 32:12-40.

Box, G.E.P. and G.M. Jenkins. 1970. Time Series Analysis. Holden-Day, San Francisco, California.

Box, G.E.P., and G.C. Tiao. 1975. "Intervention Analysis with Applications to Economic and Environmental Problems". J. Amer. Statis. Assoc. 70(349):70-79.

Box, G.E.P., and G.C. Tiao. 1976. "Comparison of Forecast and Actuality." Appl. Statistics. 25(3):195-200.

Bureau of the Census, U.S. Department of Commerce. 1967. "The X-11 Variance of the Census Method II Seasonal Adjustment Program." Technical Paper 15, J. Shishkin et al., Washington, D.C.

Cobb, D.D. 1981. "Sequential Tests for Near Real-Time Accounting." Inst. Nucl. Mater. Manag. Proceedings X:13-15.

DeBruyn, D. 1961. Cumulative Sum Tests. Hafner, New York, New York.

DOE/EIA-0108. 1977, 1978, 1979, 1980. Energy Data Reports.

DOE/EIA-0109. 1982. Petroleum Supply Monthly.

Energy Information Administration. 1983. Weekly Petroleum Status Report. DOE/EIA-0208 (83/45), Washington, D.C.

Gardenier, T.K. 1979. "Profiling for Efficacy in Long-Range Clinical Trials." Advanced Medical Systems: An Assessment of the Contributions. E. Hinman, ed., Yearbook Medical Publications, Chicago, Illinois.

Hibbs, D.A., Jr. 1977. "On Analyzing Effects of Policy Interventions: Box-Jenkins and box-Tiao Versus Structural Equation Models". In D. R. Heise (ed.), Sociological Methodology (Chapter 4, pp. 137-179.), Jossey-Bass Publishers, San Francisco, California.

Johnson, N.L. 1961. "A Simple Theoretical Approach to Cumulative Sum Control Charts." J. Amer. Statis. Assoc. 56:835-840.

Johnson, N.L., and F.C. Leone. 1962. "Cumulative Sum Control Charts: Mathematical Principles Applied to Their Construction and Use." Indust. Quality Control, pp. 15-21.

Kumbaraci, T. 1974. "Transitions in Trinomial Data: A Method for Medical Judgment." Biometrics 30:381.

Markin, J.T., and J.P. Shipley. 1982. "Timely Detection of Material Loss." Inst. Nucl. Mater. Manag. Proceedings XI:60-63.

ORNL-7660/2. 1980. Validation of the Joint Petroleum Reporting System, Volumes I and II. (Includes report by Transportation and Economic Research Associates, Inc., memoranda by EIA staff and recommended changes.)

Page, E.S. 1961. "Cumulative Sum Charts." Technometrics 3:1-9.

SAS/ETS User's Guide: Econometric and Time-Series Library. 1980. SAS Institute, Cary, North Carolina.

Sobel, M., V.R.R. Uppuluri and T. Kumbaraci. 1973. "Successive Evaluation of Improvement Pattern in the Patient Records." Unpublished manuscript.

Tiao, G.C., G.E.P. Box and W.J. Hamming. 1975. "Analysis of Los Angeles Photochemical Smog Data: A Statistical Overview." J. Air Pollution Assoc. 25(3):260-258.

Turk, T.R. 1978. "Testing the Effect of an Intervention in Sequential Ecological Data." Biometrics 34:128-129.

Wald, A. 1947. Sequential Analysis. Wiley & Sons, New York, New York.

Whittle, P. 1955. "Some Distribution and Moment Formulae for the Markov Chain." J. Royal Stat. Soc. 17(B):235-242.

Wincek, M.A., K.B. Stewart and G.F. Piepel. 1979. "Statistical Methods for Evaluating Sequential Material Base Data." NUREG/CR-0683, Washington, D.C.

Woods, A.J., and D.J. Pike. 1981. "Compromise, Credibility and Clarity in the Analysis of Inventory Data." Inst. Nucl. Mater. Manag. Proceedings 58-61.

Zellner, A. and F. Palm. 1974. "Time Series Analysis and Simultaneous Equation Econometrics Models." J. Econometrics 2:17-54.

## APPENDIX A

## A GENERAL METHOD FOR INTERVENTION IDENTIFICATION AND ADJUSTMENT

# APPENDIX A

## A GENERAL METHOD FOR INTERVENTION IDENTIFICATION AND ADJUSTMENT

### A.1 INTRODUCTION

The purpose of this section is to describe a general method that, starting with the original time series data,

1. Identifies an intervention,
2. Quantifies the intervention in terms of the time series' parameters,
3. Produces adjusted (for the post-intervention model) values for the pre-intervention data points, and
4. Estimates the error in the adjusted values.

The use of this method is illustrated with a worked example. Where no generally available software exists for the required operation, FORTRAN programs are provided and described in this and the following appendix.

The method roughly follows the steps named above with some expansion and can be outlined as:

1. Model the pre-intervention time series,
2. Test to determine if the intervention is significant,
3. Determine the nature of the intervention,
4. Estimate the magnitudes of the parameter changes,
5. Adjust the pre-intervention data to correspond to the post-intervention situation, and
6. Produce estimates of the error in the adjusted values.

Each of these steps will be more fully described in subsequent sections.

## A.2 PRE-INTERVENTION MODELING

The first step in this method can be stated simply, but is probably the most difficult both to describe or to perform correctly. What is needed is an adequate time domain (ARIMA) model of the pre-intervention time series. There are several statistical analysis packages that can be used for time series modeling, the most notable being SAS-ETS since it is available on the EIA computer. Likewise there are many books that attempt to explain time series modeling at all levels of abstraction.

Time series modeling is an iterative process of identifying possible models, estimating parameters in the entertained model, diagnostic checking of the residuals and modifying the tentative model. Though there have been several attempts to automate model selection, these methods are rather limited in the types of series they can handle effectively and thus can not be seriously recommended for routine use. There is unfortunately no substitute for experience in the area.

## A.3 TESTING FOR AN INTERVENTION

There are at least two ways in which the examination of a time series for interventions can be interpreted. First, when considering the whole history of the series one might wish to determine if any interventions have occurred. Since all time points must be considered this is in general a quite difficult problem. In this report, however, it is assumed that the location of an intervention is known and the question being asked is whether or not the intervention has significantly affected the structure of the time series.

If enough data is available the most natural procedure is to independently model both the pre- and post-intervention series and to compare the results. The asymptotically normal distributions of the parameter vectors lead immediately to tests for significant differences between the two models.

Though the above situation is obviously the most desirable, it often happens that there is not enough post-intervention data to permit effective modeling with traditional techniques. In these cases the Q-statistic defined in Section 4.1 can be used to test for the significance of the intervention. To use this technique one must previously have obtained an adequate time domain model of the pre-intervention series. From the results of this modeling it will be assumed that the parameter estimates, the residual mean square, and the residuals from the model are available for input into the intervention program TSIAA described in Appendix B. TSIAA in the intervention analysis mode requires the input of the total number of time points (both before and after the intervention), the number of the last pre-intervention time point, the model parameters, the original series, $z_t$, and the residuals, $a_t$. TSIAA will accept models with an arbitrary order of nonseasonal differencing and with a single seasonal difference also of an arbitrary order. The time series model can be anything up to a second order seasonal ARMA model, i.e. two autoregressive (AR) and/or two seasonal autoregressive (SAR) and/or two moving average (MA) and/or two seasonal moving average (SMA) parameters.

Sample output from TSIAA in the intervention analysis mode is shown in Table A.1. TSIAA uses the pre-intervention model, the data and the residuals to compute $a_t$ (one-step ahead forecast errors) for the post-intervention period. The Q-statistic is then computed from the squares of these $a_t$ divided by the mean square of the pre-intervention series. As shown TSIAA provides both the Q-statistic and its p-level, that is, the probability that this value of Q could have been obtained by chance alone. This p-level is computed from a Chi-square statistic with degrees of freedom equal to the number of time points after the intervention.

## TABLE A.1. Example Output of TSIAA in Intervention Analysis Mode

```
PROGRAM TSIAA -

TITLE OF RUN:  EXAMPLE INTERVENTION ANALYSIS

 TYPE OF RUN:  INTERVENTION ANALYSIS

MODEL:
  ((1-B)**D)*((1-B**S)**SD)*(1-AR(1)B-AR(2)B**2)*(1-SAR(1)B**S-SAR(2)B**2S)*Z(T) =
                                (1-MA(1)B-MA(2)B**2)*(1-SMA(1)B**S-SMA(2)B**2S)*A(T)

WHERE

  S =  12  D =   1  SD =   1

  PRE-INTERVENTION PARAMETERS:
           AR(1) =   0.00000E+00   AR(2) =   0.00000E+00
          SAR(1) =   0.00000E+00  SAR(2) =   0.00000E+00
           MA(1) =  -0.23700E+00   MA(2) =   0.00000E+00
          SMA(1) =   0.75500E+00  SMA(2) =   0.00000E+00

NUMBER OF TIME POINT VALUES =  84
  INTERVENTION OCCURRENCE AFTER POINT =  60

RESIDUAL MEAN SQUARE =  0.479011E+08

***    ORIGINAL DATA    ***

   T      Z(T)             A(T)

   1  0.242409E+06       ****
   2  0.250433E+06       ****
   3  0.240293E+06       ****
   4  0.226551E+06       ****
   5  0.227168E+06       ****
   6  0.227367E+06       ****
   7  0.229034E+06       ****
   8  0.232459E+06       ****
   9  0.231745E+06       ****
  10  0.228388E+06       ****
  11  0.229884E+06       ****
  12  0.233825E+06       ****
  13  0.255625E+06       ****
  14  0.258052E+06  -0.693800E+03
  15  0.264919E+06   0.159964E+05
  16  0.261065E+06   0.116150E+04
  17  0.265197E+06   0.699590E+04
  18  0.259258E+06  -0.615340E+04
  19  0.260849E+06  -0.970300E+03
  20  0.259814E+06   0.145450E+04

                  .
                  .
                  .

  60  0.263707E+06  -0.906740E+04
  61  0.279560E+06  -0.276851E+04
```

## TABLE A.1. Example Output of TSIAA in Intervention Analysis Mode (Continued)

```
 62  0.287492E+06   0.542551E+04
 63  0.288462E+06   0.393434E+04
 64  0.275759E+06  -0.529842E+04
 65  0.261541E+06  -0.659175E+04
 66  0.244246E+06  -0.138422E+05
 67  0.230429E+06  -0.123820E+05
 68  0.236156E+06   0.125547E+05
 69  0.240372E+06   0.121438E+04
 70  0.239285E+06   0.585068E+04
                  .
                  .
                  .

 80  0.227963E+06  -0.903186E+02
 81  0.235537E+06   0.654273E+04
 82  0.237117E+06   0.575093E+04
 83  0.232663E+06  -0.125547E+05
 84  0.237073E+06  -0.548911E+03

NOTE:  AFTER T =  60, A(T) IS COMPUTED FROM THE PRE-INTERVENTION MODEL

-------------------------------------------------------------------------
    Q STATISTIC =  0.305891E+02  WITH  24 DEGREES OF FREEDOM
       P-VALUE FOR A CHI-SQUARE WITH  24 DEGREES OF FREEDOM =  0.1660
-------------------------------------------------------------------------


                                                  PREDICTED ERRORS DUE TO A CHANGE IN:
                 ---------------------------------------------------------------------------------------------------------------------
 T      A(T)         MEAN         AR(1)         AR(2)        SAR(1)        SAR(2)         MA(1)         MA(2)        SMA(1)        SMA(2)

 61 -0.277E+04  0.10000E+01 -0.90674E+04  0.83563E+04  0.45855E+04 -0.23123E+04  0.90674E+04 -0.83563E+04 -0.45855E+04  0.23123E+04
 62  0.543E+04 -0.23700E+00 -0.27685E+04 -0.90674E+04  0.11312E+05 -0.47433E+04  0.61953E+03  0.11048E+05 -0.11312E+05  0.47433E+04
 63  0.393E+04  0.56169E-01  0.54255E+04 -0.27685E+04  0.13676E+05 -0.59042E+04 -0.55723E+04  0.15017E+03 -0.13676E+05  0.59042E+04
 64 -0.530E+04 -0.13312E-01  0.39343E+04  0.54255E+04 -0.73457E+04  0.64991E+04 -0.26137E+04 -0.54611E+04  0.73457E+04 -0.64991E+04
 65 -0.659E+04  0.31550E-02 -0.52984E+04  0.39343E+04 -0.22001E+04 -0.39548E+04  0.59179E+04 -0.26401E+04  0.22001E+04  0.39548E+04
 66 -0.138E+05 -0.74772E-03 -0.65918E+04 -0.52984E+04  0.69566E+04  0.87677E+04  0.51892E+04  0.59241E+04 -0.69566E+04 -0.87677E+04
 67 -0.124E+05  0.17721E-03 -0.13842E+05 -0.65918E+04 -0.10144E+05  0.66705E+04  0.12612E+05  0.51877E+04  0.10144E+05 -0.66705E+04
 68  0.126E+05 -0.41999E-04 -0.12382E+05 -0.13842E+05  0.50055E+04 -0.73008E+04  0.93928E+04  0.12613E+05 -0.50055E+04  0.73008E+04
 69  0.121E+04  0.99538E-05  0.12555E+05 -0.12382E+05 -0.10501E+04 -0.21059E+04 -0.14781E+05  0.93928E+04  0.10501E+04  0.21059E+04
 70  0.585E+04 -0.23590E-05  0.12144E+04  0.12555E+05 -0.58273E+04 -0.69081E+04  0.22887E+04 -0.14781E+05  0.58273E+04  0.69081E+04
 71  0.568E+04  0.55909E-06  0.58507E+04  0.12144E+04  0.83563E+04 -0.15040E+03 -0.63931E+04  0.22887E+04 -0.83563E+04  0.15040E+03
 72 -0.682E+04 -0.13250E-06  0.56827E+04  0.58507E+04 -0.90674E+04  0.74187E+04 -0.41675E+04 -0.63931E+04  0.90674E+04 -0.74187E+04
 73 -0.891E+04 -0.24500E+00 -0.68215E+04  0.56827E+04 -0.27685E+04  0.45855E+04  0.78092E+04 -0.41675E+04 -0.69354E+03 -0.28397E+04
 74 -0.643E+04  0.58065E-01 -0.89090E+04 -0.68215E+04  0.54255E+04  0.11312E+05  0.70582E+04  0.78092E+04 -0.13966E+05 -0.77312E+04
 75 -0.657E+04 -0.13761E-01 -0.64257E+04 -0.89090E+04  0.39343E+04  0.13676E+05  0.47529E+04  0.70582E+04 -0.14259E+05 -0.92178E+04
 76 -0.137E+05  0.32615E-02 -0.65729E+04 -0.64257E+04 -0.52984E+04 -0.73457E+04  0.54464E+04  0.47529E+04  0.10844E+05  0.24389E+04
 77  0.482E+04 -0.77296E-03 -0.13729E+05 -0.65729E+04 -0.65918E+04 -0.22001E+04  0.12438E+05  0.54464E+04  0.82528E+04  0.51860E+04
 78  0.844E+04  0.18319E-03  0.48238E+04 -0.13729E+05 -0.13842E+05  0.69566E+04 -0.77716E+04  0.12438E+05  0.85900E+04 -0.13576E+05
 79  0.658E+04 -0.43417E-04  0.84444E+04  0.48238E+04 -0.12382E+05 -0.10144E+05 -0.66025E+04 -0.77716E+04  0.20040E+05  0.51073E+04
 80 -0.903E+02  0.10290E-04  0.65844E+04  0.84444E+04  0.12555E+05  0.50055E+04 -0.50196E+04 -0.66025E+04 -0.16334E+05  0.50660E+03
 81  0.654E+04 -0.24387E-05 -0.90319E+02  0.65844E+04  0.12144E+04 -0.10501E+04  0.12800E+04 -0.50196E+04 -0.42156E+03  0.26401E+04
 82  0.575E+04  0.57796E-06  0.65427E+04 -0.90319E+02  0.58507E+04 -0.58273E+04 -0.68461E+04  0.12800E+04 -0.14511E+04  0.11043E+05
 83 -0.126E+05 -0.13698E-06  0.57509E+04  0.65427E+04  0.56827E+04  0.83563E+04 -0.41284E+04 -0.68461E+04 -0.11992E+05 -0.82427E+04
 84 -0.549E+03  0.32464E-07 -0.12555E+05  0.57509E+04 -0.68215E+04 -0.90674E+04  0.13533E+05 -0.41284E+04  0.13667E+05  0.34663E+04
```

## A.4 DETERMINING THE NATURE OF THE INTERVENTION

Assuming that the intervention has been determined to be significant, the next step is to attempt to identify which parameters or combinations of parameters have changed due to the intervention. The program TSIAA in the intervention analysis mode implements the method described in Section 4.1.

As shown in Table A.1, TSIAA provides, under the heading 'PREDICTED ERRORS DUE TO A CHANGE IN:', the information necessary to decide if any of the model's parameters have changed. To make this decision, the $a_t$'s (i.e. the one-step ahead forecast errors using the pre-intervention model and the data points after the intervention) and the predicted errors resulting from some hypothesized change in the model are examined for similarity in structure. One way to determine similarity is to plot the $a_t$'s and the hypothesized errors together and look for patterns. For example, see Figures 26 to 28. The correlation coefficient between the $a_t$'s and the different hypothesized residuals can also be examined. As described in Section 4.1, when similarities are seen between the $a_t$'s and one or more of the collections of predicted errors, linear regression can be used to estimate the change in the parameters. That is, the $a_t$'s are regressed, with no constant, on the sets of predicted errors and the estimated coefficients provide estimates of the difference between the post- and pre-intervention parameters. Thus if $\beta$ is a pre-intervention parameter, $\beta'$ is the corresponding post-intervention parameter and b is the estimated regression coefficient then

$$\beta' = \beta + b$$

As described in Section 5.3, for the current example, similar patterns were noticed between the $a_t$'s and the predicted errors due to a change in the MA(1) parameter. No other similarities were found. Thus the $a_t$'s were

## TABLE A.2. Example Output of TSIAA in Intervention Adjustment Mode

```
PROGRAM TSIAA -

TITLE OF RUN:  EXAMPLE INTERVENTION ADJUSTMENT WITHOUT MSE

 TYPE OF RUN:  INTERVENTION ADJUSTMENT WITHOUT MEAN SQUARE ERROR COMPUTATION

MODEL:
  ((1-B)**D)*((1-B**S)**SD)*(1-AR(1)B-AR(2)B**2)*(1-SAR(1)B**S-SAR(2)B**2S)*Z(T) =
                             (1-MA(1)B-MA(2)B**2)*(1-SMA(1)B**S-SMA(2)B**2S)*A(T)

WHERE

  S =  12  D =   1  SD =   1

  PRE-INTERVENTION PARAMETERS:
           AR(1) =   0.00000E+00   AR(2) =   0.00000E+00
          SAR(1) =   0.00000E+00  SAR(2) =   0.00000E+00
           MA(1) =  -0.23700E+00   MA(2) =   0.00000E+00
          SMA(1) =   0.75500E+00  SMA(2) =   0.00000E+00

  POST-INTERVENTION PARAMETERS:
           AR(1) =   0.00000E+00   AR(2) =   0.00000E+00
          SAR(1) =   0.00000E+00  SAR(2) =   0.00000E+00
           MA(1) =  -0.55500E+00   MA(2) =   0.00000E+00
          SMA(1) =   0.75500E+00  SMA(2) =   0.00000E+00

NUMBER OF TIME POINT VALUES =  84
  INTERVENTION OCCURRENCE AFTER POINT =  60

RESIDUAL MEAN SQUARE =  0.479011E+08

***    ORIGINAL DATA     ***

   T      Z(T)           A(T)

   1  0.242409E+06       ****
   2  0.250433E+06       ****
   3  0.240293E+06       ****
   4  0.226551E+06       ****
   5  0.227168E+06       ****
   6  0.227367E+06       ****
   7  0.229034E+06       ****
   8  0.232459E+06       ****
   9  0.231745E+06       ****
  10  0.228388E+06       ****
  11  0.229884E+06       ****
  12  0.233825E+06       ****
  13  0.255625E+06       ****
  14  0.258052E+06  -0.693800E+03
  15  0.264919E+06   0.159964E+05
  16  0.261065E+06   0.116150E+04
  17  0.265197E+06   0.699590E+04
  18  0.259258E+06  -0.615340E+04
```

TABLE A.2. Example Output of TSIAA in Intervention Adjustment Mode (Continued)

```
   19  0.260849E+06  -0.970300E+03
   20  0.259814E+06   0.145450E+04

                  .
                  .
                  .

   80  0.227963E+06  -0.405885E+04
   81  0.235537E+06   0.968132E+04
   82  0.237117E+06   0.173808E+04
   83  0.232663E+06  -0.132126E+05
   84  0.237073E+06   0.532337E+04

 NOTE:  AFTER T =  60, A(T) IS COMPUTED FROM THE POST-INTERVENTION MODEL


        AUTOCOVARIANCES OF THE                 PSI WEIGHTS FOR THE
 DIFFERENCED PRE-INTERVENTION SERIES       POST-INTERVENTION SERIES

        LAG    AUTOCOVARIANCES               INDEX     PSI WEIGHTS
        ---    ---------------               -----    ------------

          0      0.794301E+08                  0      0.100000E+01
          1      0.178238E+08                  1      0.155500E+01
          2      0.000000E+00                  2      0.155500E+01
          3      0.000000E+00                  3      0.155500E+01
          4      0.000000E+00                  4      0.155500E+01
          5      0.000000E+00                  5      0.155500E+01
          6      0.000000E+00                  6      0.155500E+01
          7      0.000000E+00                  7      0.155500E+01
          8      0.000000E+00                  8      0.155500E+01
          9      0.000000E+00                  9      0.155500E+01
         10      0.000000E+00                 10      0.155500E+01
         11     -0.857118E+07                 11      0.155500E+01
         12     -0.381967E+08                 12      0.180000E+01
         13     -0.857118E+07                 13      0.193598E+01
         14      0.000000E+00                 14      0.193598E+01
         15      0.000000E+00                 15      0.193598E+01
         16      0.000000E+00                 16      0.193598E+01
         17      0.000000E+00                 17      0.193598E+01
         18      0.000000E+00                 18      0.193598E+01
         19      0.000000E+00                 19      0.193598E+01
         20      0.000000E+00                 20      0.193598E+01

                                 .
                                 .
                                 .

         50      0.000000E+00                 50      0.307890E+01
         51      0.000000E+00                 51      0.307890E+01
         52      0.000000E+00                 52      0.307890E+01
         53      0.000000E+00                 53      0.307890E+01
         54      0.000000E+00                 54      0.307890E+01
         55      0.000000E+00                 55      0.307890E+01
         56      0.000000E+00                 56      0.307890E+01
         57      0.000000E+00                 57      0.307890E+01
         58      0.000000E+00                 58      0.307890E+01
         59      0.000000E+00                 59      0.307890E+01
```

TABLE A.2. Example Output of TSIAA in Intervention Adjustment Mode (Continued)

```
  T    ADJUSTED Z(T)
 ---   -------------

   1    0.285362E+06
   2    0.286866E+06
   3    0.274976E+06
   4    0.261434E+06
   5    0.260256E+06
   6    0.260327E+06
   7    0.262583E+06
   8    0.263473E+06
   9    0.261867E+06
  10    0.256286E+06
  11    0.260771E+06
  12    0.266801E+06
  13    0.284963E+06
  14    0.285774E+06
  15    0.289494E+06
  16    0.285993E+06
  17    0.292455E+06
  18    0.290255E+06
  19    0.288126E+06
  20    0.289932E+06

                .
                .
                .

  40    0.235771E+06
  41    0.232439E+06
  42    0.234634E+06
  43    0.245059E+06
  44    0.241037E+06
  45    0.234789E+06
  46    0.223288E+06
  47    0.225429E+06
  48    0.239322E+06
  49    0.264882E+06
  50    0.277406E+06
  51    0.285505E+06
  52    0.274069E+06
  53    0.264723E+06
  54    0.267691E+06
  55    0.263123E+06
  56    0.261194E+06
  57    0.261323E+06
  58    0.249510E+06
  59    0.259782E+06
  60    0.263707E+06
```

regressed on the predicted errors corresponding to a change in the MA(1) parameter resulting in an estimated change of -.318. The post intervention MA(1) parameter was therefore estimated as -.555.

## A.5 ADJUSTING THE PRE-INTERVENTION DATA

Once the nature of the intervention has been identified and the values of the post-intervention parameters estimated, the pre-intervention data may be adjusted to produce a consistent data series. TSIAA provides two modes in which to accomplish this. The first such mode permits only the adjustment of pre-intervention data, while the second also allows an assessment of the mean square error of adjustment. The first mode will be discussed in this section.

As detailed in Appendix B, TSIAA in this mode requires the input of the values of the parameters for the post-intervention model as well as the other inputs described above. Once these have been provided, TSIAA computes the estimated $a_t$'s in the post-intervention period using the post-intervention model. From these and the pre-intervention $a_t$'s TSIAA then computes the adjusted post-intervention data using the recursion formula (8) from Section 5.5 and outputs these values as shown in Table A.2.

## A.6 ESTIMATING THE ERRORS OF ADJUSTMENT

To obtain the approximate errors in the adjusted values, TSIAA implements the method described in Section 5.5. The program requires the additional input of the covariance matrix of the pi-weights as defined in Section 5.5. Since the pi-weights are rational functions of the parameters in the pre-intervention model, the variances of the parameter estimates can be propagated to estimate this covariance matrix. One such method of propagation is illustrated in Section 5.5. The degree to which the several Taylor series expansions are expanded will depend on the level of accuracy desired in the estimated errors, the complexity of the model and the distribution of the random component of the series.

From the form of the pre-intervention model, TSIAA computes the covariance structure of the differenced pre-intervention series. From the form of the post-intervention model TSIAA calculates the psi-weights as defined in Section 5.5. Formulas 27 and 21 from Section 5.5 are then used to calculate the approximate adjustment variance.

TSIAA in this mode, as well as in the previous mode outputs the estimated covariances of the differenced pre-intervention series, the psi-weights, and the adjusted values for the pre-intervention period as shown in Table A.2. Further, the estimated mean square error of adjustment and approximate 95% confidence bounds for the adjusted value are also output as shown in Table A.3.

## TABLE A.3. Additional Output of TSIAA When MSE of Adjustment is Calculated

```
                                              APPROX. 95% C. I. FOR ADJUSTED Z(T)
 T    ADJUSTED Z(T)     MSE Z(T)                LOWER LIMIT     UPPER LIMIT
---   -------------   --------------           -------------   ------------

  1    0.285362E+06    0.207983E+10             0.195976E+06    0.374748E+06
  2    0.286866E+06    0.201972E+10             0.198781E+06    0.374951E+06
  3    0.274976E+06    0.195961E+10             0.188211E+06    0.361740E+06
  4    0.261434E+06    0.189950E+10             0.176011E+06    0.346858E+06
  5    0.260256E+06    0.183939E+10             0.176195E+06    0.344317E+06
  6    0.260327E+06    0.177928E+10             0.177651E+06    0.343002E+06
  7    0.262583E+06    0.171917E+10             0.181316E+06    0.343850E+06
  8    0.263473E+06    0.165906E+10             0.183639E+06    0.343307E+06
  9    0.261867E+06    0.159898E+10             0.183492E+06    0.340242E+06
 10    0.256286E+06    0.153898E+10             0.179396E+06    0.333177E+06
 11    0.260771E+06    0.147908E+10             0.185392E+06    0.336151E+06
 12    0.266801E+06    0.138049E+10             0.193977E+06    0.339625E+06
 13    0.284963E+06    0.126523E+10             0.215246E+06    0.354680E+06
 14    0.285774E+06    0.122682E+10             0.217123E+06    0.354425E+06
 15    0.289494E+06    0.118841E+10             0.221927E+06    0.357062E+06
 16    0.285993E+06    0.115001E+10             0.219526E+06    0.352460E+06
 17    0.292455E+06    0.111160E+10             0.227107E+06    0.357803E+06
 18    0.290255E+06    0.107319E+10             0.226046E+06    0.354464E+06
 19    0.288126E+06    0.103478E+10             0.225076E+06    0.351175E+06
 20    0.289932E+06    0.996378E+09             0.228063E+06    0.351800E+06

                                          .
                                          .
                                          .

 40    0.235771E+06    0.334293E+09             0.199935E+06    0.271607E+06
 41    0.232439E+06    0.321824E+09             0.197278E+06    0.267600E+06
 42    0.234634E+06    0.309355E+09             0.200160E+06    0.269107E+06
 43    0.245059E+06    0.296887E+09             0.211287E+06    0.278830E+06
 44    0.241037E+06    0.284421E+09             0.207982E+06    0.274092E+06
 45    0.234789E+06    0.271971E+09             0.202465E+06    0.267112E+06
 46    0.223288E+06    0.259581E+09             0.191709E+06    0.254866E+06
 47    0.225429E+06    0.247233E+09             0.194611E+06    0.256248E+06
 48    0.239322E+06    0.211124E+09             0.210843E+06    0.267801E+06
 49    0.264882E+06    0.164586E+09             0.239737E+06    0.290027E+06
 50    0.277406E+06    0.159840E+09             0.252626E+06    0.302186E+06
 51    0.285505E+06    0.155093E+09             0.261096E+06    0.309914E+06
 52    0.274069E+06    0.150347E+09             0.250036E+06    0.298102E+06
 53    0.264723E+06    0.145601E+09             0.241073E+06    0.288373E+06
 54    0.267691E+06    0.140855E+09             0.244429E+06    0.290953E+06
 55    0.263123E+06    0.136109E+09             0.240257E+06    0.285989E+06
 56    0.261194E+06    0.131366E+09             0.238729E+06    0.283659E+06
 57    0.261323E+06    0.126641E+09             0.239266E+06    0.283380E+06
 58    0.249510E+06    0.121993E+09             0.227862E+06    0.271158E+06
 59    0.259782E+06    0.117589E+09             0.238528E+06    0.281036E+06
 60    0.263707E+06    0.479011E+08             0.250142E+06    0.277272E+06
```

# APPENDIX B

# TIME-SERIES INTERVENTION ANALYSIS AND ADJUSTEMENT COMPUTER PROGRAM

# APPENDIX B

# TIME SERIES INTERVENTION ANALYSIS AND ADJUSTMENT COMPUTER PROGRAM

## B.1 INTRODUCTION

The Time Series Intervention Analysis and Adjustment (TSIAA) program is a FORTRAN computer code developed to assist the user in time series intervention analysis. TSIAA can handle up to a second order seasonal ARMA time series model with arbitrary order of seasonal and nonseasonal differencing. Program TSIAA is designed to be used with a computerized time series modeling package (e.g. the SAS ETS package) and is executed in two steps. In the intervention analysis step, the user supplies as program inputs the original time series and the pre-intervention model parameters and residuals. Program TSIAA computes and outputs the Q statistic and the predicted errors due to a change in level and model parameters. In the intervention adjustment step, the user additionally supplies the post-intervention model parameters and optionally supplies the covariance matrix of the estimated pi weights of the pre-intervention model. Program TSIAA then computes and outputs the adjusted time series. If the optional covariance matrix of the estimated pi weights was supplied, then TSIAA also calculates and outputs the mean square errors of the adjusted time series values and the approximate 95% confidence limits about each adjusted value.

The remainder of this appendix is divided into five sections. The first section describes the data input to program TSIAA. The outputs from TSIAA are presented in the second section. The third section describes how to use TSIAA on the EIA computer. The program's structure and its modules are described in the fourth section and the last section provides a listing of the program's FORTRAN source code.

## B.2 PROGRAM INPUT

Input to program TSIAA consists of a data file. The input data file has a specific structure which is defined in Table B.1. Table B.1 lists the card image sequence within the input data file and the field(s) of each card image in terms of the variable name that program TSIAA assigns to a field, a description of the field, and the FORTRAN format specification used by program TSIAA to read a field's value from the card image.

The amount of information that must be supplied in the input data file will depend on the purpose for which program TSIAA is to be used. If an intervention analysis is desired, then, in reference to Table B.1, only cards 1 through 12 need be supplied. If an intervention adjustment analysis is desired, then, in addition to cards 1 through 12, cards 13 through 16 must also be supplied. If, for an intervention adjustment analysis, the mean square errors and 95% confidence limits for the adjusted time series values are desired, then cards 17 and 18 must also be supplied.

The value of the field IPT on card 2 indicates to program TSIAA which of the three types of program usage (i.e., intervention analysis, intervention adjustment without mean square error computations, or intervention adjustment with mean square error computations) is in effect and directs how TSIAA reads the data file. Thus, when proceeding from an intervention analysis run to a intervention adjustment run for a given time series, modifications to the data file would involve changing the value of IPT from 1 to 2 or 3, as well as appending the additionally required card images to the existing data file.

For an example time series, Table B.2, Table B.3, and Table B.4 illustrate the forms of the input data file corresponding to the three TSIAA run types of intervention analysis, intervention adjustment without mean square error computations, and intervention adjustment with mean square error

## TABLE B.1. Input Data File Structure

| Card | Variable | Description | Format |
|---|---|---|---|
| 1 | TITLE | The title of the run. | 20A4 |
| 2 | IPT | Type of run. IPT = 1 indicates an intervention analysis run. IPT = 2 indicates an intervention adjustment run WITHOUT mean square error computations. IPT = 3 indicates an intervention adjustment run WITH mean square error computations. | I5 |
|  | IPRT | Print flag for the VP matrix (defined on card 18). If IPRT = 1, then the VP matrix is output on the report. (IPRT is used only when IPT = 3). | I5 |
| 3 | N | Number of input time points. | I5 |
|  | IV | Time point after which the intervention occurred. | I5 |
| 4 | IS | Order of seasonality. | I5 |
|  | ID | Degree of non-seasonal differencing. | I5 |
|  | ICD | Degree of seasonal differencing. | I5 |
| 5 | PHI(1) | AR(1) pre-intervention model parameter. | F10.0 |
|  | PHI(2) | AR(2) pre-intervention model parameter. | F10.0 |
| 6 | THETA(1) | MA(1) pre-intervention model parameter. | F10.0 |
|  | THETA(2) | MA(2) pre-intervention model parameter. | F10.0 |
| 7 | CPHI(1) | SAR(1) pre-intervention model parameter. | F10.0 |
|  | CPHI(2) | SAR(2) pre-intervention model parameter. | F10.0 |
| 8 | CTHETA(1) | SMA(1) pre-intervention model parameter. | F10.0 |
|  | CTHETA(2) | SMA(2) pre-intervention model parameter. | F10.0 |
| 9 | FMT | FORTRAN format specification for the following Z array. For example: (8F10.2) | 20A4 |
| 10* | Z(I)<br>(*) | Original time series (where I = 1,...N). Card 10 is repeated as necessary to read all the Z(I). | FMT |

## TABLE B.1. Input Data File Structure (Continued)

| Card | Variable | Description | Format |
|---|---|---|---|
| 11 | FMT | FORTRAN format specification for the following A array. | 20A4 |
| 12* | A(I) | Pre-intervention residuals (where I = 1,...IV). | FMT |
| | (*) | Card 12 is repeated as necessary to read all the A(I). | |
| ******** | | Cards 13 - 16 are input only when IPT = 2 or 3 | ******** |
| 13 | PHIN(1) | AR(1) post-intervention model parameter. | F10.0 |
| | PHIN(2) | AR(2) post-intervention model parameter. | F10.0 |
| 14 | THETN(1) | MA(1) post-intervention model parameter. | F10.0 |
| | THETN(2) | MA(2) post-intervention model parameter. | F10.0 |
| 15 | CPHIN(1) | SAR(1) post-intervention model parameter. | F10.0 |
| | CPHIN(2) | SAR(2) post-intervention model parameter. | F10.0 |
| 16 | CTHETN(1) | SMA(1) post-intervention model parameter. | F10.0 |
| | CTHETN(2) | SMA(2) post-intervention model parameter. | F10.0 |
| ******** | | Cards 17 and 18 are input only when IPT = 3 | ******** |
| 17 | FMT | FORTRAN format specification for the following VP array. | 20A4 |
| 18* | VP(I,J) | Covariance matrix of the estimated pi weights of the pre-intervention model (where I = 1,...IV and J = I,...IV). This matrix is read in upper triangular form a row at a time. | FMT |
| | (*) | Card 18 is repeated as necessary to read all the VP(I,J). | |

TABLE B.2. Example Input Data File - Intervention Analysis (IPT = 1)

```
EXAMPLE INTERVENTION ANALYSIS            --  TITLE
    1    0                               --  IPT (1=ANALYSIS), IPRT
   84   60                               --  N, IV
   12    1    1                          --  IS, ID, ICD
 0.         0.                           --  PHI(1), PHI(2)
-.237       0.                           --  THETA(1), THETA(2)
 0.         0.                           --  CPHI(1), CPHI(2)
.755        0.                           --  CTHETA(1), CTHETA(2)
(6X,F12.0)                               --  FMT
     1      242409.          0.          --  Z(I)
     2      250433.          0.
     3      240293.          0.
     4      226551.          0.
     5      227168.          0.
     6      227367.          0.
     7      229034.          0.
     8      232459.          0.
     9      231745.          0.
    10      228388.          0.
    11      229884.          0.
    12      233825.          0.
    13      255625.          0.
    14      258052.       -693.8
    15      264919.      15996.4

               .
               .
               .

    75      248928.     -14244.0
    76      224235.     -22350.5
    77      217512.      -1736.7
    78      221435.       9770.6
    79      228029.       9020.0
    80      227963.       6344.2
    81      235537.       6671.4
    82      237117.       9164.2
    83      232663.      -7449.0
    84      237073.      -7543.9
```

## TABLE 2. Example Input Data File - Intervention Analysis (IPT = 1) (Continued)

```
(18X,F12.0)                                            --  FMT
       1      242409.          0.                      --  A(I)
       2      250433.          0.
       3      240293.          0.
       4      226551.          0.
       5      227168.          0.
       6      227367.          0.
       7      229034.          0.
       8      232459.          0.
       9      231745.          0.
      10      228388.          0.
      11      229884.          0.
      12      233825.          0.
      13      255625.          0.
      14      258052.     -693.8
      15      264919.    15996.4

                 .
                 .
                 .

      50      277406.    11312.4
      51      285505.    13675.5
      52      274069.    -7345.7
      53      264723.    -2200.1
      54      267691.     6956.6
      55      263123.   -10143.5
      56      261194.     5005.5
      57      261323.    -1050.1
      58      249510.    -5827.3
      59      259782.     8356.3
      60      263707.    -9067.4
```

## TABLE B.3. Example Input Data File - Intervention Adjustment Without MSE (IPT = 2)

```
EXAMPLE INTERVENTION ADJUSTMENT WITHOUT MSE     --  TITLE
   2    0                                       --  IPT (2=ADJUST WITHOUT MSE), IPRT
  84   60                                       --  N, IV
  12    1    1                                  --  IS, ID, ICD
 0.         0.                                  --  PHI(1), PHI(2)
-.237       0.                                  --  THETA(1), THETA(2)
 0.         0.                                  --  CPHI(1), CPHI(2)
.755        0.                                  --  CTHETA(1), CTHETA(2)
(6X,F12.0)                                      --  FMT
     1      242409.          0.                 --  Z(I)
     2      250433.          0.
     3      240293.          0.
     4      226551.          0.
     5      227168.          0.
     6      227367.          0.
     7      229034.          0.
     8      232459.          0.
     9      231745.          0.
    10      228388.          0.
    11      229884.          0.
    12      233825.          0.
    13      255625.          0.
    14      258052.       -693.8
    15      264919.      15996.4

              .
              .
              .

    75      248928.     -14244.0
    76      224235.     -22350.5
    77      217512.      -1736.7
    78      221435.       9770.6
    79      228029.       9020.0
    80      227963.       6344.2
    81      235537.       6671.4
    82      237117.       9164.2
    83      232663.      -7449.0
    84      237073.      -7543.9
```

TABLE B.3. Example Input Data File - Intervention Adjustment Without MSE (IPT = 2) (Continued)

```
(18X,F12.0)                                                 --  FMT
     1       242409.          0.                            --  A(I)
     2       250433.          0.
     3       240293.          0.
     4       226551.          0.
     5       227168.          0.
     6       227367.          0.
     7       229034.          0.
     8       232459.          0.
     9       231745.          0.
    10       228388.          0.
    11       229884.          0.
    12       233825.          0.
    13       255625.          0.
    14       258052.      -693.8
    15       264919.     15996.4

                .
                .
                .

    50       277406.     11312.4
    51       285505.     13675.5
    52       274069.     -7345.7
    53       264723.     -2200.1
    54       267691.      6956.6
    55       263123.    -10143.5
    56       261194.      5005.5
    57       261323.     -1050.1
    58       249510.     -5827.3
    59       259782.      8356.3
    60       263707.     -9067.4
0.           0.                                             --  PHIN(1), PHIN(2)
-.555        0.                                             --  THETN(1), THETN(2)
0.           0.                                             --  CPHIN(1), CPHIN(2)
.755         0.                                             --  CTHETN(1), CTHETN(2)
```

TABLE B.4. Example Input Data File - Intervention Adjustment with MSE (IPT = 3)

```
EXAMPLE INTERVENTION ADJUSTMENT WITH MSE       --  TITLE
    3    0                                     --  IPT (3=ADJUST WITH MSE), IPRT
   84   60                                     --  N, IV
   12    1    1                                --  IS, ID, ICD
 0.        0.                                  --  PHI(1), PHI(2)
-.237      0.                                  --  THETA(1), THETA(2)
 0.        0.                                  --  CPHI(1), CPHI(2)
.755       0.                                  --  CTHETA(1), CTHETA(2)
(6X,F12.0)                                     --  FMT
     1     242409.          0.                 --  Z(I)
     2     250433.          0.
     3     240293.          0.
     4     226551.          0.
     5     227168.          0.
     6     227367.          0.
     7     229034.          0.
     8     232459.          0.
     9     231745.          0.
    10     228388.          0.
    11     229884.          0.
    12     233825.          0.
    13     255625.          0.
    14     258052.      -693.8
    15     264919.     15996.4

              .
              .
              .

    75     248928.    -14244.0
    76     224235.    -22350.5
    77     217512.     -1736.7
    78     221435.      9770.6
    79     228029.      9020.0
    80     227963.      6344.2
    81     235537.      6671.4
    82     237117.      9164.2
    83     232663.     -7449.0
    84     237073.     -7543.9
```

## TABLE B.4. Example Input Data File - Intervention Adjustment with MSE (IPT = 3) (Continued)

```
(18X,F12.0)                                                  -- FMT
      1       242409.          0.                            -- A(I)
      2       250433.          0.
      3       240293.          0.
      4       226551.          0.
      5       227168.          0.
      6       227367.          0.
      7       229034.          0.
      8       232459.          0.
      9       231745.          0.
     10       228388.          0.
     11       229884.          0.
     12       233825.          0.
     13       255625.          0.
     14       258052.      -693.8
     15       264919.     15996.4

                 .
                 .
                 .

     50       277406.     11312.4
     51       285505.     13675.5
     52       274069.     -7345.7
     53       264723.     -2200.1
     54       267691.      6956.6
     55       263123.    -10143.5
     56       261194.      5005.5
     57       261323.     -1050.1
     58       249510.     -5827.3
     59       259782.      8356.3
     60       263707.     -9067.4
0.            0.                                             -- PHIN(1), PHIN(2)
-.555         0.                                             -- THETN(1), THETN(2)
0.            0.                                             -- CPHIN(1), CPHIN(2)
.755          0.                                             -- CTHETN(1), CTHETN(2)
```

TABLE B.4. Example Input Data File - Intervention Adjustment with MSE (IPT = 3) (Continued)

```
(8F10.0)                                                                 --  FMT
 0.222E-01-0.105E-01 0.374E-02-0.118E-02 0.350E-03-0.996E-04 0.275E-04-0.746E-05  --  VP (FIRST ROW OF UPPER TRIANGULAR FORM)
 0.199E-05-0.524E-06 0.137E-06-0.700E-03 0.169E-01-0.798E-02 0.283E-02-0.895E-03
 0.265E-03-0.753E-04 0.208E-04-0.564E-05 0.150E-05-0.396E-06 0.103E-06-0.106E-02
 0.129E-01-0.606E-02 0.215E-02-0.677E-03 0.200E-03-0.570E-04 0.157E-04-0.426E-05
 0.114E-05-0.299E-06 0.780E-07-0.120E-02 0.984E-02-0.460E-02 0.163E-02-0.513E-03
 0.152E-03-0.431E-04 0.119E-04-0.322E-05 0.859E-06-0.226E-06 0.589E-07-0.121E-02
 0.750E-02-0.349E-02 0.123E-02-0.388E-03 0.115E-03-0.326E-04 0.900E-05-0.244E-05
 0.649E-06-0.171E-06 0.445E-07-0.114E-02
 0.647E-02-0.177E-02 0.560E-03-0.166E-03 0.472E-04-0.131E-04 0.354E-05-0.943E-06  --  VP (SECOND ROW OF UPPER TRIANGULAR FORM)
 0.248E-06-0.647E-07 0.332E-03-0.802E-02 0.378E-02-0.134E-02 0.424E-03-0.126E-03
 0.357E-04-0.987E-05 0.267E-05-0.713E-06 0.188E-06-0.489E-07 0.501E-03-0.612E-02
 0.287E-02-0.102E-02 0.321E-03-0.950E-04 0.270E-04-0.746E-05 0.202E-05-0.539E-06
 0.142E-06-0.370E-07 0.568E-03-0.466E-02 0.218E-02-0.771E-03 0.243E-03-0.719E-04
 0.204E-04-0.564E-05 0.153E-05-0.407E-06 0.107E-06-0.279E-07 0.571E-03-0.355E-02
 0.165E-02-0.584E-03 0.184E-03-0.544E-04 0.154E-04-0.427E-05 0.115E-05-0.308E-06
 0.810E-07-0.211E-07 0.539E-03

                                                                                  --  VP (ETC)
                .
                .
                .
```

computations, respectively. Note that in these tables, the variable names are placed to the right of the data file records only for readability purposes and they are not a part of the actual data file.

## B.3 PROGRAM OUTPUT

Outputs from a TSIAA program execution consist of a report file and a data file. The contents of both files vary depending on the type of TSIAA run. The following describes the program's outputs for the three TSIAA run types of intervention analysis, intervention adjustment without mean square error computations, and intervention adjustment with mean square error computations.

### Intervention Analysis - (IPT = 1)

Table B.5 presents an example report file from a TSIAA intervention analysis run. This report file corresponds to the input data file given in Table B.2. For any type of run, the TSIAA report file first lists the user supplied title and the type of run selected. Next, the general form of the second order seasonal ARMA model is displayed, followed by the values in effect for seasonality (S), nonseasonal differencing (D), seasonal differencing (SD), and the pre-intervention model parameters (AR(1), AR(2), SAR(1), SAR(2), MA(1), MA(2), SMA(1), and SMA(2)).

For an intervention analysis run, the report file then gives the number of time points input, the intervention time point index, and the calculated residual mean square of the pre-intervention time series. Next the input data series, Z(T), and the residuals, A(T), are listed. For points beyond the intervention, the values for A(T) are computed using the pre-intervention model. Next the Q-statistic (described in Section 4.1) and the p-value for a Chi-square with degrees of freedom equal to the number of time points after

## TABLE B.5. Example Report File - Intervention Analysis (IPT = 1)

```
PROGRAM TSIAA -

TITLE OF RUN:  EXAMPLE INTERVENTION ANALYSIS

 TYPE OF RUN:  INTERVENTION ANALYSIS

MODEL:
  ((1-B)**D)*((1-B**S)**SD)*(1-AR(1)B-AR(2)B**2)*(1-SAR(1)B**S-SAR(2)B**2S)*Z(T)  =
                          (1-MA(1)B-MA(2)B**2)*(1-SMA(1)B**S-SMA(2)B**2S)*A(T)

WHERE

  S =  12  D =   1  SD =   1

  PRE-INTERVENTION PARAMETERS:
           AR(1) =   0.00000E+00   AR(2) =   0.00000E+00
          SAR(1) =   0.00000E+00  SAR(2) =   0.00000E+00
           MA(1) =  -0.23700E+00   MA(2) =   0.00000E+00
          SMA(1) =   0.75500E+00  SMA(2) =   0.00000E+00

NUMBER OF TIME POINT VALUES =  84
  INTERVENTION OCCURRENCE AFTER POINT =  60

RESIDUAL MEAN SQUARE =  0.479011E+08

***    ORIGINAL DATA    ***

   T     Z(T)          A(T)

   1  0.242409E+06      ****
   2  0.250433E+06      ****
   3  0.240293E+06      ****
   4  0.226551E+06      ****
   5  0.227168E+06      ****
   6  0.227367E+06      ****
   7  0.229034E+06      ****
   8  0.232459E+06      ****
   9  0.231745E+06      ****
  10  0.228388E+06      ****
  11  0.229884E+06      ****
  12  0.233825E+06      ****
  13  0.255625E+06      ****
  14  0.258052E+06  -0.693800E+03
  15  0.264919E+06   0.159964E+05
  16  0.261065E+06   0.116150E+04
  17  0.265197E+06   0.699590E+04
  18  0.259258E+06  -0.615340E+04
  19  0.260849E+06  -0.970300E+03
  20  0.259814E+06   0.145450E+04

                .
                .
                .

  60  0.263707E+06  -0.906740E+04
  61  0.279560E+06  -0.276851E+04
```

## TABLE B.5. Example Report File - Intervention Analysis (IPT = 1) (Continued)

```
 62  0.287492E+06   0.542551E+04
 63  0.288462E+06   0.393434E+04
 64  0.275759E+06  -0.529842E+04
 65  0.261541E+06  -0.659175E+04
 66  0.244246E+06  -0.138422E+05
 67  0.230429E+06  -0.123820E+05
 68  0.236156E+06   0.125547E+05
 69  0.240372E+06   0.121438E+04
 70  0.239285E+06   0.585068E+04
                 .
                 .
                 .

 80  0.227963E+06  -0.903186E+02
 81  0.235537E+06   0.654273E+04
 82  0.237117E+06   0.575093E+04
 83  0.232663E+06  -0.125547E+05
 84  0.237073E+06  -0.548911E+03

NOTE:  AFTER T =  60, A(T) IS COMPUTED FROM THE PRE-INTERVENTION MODEL

------------------------------------------------------------------------
   Q STATISTIC =  0.305891E+02  WITH  24 DEGREES OF FREEDOM
       P-VALUE FOR A CHI-SQUARE WITH  24 DEGREES OF FREEDOM =  0.1660
------------------------------------------------------------------------

                                                   PREDICTED ERRORS DUE TO A CHANGE IN:
              ---------------------------------------------------------------------------------------------------------------
 T     A(T)         MEAN         AR(1)        AR(2)        SAR(1)       SAR(2)        MA(1)        MA(2)        SMA(1)       SMA(2)

 61 -0.277E+04  0.10000E+01 -0.90674E+04  0.83563E+04  0.45855E+04 -0.23123E+04  0.90674E+04 -0.83563E+04 -0.45855E+04  0.23123E+04
 62  0.543E+04 -0.23700E+00 -0.27685E+04 -0.90674E+04  0.11312E+05 -0.47433E+04  0.61953E+03  0.11048E+05 -0.11312E+05  0.47433E+04
 63  0.393E+04  0.56169E-01  0.54255E+04 -0.27685E+04  0.13676E+05 -0.59042E+04 -0.55723E+04  0.15017E+03 -0.13676E+05  0.59042E+04
 64 -0.530E+04 -0.13312E-01  0.39343E+04  0.54255E+04 -0.73457E+04  0.64991E+04 -0.26137E+04 -0.54611E+04  0.73457E+04 -0.64991E+04
 65 -0.659E+04  0.31550E-02 -0.52984E+04  0.39343E+04 -0.22001E+04 -0.39548E+04  0.59179E+04 -0.26401E+04  0.22001E+04  0.39548E+04
 66 -0.138E+05 -0.74772E-03 -0.65918E+04 -0.52984E+04  0.69566E+04  0.87677E+04  0.51892E+04  0.59241E+04 -0.69566E+04 -0.87677E+04
 67 -0.124E+05  0.17721E-03 -0.13842E+05 -0.65918E+04 -0.10144E+05  0.66705E+04  0.12612E+05  0.51877E+04  0.10144E+05 -0.66705E+04
 68  0.126E+05 -0.41999E-04 -0.12382E+05 -0.13842E+05  0.50055E+04 -0.73008E+04  0.93928E+04  0.12613E+05 -0.50055E+04  0.73008E+04
 69  0.121E+04  0.99538E-05  0.12555E+05 -0.12382E+05 -0.10501E+04 -0.21059E+04 -0.14781E+05  0.93928E+04  0.10501E+04  0.21059E+04
 70  0.585E+04 -0.23590E-05  0.12144E+04  0.12555E+05 -0.58273E+04 -0.69081E+04  0.22887E+04 -0.14781E+05  0.58273E+04  0.69081E+04
 71  0.568E+04  0.55909E-06  0.58507E+04  0.12144E+04  0.83563E+04 -0.15040E+03 -0.63931E+04  0.22887E+04 -0.83563E+04  0.15040E+03
 72 -0.682E+04 -0.13250E-06  0.56827E+04  0.58507E+04 -0.90674E+04  0.74187E+04 -0.41675E+04 -0.63931E+04  0.90674E+04 -0.74187E+04
 73 -0.891E+04 -0.24500E+00 -0.68215E+04  0.56827E+04 -0.27685E+04  0.45855E+04  0.78092E+04 -0.41675E+04 -0.69354E+03 -0.28397E+04
 74 -0.643E+04  0.58065E-01 -0.89090E+04 -0.68215E+04  0.54255E+04  0.11312E+05  0.70582E+04  0.78092E+04 -0.13966E+05 -0.77312E+04
 75 -0.657E+04 -0.13761E-01 -0.64257E+04 -0.89090E+04  0.39343E+04  0.13676E+05  0.47529E+04  0.70582E+04 -0.14259E+05 -0.92178E+04
 76 -0.137E+05  0.32615E-02 -0.65729E+04 -0.64257E+04 -0.52984E+04 -0.73457E+04  0.54464E+04  0.47529E+04  0.10844E+05  0.24389E+04
 77  0.482E+04 -0.77296E-03 -0.13729E+05 -0.65729E+04 -0.65918E+04 -0.22001E+04  0.12438E+05  0.54464E+04  0.82528E+04  0.51860E+04
 78  0.844E+04  0.18319E-03  0.48238E+04 -0.13729E+05 -0.13842E+05  0.69566E+04 -0.77716E+04  0.12438E+05  0.85900E+04 -0.13576E+05
 79  0.658E+04 -0.43417E-04  0.84444E+04  0.48238E+04 -0.12382E+05 -0.10144E+05 -0.66025E+04 -0.77716E+04  0.20040E+05  0.51073E+04
 80 -0.903E+02  0.10290E-04  0.65844E+04  0.84444E+04  0.12555E+05  0.50055E+04 -0.50196E+04 -0.66025E+04 -0.16334E+05  0.50660E+03
 81  0.654E+04 -0.24387E-05 -0.90319E+02  0.65844E+04  0.12144E+04 -0.10501E+04  0.12800E+04 -0.50196E+04 -0.42156E+03  0.26401E+04
 82  0.575E+04  0.57796E-06  0.65427E+04 -0.90319E+02  0.58507E+04 -0.58273E+04 -0.68461E+04  0.12800E+04 -0.14511E+04  0.11043E+05
 83 -0.126E+05 -0.13698E-06  0.57509E+04  0.65427E+04  0.56827E+04  0.83563E+04 -0.41284E+04 -0.68461E+04 -0.11992E+05 -0.82427E+04
 84 -0.549E+03  0.32464E-07 -0.12555E+05  0.57509E+04 -0.68215E+04 -0.90674E+04  0.13533E+05 -0.41284E+04  0.13667E+05  0.34663E+04
```

the intervention are given. Lastly, the intervention analysis report file lists the one-step ahead forecast errors A(T)'s) and the corresponding predicted errors due to a change in level (MEAN) and model parameters (AR(1), AR(2), SAR(1), SAR(2), MA(1), MA(2), SMA(1), and SMA(2)).

The data file produced by program TSIAA in the intervention analysis mode consists of T, A(T), and the predicted errors due to a change in level and model parameters for T = IV+1, IV+2, ..., N, where IV is the time point after which the intervention occurred and N is the number of time points input. That is, the data file contains the last eleven columns listed on the report file. These values are written to the data file in order to facilitate their input to subsequent plotting or analysis procedures. This data file contains N-IV records corresponding to the N-IV time points after the intervention. The FORTRAN format by which each record is written is (1X,I3,1X,E10.3,9E13.5).

### Intervention Adjustment without Mean Square Errors - (IPT = 2)

Table B.6 presents an example report file from a TSIAA intervention adjustment run without the mean square error computations. This report file corresponds to the data input file given in Table B.3. As in the intervention analysis mode, this report file gives the run title, run type, general model, seasonality, differencing values and the pre-intervention model parameters. However, immediately following the pre-intervention parameters, the post-intervention model parameters are also given. Then, as in the intervention analysis mode, the number of time points, intervention time point, residual mean square, Z(T)'s, and A(T)'s are listed. However, in the intervention adjustment mode, the A(T)'s, for T after the intervention, are computed using the post-intervention model parameters.

For an intervention adjustment run, the report file then lists the autocovariances of the differenced pre-intervention time series and the psi

# TABLE B.6. Example Report File - Intervention Adjustment Without MSE (IPT = 2)

```
PROGRAM TSIAA -

TITLE OF RUN:  EXAMPLE INTERVENTION ADJUSTMENT WITHOUT MSE

 TYPE OF RUN:  INTERVENTION ADJUSTMENT WITHOUT MEAN SQUARE ERROR COMPUTATION

MODEL:
  ((1-B)**D)*((1-B**S)**SD)*(1-AR(1)B-AR(2)B**2)*(1-SAR(1)B**S-SAR(2)B**2S)*Z(T) =
                            (1-MA(1)B-MA(2)B**2)*(1-SMA(1)B**S-SMA(2)B**2S)*A(T)

WHERE

  S =  12  D =   1  SD =   1

  PRE-INTERVENTION PARAMETERS:
           AR(1) =   0.00000E+00   AR(2) =   0.00000E+00
          SAR(1) =   0.00000E+00  SAR(2) =   0.00000E+00
           MA(1) =  -0.23700E+00   MA(2) =   0.00000E+00
          SMA(1) =   0.75500E+00  SMA(2) =   0.00000E+00

  POST-INTERVENTION PARAMETERS:
           AR(1) =   0.00000E+00   AR(2) =   0.00000E+00
          SAR(1) =   0.00000E+00  SAR(2) =   0.00000E+00
           MA(1) =  -0.55500E+00   MA(2) =   0.00000E+00
          SMA(1) =   0.75500E+00  SMA(2) =   0.00000E+00

NUMBER OF TIME POINT VALUES =  84
  INTERVENTION OCCURRENCE AFTER POINT =  60

RESIDUAL MEAN SQUARE =  0.479011E+08

***    ORIGINAL DATA    ***

   T      Z(T)          A(T)

   1  0.242409E+06      ****
   2  0.250433E+06      ****
   3  0.240293E+06      ****
   4  0.226551E+06      ****
   5  0.227168E+06      ****
   6  0.227367E+06      ****
   7  0.229034E+06      ****
   8  0.232459E+06      ****
   9  0.231745E+06      ****
  10  0.228388E+06      ****
  11  0.229884E+06      ****
  12  0.233825E+06      ****
  13  0.255625E+06      ****
  14  0.258052E+06  -0.693800E+03
  15  0.264919E+06   0.159964E+05
  16  0.261065E+06   0.116150E+04
  17  0.265197E+06   0.699590E+04
  18  0.259258E+06  -0.615340E+04
```

## TABLE B.6. Example Report File - Intervention Adjustment Without MSE (IPT = 2) (Continued)

```
  19  0.260849E+06  -0.970300E+03
  20  0.259814E+06   0.145450E+04

                   .
                   .
                   .

  80  0.227963E+06  -0.405885E+04
  81  0.235537E+06   0.968132E+04
  82  0.237117E+06   0.173808E+04
  83  0.232663E+06  -0.132126E+05
  84  0.237073E+06   0.532337E+04

NOTE:  AFTER T =  60, A(T) IS COMPUTED FROM THE POST-INTERVENTION MODEL


       AUTOCOVARIANCES OF THE                PSI WEIGHTS FOR THE
DIFFERENCED PRE-INTERVENTION SERIES      POST-INTERVENTION SERIES

       LAG   AUTOCOVARIANCES              INDEX    PSI WEIGHTS
       ---   ---------------              -----    -------------

         0     0.794301E+08                   0     0.100000E+01
         1     0.178238E+08                   1     0.155500E+01
         2     0.000000E+00                   2     0.155500E+01
         3     0.000000E+00                   3     0.155500E+01
         4     0.000000E+00                   4     0.155500E+01
         5     0.000000E+00                   5     0.155500E+01
         6     0.000000E+00                   6     0.155500E+01
         7     0.000000E+00                   7     0.155500E+01
         8     0.000000E+00                   8     0.155500E+01
         9     0.000000E+00                   9     0.155500E+01
        10     0.000000E+00                  10     0.155500E+01
        11    -0.857118E+07                  11     0.155500E+01
        12    -0.381967E+08                  12     0.180000E+01
        13    -0.857118E+07                  13     0.193598E+01
        14     0.000000E+00                  14     0.193598E+01
        15     0.000000E+00                  15     0.193598E+01
        16     0.000000E+00                  16     0.193598E+01
        17     0.000000E+00                  17     0.193598E+01
        18     0.000000E+00                  18     0.193598E+01
        19     0.000000E+00                  19     0.193598E+01
        20     0.000000E+00                  20     0.193598E+01

                                  .
                                  .
                                  .

        50     0.000000E+00                  50     0.307890E+01
        51     0.000000E+00                  51     0.307890E+01
        52     0.000000E+00                  52     0.307890E+01
        53     0.000000E+00                  53     0.307890E+01
        54     0.000000E+00                  54     0.307890E+01
        55     0.000000E+00                  55     0.307890E+01
        56     0.000000E+00                  56     0.307890E+01
        57     0.000000E+00                  57     0.307890E+01
        58     0.000000E+00                  58     0.307890E+01
        59     0.000000E+00                  59     0.307890E+01
```

TABLE B.6. Example Report File - Intervention Adjustment Without MSE (IPT = 2) (Continued)

```
  T    ADJUSTED Z(T)
 ---   -------------

   1     0.285362E+06
   2     0.286866E+06
   3     0.274976E+06
   4     0.261434E+06
   5     0.260256E+06
   6     0.260327E+06
   7     0.262583E+06
   8     0.263473E+06
   9     0.261867E+06
  10     0.256286E+06
  11     0.260771E+06
  12     0.266801E+06
  13     0.284963E+06
  14     0.285774E+06
  15     0.289494E+06
  16     0.285993E+06
  17     0.292455E+06
  18     0.290255E+06
  19     0.288126E+06
  20     0.289932E+06

               .
               .
               .

  40     0.235771E+06
  41     0.232439E+06
  42     0.234634E+06
  43     0.245059E+06
  44     0.241037E+06
  45     0.234789E+06
  46     0.223288E+06
  47     0.225429E+06
  48     0.239322E+06
  49     0.264882E+06
  50     0.277406E+06
  51     0.285505E+06
  52     0.274069E+06
  53     0.264723E+06
  54     0.267691E+06
  55     0.263123E+06
  56     0.261194E+06
  57     0.261323E+06
  58     0.249510E+06
  59     0.259782E+06
  60     0.263707E+06
```

TABLE B.7. Example Report File - Intervention Adjustment with MSE (IPT = 3)

```
                                     APPROX. 95% C. I. FOR ADJUSTED Z(T)
 T   ADJUSTED Z(T)     MSE Z(T)               LOWER LIMIT    UPPER LIMIT
---  -------------  --------------            -------------  ------------

  1   0.285362E+06   0.207983E+10             0.195976E+06   0.374748E+06
  2   0.286866E+06   0.201972E+10             0.198781E+06   0.374951E+06
  3   0.274976E+06   0.195961E+10             0.188211E+06   0.361740E+06
  4   0.261434E+06   0.189950E+10             0.176011E+06   0.346858E+06
  5   0.260256E+06   0.183939E+10             0.176195E+06   0.344317E+06
  6   0.260327E+06   0.177928E+10             0.177651E+06   0.343002E+06
  7   0.262583E+06   0.171917E+10             0.181316E+06   0.343850E+06
  8   0.263473E+06   0.165906E+10             0.183639E+06   0.343307E+06
  9   0.261867E+06   0.159898E+10             0.183492E+06   0.340242E+06
 10   0.256286E+06   0.153898E+10             0.179396E+06   0.333177E+06
 11   0.260771E+06   0.147908E+10             0.185392E+06   0.336151E+06
 12   0.266801E+06   0.138049E+10             0.193977E+06   0.339625E+06
 13   0.284963E+06   0.126523E+10             0.215246E+06   0.354680E+06
 14   0.285774E+06   0.122682E+10             0.217123E+06   0.354425E+06
 15   0.289494E+06   0.118841E+10             0.221927E+06   0.357062E+06
 16   0.285993E+06   0.115001E+10             0.219526E+06   0.352460E+06
 17   0.292455E+06   0.111160E+10             0.227107E+06   0.357803E+06
 18   0.290255E+06   0.107319E+10             0.226046E+06   0.354464E+06
 19   0.288126E+06   0.103478E+10             0.225076E+06   0.351175E+06
 20   0.289932E+06   0.996378E+09             0.228063E+06   0.351800E+06

                                    .
                                    .
                                    .

 40   0.235771E+06   0.334293E+09             0.199935E+06   0.271607E+06
 41   0.232439E+06   0.321824E+09             0.197278E+06   0.267600E+06
 42   0.234634E+06   0.309355E+09             0.200160E+06   0.269107E+06
 43   0.245059E+06   0.296887E+09             0.211287E+06   0.278830E+06
 44   0.241037E+06   0.284421E+09             0.207982E+06   0.274092E+06
 45   0.234789E+06   0.271971E+09             0.202465E+06   0.267112E+06
 46   0.223288E+06   0.259581E+09             0.191709E+06   0.254866E+06
 47   0.225429E+06   0.247233E+09             0.194611E+06   0.256248E+06
 48   0.239322E+06   0.211124E+09             0.210843E+06   0.267801E+06
 49   0.264882E+06   0.164586E+09             0.239737E+06   0.290027E+06
 50   0.277406E+06   0.159840E+09             0.252626E+06   0.302186E+06
 51   0.285505E+06   0.155093E+09             0.261096E+06   0.309914E+06
 52   0.274069E+06   0.150347E+09             0.250036E+06   0.298102E+06
 53   0.264723E+06   0.145601E+09             0.241073E+06   0.288373E+06
 54   0.267691E+06   0.140855E+09             0.244429E+06   0.290953E+06
 55   0.263123E+06   0.136109E+09             0.240257E+06   0.285989E+06
 56   0.261194E+06   0.131366E+09             0.238729E+06   0.283659E+06
 57   0.261323E+06   0.126641E+09             0.239266E+06   0.283380E+06
 58   0.249510E+06   0.121993E+09             0.227862E+06   0.271158E+06
 59   0.259782E+06   0.117589E+09             0.238528E+06   0.281036E+06
 60   0.263707E+06   0.479011E+08             0.250142E+06   0.277272E+06
```

weights of the post-intervention time series. Lastly, in the case of intervention adjustment without the mean square error computations, the adjusted Z(T) are listed for T = 1, 2, ..., IV.

The data file produced by program TSIAA in the intervention adjustment mode without the mean square error computations consists of T and the adjusted Z(T) for T = 1, 2, ..., IV. That is, the data file contains the last two columns listed on the report file. These values are written to the data file in order to facilitate their input to subsequent plotting or analysis procedures. This data file contains IV records corresponding to the IV time points prior to the intervention. The FORTRAN format by which each record is written is (1X,I3,2X,E13.6).

### Intervention Adjustment with Mean Square Errors - (IPT = 3)

For the intervention adjustment run with the mean square error computations, the report file is essentially identical (with the exception of the run type and possibly the user supplied run title) to the report file for the intervention adjustment without the mean square error computations up to the point at which the adjusted Z(T)'s are listed. At this point the covariance matrix of the estimated pi weights (VP matrix) will be displayed if requested (i.e., the input field IPRT is set equal to one). This matrix is displayed on the report file in lower triangular form. Then, in addition to T and the adjusted Z(T), the report file also lists the mean square error of the adjusted Z(T) and the approximate 95% confidence interval's lower and upper limits for the adjusted Z(T). Table B.7 presents the last portion of an example report file from a TSIAA intervention adjustment run with the mean square error computations.

The data file produced by program TSIAA in the intervention adjustment mode with the mean square error computations consists of T, adjusted Z(T), the mean square error of the adjusted Z(T), and the lower and upper 95% confidence interval limits for the adjusted Z(T) for T = 1, 2, ..., IV. That is, the data file contains the last five columns listed on the report file. These values are written to the data file in order to facilitate their input to subsequent plotting or analysis procedures. This data file contains IV records corresponding to the IV time points prior to the intervention. The FORTRAN format by which each record is written is

```
(1X,I3,2X,E13.6,2X,E13.6,7X,E13.6,2X,E13.6)
```

## B.4 PROGRAM OPERATION

This section describes how to use program TSIAA together with other facilities available on the EIA computer to perform time series intervention analysis and adjustment. This description follows the methodology presented in Appendix A.

### Pre-intervention Modeling

If the original time series data are available on a file, the SAS-ETS procedure ARIMA can be used to perform the pre-intervention modeling. In addition, the resulting residuals can be output to a data file with this procedure. Table B.B gives example JCL and SAS commands to perform the modeling and to output both the original time series and the residuals to a data file. In this example, the file SERIES.DATA contains the original time series and the file PREINT.DATA contains both the original time series and the residuals.

### Testing for an Intervention

Before using TSIAA in the intervention analysis mode, it is necessary to prepare the data file containing the original time series and the residuals for input to TSIAA. This involves editing the data file so that it includes the

TABLE B.8. Example JCL and SAS Commands for Pre-Intervention Modeling

```
//STEP1        EXEC SAS,OPTIONS='S=72 NOCENTER',
//             TIME.SAS=(1,20),REGION=264K
//***
//INF          DD DSN=SERIES.DATA,DISP=OLD
//***
//OUTF         DD DSN=PREINT.DATA,DISP=(NEW,CATLG)...
//***
//SYSIN        DD *
               DATA SERIES;
                  INFILE INF;
                  INPUT @10 Z 9.;
                  PROC ARIMA DATA=SERIES;
                       IDENTIFY VAR=Z(1,12);
                       ESTIMATE Q=(1)(12);
                       FORECAST LEAD=24 OUT=PREINT;
               DATA _NULL_;
                  SET PREINT;
                  FILE OUTF;
                     PUT @7 Z 12. @19 RESIDUAL 12.1;
```

TITLE, IPT, N, IV, IS, ID, ICD, PHI, THETA, CPHI, CTHETA and FMT fields described in the Program Input section of this appendix. This type of file editing can be performed on the EIA computer with the SUPERWYLBUR utility.

Once the TSIAA input data file is ready, TSIAA can be executed. In preparing the JCL to run TSIAA, note that the program reads the input data on FORTRAN unit 5 and writes the report on unit 6 and the output data on unit 1. Table B.9 gives example JCL commands to run TSIAA. In this example the input data file is named TSIAA.IN, the output data file is named TSIAA.DAT and the report is on TSIAA.RPT. The TSIAA report file contains the Q-statistic and its p-level. From this information it can be determined if the intervention is significant.

### Determining the Nature of the Intervention

If, from the results of the TSIAA intervention analysis run, it is decided that the intervention is significant, then the TSIAA output data file (which contains the one-step ahead forecast errors and the predicted errors due to a change in level and model parameters) can be input to the SAS procedures PLOT and REG. The PLOT procedure is used to plot the predicted errors versus the one-step ahead forecast errors to observe similarities. If similarities are seen, then the REG procedure is used to regress the one-step ahead forecast errors on the sets of predicted errors to estimate changes in the parameters. The TSIAA data file can be read directly by SAS. The record format of this file is described in the Program Output section of this appendix.

### Adjusting the Pre-intervention Data and Estimating the Errors of Adjustment

Once the post-intervention model parameters have been determined, TSIAA can then be used to adjust the pre-intervention data. This involves adding the post-intervention parameters PHIN, THETN, CPHIN, and CTHETN to the TSIAA intervention analysis input data file. In addition, the IPT field is modified

TABLE B.9. Example JCL Commands to Run TSIAA

```
//AAAAZZZZ JOB (1234,FOR,1,99),'PNAME',TIME=(0,20),REGION=300K
//***
//***
//STEP1 EXEC     FORTGCLG
//FORT.SYSIN     DD DSN=TSIAA.FOR,DISP=OLD
//***
//GO.FT05F001    DD DSN=TSIAA.IN,DISP=OLD
//***
//GO.FT06F001    DD DSN=TSIAA.RPT,DISP=(NEW,CATLG),
//                  UNIT=DASD,SPACE=(TRK,10),
//                  DCB=(RECFM=FB,BLKSIZE=13200,LRECL=132)
//***
//FT01F001       DD DSN=TSIAA.DAT,DISP=(NEW,CATLG),
//                  UNIT=DASD,SPACE=(TRK,10),
//                  DCB=(RECFM=FB,BLKSIZE=13200,LRECL=132)
//***
```

to indicate an intervention adjustment run. If only the adjustment of the pre-intervention data is desired, then IPT is set equal to 2. IPT is set equal to 3 if, in addition to the adjusted data, the adjustment errors are desired. However, this requires that the fields VP and FMT (the covariance matrix of the pi weights and its format) be added to the input data file. The necessary editing of the data file can be performed by using SUPERWYLBUR. The ordering and formats of the additional fields for an intervention adjustment run are described in the Program Input section of this appendix. Once the input data file for the TSIAA intervention adjustment run is prepared, TSIAA is executed in the same manner as for the intervention analysis run (refer to Table B.9). The TSIAA report file contains the adjusted data and, if requested, the adjusted errors. The TSIAA data file in the intervention adjustment mode also contains the adjusted values and, if requested, the adjusted errors. If desired, this data file can be input directly to SAS for plotting. Its record format is given in the Program Output section of this appendix.

## B.5 PROGRAM DESCRIPTION

Program TSIAA consists of 13 modules. These are the main program TSIAA, eleven subroutine subprograms and one function subprogram. A description of each module is given below. Table B.10 lists the major program variables together with a short description of each variable. To simplify the variable descriptions and to provide a link between the computer program and the theory presented in Section 4.1 and Section 5.5, Table B.10 also includes equivalent symbolic representations for many of the program variables. Thus, variables and their symbols are used interchangeably in the descriptions that follow.

TABLE B.10. Major TSIAA Program Variables
(In Order of Occurrence in the Program)

| VARIABLE | SYMBOL | DESCRIPTION |
|---|---|---|
| NMAX | | Maximum number of time points allowed by TSIAA |
| IU | | Data input unit number |
| IOU | | Data output unit number |
| IPT | | TSIAA run type |
| IPRT | | VP matrix print flag |
| N | $N$ | Number of original time points |
| IV | $t_o$ | Time point after which intervention occurred |
| IS | $s$ | Order of seasonality |
| ID | $d$ | Degree of non-seasonal differencing |
| ICD | $D$ | Degree of seasonal differencing |
| PHI(1) | $\phi_1$ | AR1, pre-intervention model |
| PHI(2) | $\phi_2$ | AR2, pre-intervention model |
| THETA(1) | $\theta_1$ | MA1, pre-intervention model |
| THETA(2) | $\theta_2$ | MA2, pre-intervention model |
| CPHI(1) | $\Phi_1$ | SAR1, pre-intervention model |
| CPHI(2) | $\Phi_2$ | SAR2, pre-intervention model |
| CTHETA(1) | $\Theta_1$ | SMA1, pre-intervention model |

### TABLE B.10. Major TSIAA Program Variables (In Order of Occurrence in the Program) (Continued)

| VARIABLE | SYMBOL | DESCRIPTION |
|---|---|---|
| CTHETA(2) | $\Theta_2$ | SMA2, pre-intervention model |
| Z(I) | $z_i$ | time series |
| A(I) | $a_i$ | residuals |
| PHIN(1) | $\phi'_1$ | AR1, post-intervention model |
| PHIN(2) | $\phi'_2$ | AR2, post-intervention model |
| THETN(1) | $\theta'_1$ | MA1, post-intervention model |
| THETN(2) | $\theta'_2$ | MA2, post-intervention model |
| CPHIN(1) | $\Phi'_1$ | SAR1, post-intervention model |
| CPHIN(2) | $\Phi'_2$ | SAR2, post-intervention model |
| CTHETN(1) | $\Theta'_1$ | SMA1, post-intervention model |
| CTHETN(2) | $\Theta'_2$ | SMA2, post-intervention model |
| VP(I,J) | $\Pi_{ij}$ | Covariance matrix of the estimated pi weights of the pre-intervention model |
| IDIF |  | Total degree of differencing (IDIF = d + Ds) |
| XP |  | Number of non-zero pre-intervention model parameters |
| SSQA | $\sigma_a^2$ | Residual mean square, pre-intervention model |

TABLE B.10. Major TSIAA Program Variables
(In Order of Occurrence in the Program)
(Continued)

| VARIABLE | SYMBOL | DESCRIPTION |
|---|---|---|
| DCOEF(I) | | Coefficients of polynomial $(1-B)^d$ |
| CDCOEF(I) | | Coefficients of polynomial $(1-B^s)^D$ |
| WCOEF(I,J) | | Coefficients of polynomial $(1-B)^d\ (1-B^s)^D$ |
| W(I) | $w_i$ | differenced $z_i$<br>$(w_i = (1-B)^d(1-B^s)^D z_i)$ |
| Q | Q | Q-statistic |
| IDF | | Degrees of freedom for Q-statistic ($IDF = N-t_o$) |
| PVAL | | p-value for Q-statistic |
| CX(I) | X | Predicted errors due to a change in mean level |
| CWPHI(I,1) | $W_i, \phi_1$ | Predicted errors due to a change in $\phi_1$ |
| CWPHI(I,2) | $W_i, \phi_2$ | Predicted errors due to a change in $\phi_2$ |
| CWTHE(I,1) | $W_i, \theta_1$ | Predicted errors due to a change in $\theta_1$ |
| CWTHE(I,2) | $W_i, \theta_2$ | Predicted errors due to a change in $\theta_2$ |
| CWCPHI(I,1) | $W_i, \Phi_1$ | Predicted errors due to a change in $\Phi_1$ |
| CWCPHI(I,1) | $W_i, \Phi_2$ | Predicted errors due to a change in $\Phi_2$ |
| CWCTHE(I,1) | $W_i, \Theta_1$ | Predicted errors due to a change in $\Theta_1$ |
| CWCTHE(I,2) | $W_i, \Theta_2$ | Predicted errors due to a change in $\Theta_2$ |

TABLE B.10. Major TSIAA Program Variables (In Order of Occurrence in the Program) (Continued)

| VARIABLE | SYMBOL | DESCRIPTION |
|---|---|---|
| PSI (I) | $\psi_i$ | Psi weights for the post-intervention series |
| GAMMA(I) | $W_{ij}$ | Autocovariances of the differenced pre-intervention series ($W_{ij}$ = GAMMA($\lvert$I-J$\rvert$ + 1)) |
| XMSE (I) | | Mean square error of the adjusted $z_i$ |

## Main Program TSIAA

Main program TSIAA contains the dimension statements for all arrays which depend on the number of points in the time series. Currently the maximum number of time points allowed by the program is 150. If intervention analysis and adjustment are desired for a time series with greater than 150 points, then the program must be modified. That is, all arrays with a dimension of 150 must have the dimension increased to the desired number. In addition, the variable NMAX (which is set in the first assignment statement of the program) must be assigned the new value. Main program TSIAA also contains the assignment statements for the unit number variables of the input data file (IU = 5) and the output data file (IOU = 1). The program directs all report output to unit 6.

Upon program initiation, TSIAA calls subroutine INDAT to read the input data file. Next TSIAA computes the residual mean square of the input $a_i$'s and calls subroutine DIFFZ to difference the input $z_i$'s to get the $w_i$'s. Then TSIAA calculates the values of $a_i$ from $i = t_0 + 1$ to $N$ using the pre-intervention model in the intervention analysis mode or the post-intervention model in the intervention adjustment mode.

If the intervention analysis mode is in effect, then TSIAA computes the Q-statistic according to equation 4 in Section 4.1. TSIAA calls subroutine GAMCUM to calculate the Chi-square p-value for the Q-statistic with $N - t_0$ degrees of freedom. Lastly, in the intervention analysis mode, TSIAA computes the predicted errors due to a change in mean level (as described in Section 4.1) and calls subroutine PARAMC to calculate the predicted errors due to changes in the model parameters.

If the intervention adjustment mode is in effect, then TSIAA calls subroutine ADJSTW to perform the intervention adjustment on the $w_i$'s. TSIAA undifferences the adjusted $w_i$'s to get the adjusted $z_i$'s. Next, TSIAA computes the autocovariances of the differenced pre-intervention time series ($W_{ij}$'s) and the psi weights for the post-intervention series ($\Psi_i$'s). Lastly, if requested, TSIAA calculates the mean square errors of the adjusted $z_i$'s (using the $W_{ij}$'s, $\Psi_i$'s, and $\Pi_{ij}$'s according to equations 21 and 27 in Section 5.5) and the 95% confidence interval limits about the adjusted values.

Subroutine ADJSTW

ADJSTW is called by TSIAA to adjust the $w_i$'s for $i = t_o, t_o - 1, \ldots 1$ according to the method described in Section 5.5.

Subroutine AMISS

AMISS is called by TSIAA to determine the A(I)'s which are undefined based on the current model (i.e., order of seasonality, degrees of seasonal and non-seasonal differencing, and the non-zero model parameters).

Subroutine BINEXP

BINEXP is called by DIFFZ to compute the coefficients of the polynomial

$$(1 - x)^{\alpha} .$$

Subroutine DIFFEQ

DIFFEQ is called by TSIAA to compute $a_i$ using the equality

$$(1 - \phi_1 B - \phi_2 B^2) \; (1 - \Phi_1 B^s - \Phi_2 B^{2s}) \; w_i =$$
$$(1 - \theta_1 B - \theta_2 B^2) \; (1 - \Theta_1 B^s - \Theta_2 B^{2s}) \; a_i \; .$$

Where

$\phi_1$, $\phi_2$, $\Phi_1$, $\Phi_2$, $\theta_1$, $\theta_2$, $\Theta_1$, $\Theta_2$, s, $w_i$ and $a_i$ are input parameters.

Subroutine DIFFZ

DIFFZ is called by TSIAA to compute

$$w_i = (1 - B^d) \; (1 - B^s)^D \; z_i \quad .$$

### Subroutine GAMCUM

GAMCUM is called by TSIAA to compute the Chi-square p-value for the Q-statistic.

### Function GAMLN

GAMLN is used by GAMCUM to compute the natural logarithm of the gamma function with argument IDF/2.

### Subroutine INDAT

INDAT is called by TSIAA to read the input data file on unit IU. The number of input variable values read depends on the value of the input variable IPT. INDAT checks the input values for consistency and writes the values of the input variables TITLE, IS, ID, ICD, PHI, CPHI, THETA, CTHETA, PHIN, THETN, CPHIN, CTHETN, N, and IV to the report file.

### Subroutine PARAMC

PARAMC is called by TSIAA to compute the predicted errors due to changes in the model according to the method described in Section 4.1. That is, PARAMC computes

$$W_{i,\phi_k} = \phi_1 W_{i-1,\phi_k} + \phi_2 W_{i-2,\phi_k} + a_{i-k}$$

$$W_{i,\Phi_k} = \Phi_1 W_{i-s,\Phi_k} + \Phi_2 W_{i-2s,\Phi_k} + a_{i-ks}$$

$$W_{i,\theta_k} = \theta_1 W_{i-1,\theta_k} + \theta_2 W_{i-2,\theta_k} - a_{i-k}$$

$$W_{i,\Theta_k} = \Theta_1 W_{i-s,\Theta_k} + \Theta_2 W_{i-2s,\Theta_k} - a_{i-ks}$$

for $k = 1,2$.

### Subroutine POLYD

POLYD is called by TSIAA to perform polynomial division.

### Subroutine PRTORG

PRTORG is called by TSIAA to write the input values of Z and VP to the report file. The input values of A and the values of A computed from the current model (i.e., for the points occurring after the intervention), are also written to the report file.

### Subroutine RATIOC

RATIOC is called by TSIAA to compute the coefficients of the polynomials

$$(1 - \theta_1 B - \theta_2 B^2)\ (1 - \Theta_1 B^s - \Theta_2 B^{2s})$$

and

$$(1 - \phi_1 B - \phi_2 B^2)\ (1 - \Phi_1 B^s - \Phi_2 B^{2s})$$

where

$\theta_1$, $\theta_2$, $\Theta_1$, $\Theta_2$, $\phi_1$, $\phi_2$, $\Phi_1$, $\Phi_2$ and s are input parameters.

## B.6 PROGRAM LISTING

The following pages comprise the FORTRAN source code listing of program TSIAA.

```
C     PROGRAM TSIAA
C
C
C     -    -      -            -            -
C     TIME SERIES INTERVENTION ANALYSIS AND ADJUSTMENT
C
C
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION Z(150),A(150),W(150),WCOEF(150,150)
      DIMENSION CDCOEF(150),DCOEF(150),IAMS(150)
      DIMENSION XTMP(150),X(150),CX(150)
      DIMENSION PHI(2),THETA(2),CPHI(2),CTHETA(2)
      DIMENSION PHIN(2),THETN(2),CPHIN(2),CTHETN(2)
      DIMENSION KP(2),KT(2),KCP(2),KCT(2)
      DIMENSION CWPHI(150,2),CWCPHI(150,2)
      DIMENSION CWTHE(150,2),CWCTHE(150,2)
      DIMENSION ANUM(150),ADEN(150),PSIP(150)
      DIMENSION GAMMA(150),ADEN2(150),PDEN(150)
      DIMENSION PSI(150),VP(150,150),XMSE(150)
      DIMENSION PARAM(2)
C
C     FOR THE Q-STATISTIC, CHI-SQUARE TEST --
C
C       PARAM(1) = ALPHA = DF/2 (DF = DEGREES OF FREEDOM) AND
C       PARAM(2) =  BETA =  1/2  IN THE GAMMA DISTRIBUTION:
C
C         X**(ALPHA-1) * EXP(BETA*X)
C
C
      DATA PARAM(2) / 0.5 /
      DATA NOPRAM / 2 /
C
C
C
C
C
C     NMAX = MAXIMUM NUMBER OF TIME POINTS ALLOWED
      NMAX=150
C
C
C     IU = DATA INPUT UNIT NUMBER
      IU=5
C
C     IOU = DATA OUTPUT UNIT NUMBER
      IOU=1
C
      WRITE (6,200)
C
C
C     READ DATA
      CALL INDAT (IU,NMAX,IPT,IPRT,N,IV,IS,ID,ICD,PHI,THETA,CPHI,
     *            CTHETA,Z,A,PHIN,THETN,CPHIN,CTHETN,VP)
C
C
C     IDIF = TOTAL DEGREE OF DIFFERENCING
      IDIF=ID + ICD*IS
      IVP1=IV + 1
```

```
      IDP1=ID + 1
      ICDP1=ICD + 1
C
C     DETERMINE MISSING VALUES OF A (PRE-INTERVENTION MODEL)
      CALL AMISS (N,IV,A,IAMS,IS,IDIF,PHI,THETA,CPHI,CTHETA,
     *            KP,KT,KCP,KCT)
C
C     CALCULATE SSQA (RESIDUAL MEAN SQUARE, PRE-INTERVENTION MODEL)
C
      XN=0
      SUM=0
      DO 10 I=1,IV
         IF (IAMS(I) .EQ. 0) SUM=SUM + A(I)
         IF (IAMS(I) .EQ. 0) XN=XN + 1
   10 CONTINUE
C
      XMEAN=SUM/XN
C
C     XP =  NUMBER OF NON-ZERO MODEL PARAMETERS
      XP=0
      DO 13 I=1,2
         XP=XP + KP(I) + KT(I) + KCP(I) + KCT(I)
   13 CONTINUE
C
      SSQA=0
      DO 15 I=1,IV
         IF (IAMS(I) .EQ. 0)
     *   SSQA=SSQA + (A(I) - XMEAN)*(A(I) - XMEAN)
   15 CONTINUE
C
      IF (XN .LE. XP) SSQA=0
      IF (XN .GT. XP) SSQA=SSQA/(XN - XP)
C
      WRITE (6,210) SSQA
C
C     DIFFERENCE Z TO GET W
      CALL DIFFZ (NMAX,N,Z,ID,ICD,IS,DCOEF,CDCOEF,WCOEF,W)
C
      IF (IPT .EQ. 2 .OR. IPT .EQ. 3) GO TO 20
C
C     CALCULATE A(T) FROM IV+1 ON (PRE-INTERVENTION MODEL)
      CALL DIFFEQ (W,N,IVP1,PHI,THETA,CPHI,CTHETA,IS,A,IAMS)
C
      GO TO 25
C
   20 CONTINUE
C
C     DETERMINE MISSING VALUES OF A (POST-INTERVENTION MODEL)
      CALL AMISS (N,IV,A,IAMS,IS,IDIF,PHIN,THETN,CPHIN,CTHETN,
     *            KP,KT,KCP,KCT)
C
C     CALCULATE A(T) FROM IV+1 ON (POST-INTERVENTION MODEL)
      CALL DIFFEQ (W,N,IVP1,PHIN,THETN,CPHIN,CTHETN,IS,A,IAMS)
C
C
   25 CONTINUE
C
C
C     PRINT REMAINING INPUT DATA AND THE A(T) FROM IV+1 ON
      CALL PRTORG (NMAX,N,IV,Z,A,IAMS,IPT,IPRT,VP)
```

```
C
      IF (IPT .EQ. 2 .OR. IPT .EQ. 3) GO TO 70
C
C
C     CONTINUATION OF THE INTERVENTION ANALYSIS
C
C     CALCULATE THE Q STATISTIC AND CHI-SQUARE P-VALUE
C
      Q=0
      DO 30 I=IVP1,N
         Q=Q + A(I)*A(I)
   30 CONTINUE
C
      IF (SSQA .EQ. 0.0) Q=0.0
      IF (SSQA .NE. 0.0) Q=Q/SSQA
C
      IDF=N - IV
      PARAM(1)=FLOAT(IDF)/2.0
C
      CALL GAMCUM (NOPRAM,PARAM,Q,PVAL)
      PVAL=1.0 - PVAL
C
      WRITE (6,215) Q, IDF, IDF, PVAL
C
C
C     COMPUTE PREDICTED ERRORS DUE TO CHANGE IN LEVEL OR PARAMETERS
C
C
C     LEVEL CHANGE
C
C     FORM XTMP = ARRAY OF ZEROS AND ONES CORRESPONDING TO
C                 PRE-INTERVENTION AND POST-INTERVENTION TIME POINTS
      DO 40 I=1,N
         CX(I)=0.
         IF (I .LE. IV) XTMP(I)=0
         IF (I .GT. IV) XTMP(I)=1
   40 CONTINUE
C
C     DIFFERENCE XTMP TO GET X
      DO 50 K=1,N
         X(K)=0
         IF (K .LE. IV) GO TO 50
         DO 45 I=1,IDP1
         DO 45 J=1,ICDP1
            KNDX=(I-1) + (J-1)*IS
            X(K)=X(K) + WCOEF(I,J)*XTMP(K-KNDX)
   45    CONTINUE
   50 CONTINUE
C
C
C     CALCULATE CX FROM X USING THE DIFFERENCE EQUATION
      IB=1
      CALL DIFFEQ (X,N,IB,PHI,THETA,CPHI,CTHETA,IS,CX,IAMS)
C
C
C
C     CHANGE IN PARAMETERS
C
      CALL PARAMC (NMAX,N,IV,IS,A,PHI,THETA,CPHI,CTHETA,KP,KT,KCP,KCT,
     *             CWPHI,CWTHE,CWCPHI,CWCTHE)
```

```
C
C     PRINT OUT PREDICTED ERRORS AND ALSO WRITE
C     THEM TO THE OUTPUT DATA FILE
      WRITE (6,220)
C
      DO 60 I=IVP1,N
         WRITE (6,225) I,A(I),CX(I),CWPHI(I,1),CWPHI(I,2),
     *                 CWCPHI(I,1),CWCPHI(I,2),CWTHE(I,1),CWTHE(I,2),
     *                 CWCTHE(I,1),CWCTHE(I,2)
         WRITE (IOU,225) I,A(I),CX(I),CWPHI(I,1),CWPHI(I,2),
     *                   CWCPHI(I,1),CWCPHI(I,2),CWTHE(I,1),CWTHE(I,2),
     *                   CWCTHE(I,1),CWCTHE(I,2)
   60 CONTINUE
C
C
C     END OF THE INTERVENTION ANALYSIS
C
      GO TO 9999
C
C
   70 CONTINUE
C
C     CONTINUATION OF THE INTERVENTION ADJUSTMENT
C
C
C     PERFORM ADJUSTMENT TO W
      CALL ADJSTW (N,A,IV,IS,PHIN,THETN,CPHIN,CTHETN,W)
C
C     UNDIFFERENCE ADJUSTED W TO GET ADJUSTED Z
C
      DO 75 J=1,IV
C        WORK BACKWARDS
         I=IV - J + 1
         INDX=IDIF
         SUM=W(I+INDX)
         DO 72 K=1,IDP1
         DO 72 L=1,ICDP1
            IF (K .EQ. IDP1 .AND. L .EQ. ICDP1) DIV=WCOEF(K,L)
            IF (K .EQ. IDP1 .AND. L .EQ. ICDP1) GO TO 73
            KNDX=(K-1) + (L-1)*IS
C           MOVE TO RHS (SUBTRACT)
            SUM=SUM - WCOEF(K,L)*Z(I+INDX-KNDX)
   72    CONTINUE
   73    CONTINUE
         Z(I)=SUM/DIV
   75 CONTINUE
C
C
C
C     ERROR BOUND COMPUTATION
C
C
C     COMPUTE AUTOCOVARIANCES OF THE DIFFERENCED PRE-INTERVENTION
C     TIME SERIES (GAMMA)
C
C     MAX DIMENSION OF PSIP POLYNOMIAL
      IPDIM=MAX0(IV,2*IS+2)
      IF (IPDIM .GT. NMAX) IPDIM=NMAX
C
C     GET NUMERATOR AND DENOMINATOR POLYNOMIAL COEFFICIENTS OF PSIP
```

```
      CALL RATIOC (IPDIM,IS,PHI,THETA,CPHI,CTHETA,ANUM,ADEN)
C
C     COMPUTE PSIP=ANUM/ADEN
      CALL POLYD (ANUM,ADEN,IPDIM,PSIP)
C
C     COMPUTE POSITIVE TERMS OF GAMMA(B)=SSQA*PSIP(B)*PSIP(F)
      DO 77 I=1,IPDIM
         GAMMA(I)=0.
         K=0
         DO 76 J=I,IPDIM
            K=K + 1
            GAMMA(I)=GAMMA(I) + PSIP(K)*PSIP(J)
  76     CONTINUE
         GAMMA(I)=SSQA*GAMMA(I)
  77  CONTINUE
C
C
C
C     COMPUTE MEAN SQUARE ERROR OF THE ADJUSTED Z
C
C
C     MAX DIMENSION OF POLYNOMIAL PSI = ANUM/(ADEN*ADEN2)
      IDIM=MAX0(IDIF+2*IS+2,IV)
      IF (IDIM .GT. NMAX) IDIM=NMAX
C
C     GET COEFFICIENTS OF THE POLYNOMIALS ANUM AND ADEN
      CALL RATIOC (IDIM,IS,PHIN,THETN,CPHIN,CTHETN,ANUM,ADEN)
C
C     COMPUTE COEFFICIENTS (ADEN2) OF THE POLYNOMIAL
C     ((1-B)**ID)((1-B**IS)**ICD)
      DO 79 K=1,IDIM
         ADEN2(K)=0.
         DO 78 I=1,IDP1
         DO 78 J=1,ICDP1
            IF ((I-1) + IS*(J-1) + 1 .EQ. K)
     *         ADEN2(K)=ADEN2(K) + WCOEF(I,J)
  78     CONTINUE
  79  CONTINUE
C
C     GET COEFFICIENTS OF PDEN=ADEN*ADEN2
      DO 81 I=1,IDIM
         PDEN(I)=0.
         K=I
         DO 80 J=1,I
            PDEN(I)=PDEN(I) + ADEN2(J)*ADEN(K)
            K=K - 1
  80     CONTINUE
  81  CONTINUE
C
C     COMPUTE PSI=ANUM/PDEN
      CALL POLYD (ANUM,PDEN,IDIM,PSI)
C
C
C     PRINT OUT GAMMA AND PSI
      WRITE (6,230)
C
      DO 85 I=1,IV
            INDX=I - 1
         WRITE (6,235) INDX, GAMMA(I), INDX, PSI(I)
  85  CONTINUE
```

```
C
C
      IF (IPT .EQ. 3) GO TO 90
C
C     PRINT OUT ADJUSTED Z'S AND ALSO WRITE
C     THEM TO THE OUTPUT DATA FILE
C
      WRITE (6,240)
      DO 88 I=1,IV
         WRITE (6,242) I, Z(I)
         WRITE (IOU,242) I, Z(I)
   88 CONTINUE
C
C     END OF THE INTERVENTION ADJUSTMENT WITHOUT MSE COMPUTATIONS
C
      GO TO 9999
C
C
   90 CONTINUE
C
C
C
C     XMSE(T) = MEAN SQUARE ERROR OF ADJUSTED Z(T)
C
C
      DO 95 I=1,IV
         INDX=IV - I + 1
         XMSE(INDX)=0.
         DO 94 K=1,I
            KNDX=IV - I + K
            SUM=0.
            SUM2=0.
            IF (KNDX .EQ. IV) SUM=SSQA
            IF (KNDX .EQ. IV) GO TO 93
            KNDXM1=IV - KNDX
            DO 92 L=1,KNDXM1
               SUM=SUM + GAMMA(1)*VP(L,L)
               IF (L .EQ. KNDXM1) GO TO 92
               LP1=L + 1
               DO 91 M=LP1,KNDXM1
                  SUM2=SUM2 + GAMMA(IABS(L-M)+1)*VP(L,M)
   91          CONTINUE
   92       CONTINUE
C
   93       CONTINUE
            XMSE(INDX)=XMSE(INDX) + PSI(K)*PSI(K)*(SUM + 2.*SUM2)
   94    CONTINUE
   95 CONTINUE
C
C     PRINT OUT THE ADJUSTED Z'S WITH MSE
C     ALSO WRITE THEM TO THE OUTPUT DATA FILE
      WRITE (6,250)
C
      DO 98 I=1,IV
         CI=1.96*SQRT(XMSE(I))
         XL=Z(I) - CI
         XU=Z(I) + CI
         WRITE (6,255) I,Z(I),XMSE(I),XL,XU
         WRITE (IOU,255) I,Z(I),XMSE(I),XL,XU
   98 CONTINUE
```

```
C
C
C
 9999 CONTINUE
C
C
      STOP
C
C
 200  FORMAT ('0PROGRAM TSIAA - ')
 210  FORMAT ('0RESIDUAL MEAN SQUARE = ',E13.6)
 215  FORMAT (1X,/,'0',74('-'),/,
     1         '      Q STATISTIC = ',E13.6,'  WITH ',I3,
     2         ' DEGREES OF FREEDOM',/,
     3         '          P-VALUE FOR A CHI-SQUARE WITH ',I3,
     4         ' DEGREES OF FREEDOM = ',F7.4,/,
     5         ' ',74('-'),/,1X)
 220  FORMAT ('0',50X,'PREDICTED ERRORS DUE TO A CHANGE IN:',/,
     1         17X,115('-'),/,
     1         2X,'T',2X,3X,'A(T)',3X,5X,'MEAN',4X,4X,'AR(1)',4X,
     2         4X,'AR(2)',4X,4X,'SAR(1)',3X,4X,'SAR(2)',3X,
     3         4X,'MA(1)',4X,4X,'MA(2)',4X,4X,'SMA(1)',3X,
     4         4X,'SMA(2)',/,1X)
 225  FORMAT (1X,I3,1X,E10.3,9E13.5)
 230  FORMAT (1X,/,'0','       AUTOCOVARIANCES OF THE      ',5X,
     1             '   PSI WEIGHTS FOR THE',/,
     2         ' ','DIFFERENCED PRE-INTERVENTION SERIES',5X,
     3             'POST-INTERVENTION SERIES',/,
     4         '0','       LAG',3X,'AUTOCOVARIANCES       ',5X,
     5             '  INDEX',3X,' PSI WEIGHTS ',/,
     6             8X,'---',3X,15('-'),12X,2X,'-----',3X,13('-'),/,1X)
 235  FORMAT (8X,I3,3X,1X,E13.6,15X,1X,I3,1X,3X,E13.6)
 240  FORMAT (1X,/,'0',' T ',2X,'ADJUSTED Z(T)',/,
     1              ' ','---',2X,'-------------',/,1X)
 242  FORMAT (1X,I3,2X,E13.6)
 250  FORMAT (1X,/,'0',37X,'APPROX. 95% C. I.',
     1         ' FOR ADJUSTED Z(T)',/,
     2         1X,' T ',2X,'ADJUSTED Z(T)',2X,'   MSE Z(T)  ',7X,
     3         ' LOWER LIMIT ',2X,' UPPER LIMIT',/,
     4         1X,'---',2X,'-------------',2X,'---------------',7X,
     5         '-------------',2X,'------------',/,1X)
 255  FORMAT (1X,I3,2X,E13.6,2X,E13.6,7X,E13.6,2X,E13.6)
C
      END
```

```
      SUBROUTINE ADJSTW (N,A,IV,IS,PHIN,THETN,CPHIN,CTHETN,W)
C
C
C     ROUTINE TO ADJUST W FROM POINT IV TO 1 USING THE DIFFERENCE
C       EQUATION:
C
C      W(T)-PHIN(1)*W(T-1)-PHIN(2)*W(T-2)-CPHIN(1)*W(T-IS)
C       +PHIN(1)*CPHIN(1)*W(T-(IS+1))+PHIN(2)*CPHIN(1)*W(T-(IS+2))
C       -CPHIN(2)*W(T-2*IS)+PHIN(1)*CPHIN(2)*W(T-(2*IS+1))
C       +PHIN(2)*CPHIN(2)*W(T-(2*IS+2)) =
C      A(T)-THETN(1)*A(T-1)-THETN(2)*A(T-2)-CTHETN(1)*A(T-IS)
C       +THETN(1)*CTHETN(1)*A(T-(IS+1))+THETN(2)*CTHETN(1)*A(T-(IS+2))
C       -CTHETN(2)*A(T-2*IS)+THETN(1)*CTHETN(2)*A(T-(2*IS+1))
C       +THETN(2)*CTHETN(2)*A(T-(2*IS+2))
C
C
C    VARIABLES:
C
C         N  -  NUMBER OF TIME POINTS
C         A  -  RESIDUALS (INPUT FOR TIME 1 TO IV, COMPUTED VIA THE
C                 POST-INTERVENTION MODEL FOR TIME IV+1 TO N)
C        IV  -  TIME POINT AFTER WHICH INTERVENTION OCCURRED
C        IS  -  DEGREE OF SEASONALITY
C     ****  POST-INTERVENTION MODEL PARAMETERS
C      PHIN  -  COEFFICIENTS (2) OF THE PHI POLYNOMIAL
C     THETN  -  COEFFICIENTS (2) OF THE THETA POLYNOMIAL
C     CPHIN  -  COEFFICIENTS (2) OF THE CAPITIAL PHI POLYNOMIAL
C    CTHETN  -  COEFFICIENTS (2) OF THE CAPITAL THETA POLYNOMIAL
C     ****
C         W  -  DIFFERENCED INPUT TIME SERIES
C
C    LOCAL VARIABLES:
C
C      COEL  -  COEFFICIENTS OF THE LEFT HAND SIDE OF THE EQUATION
C      COER  -  COEFFICIENTS OF THE RIGHT HAND SIDE OF THE EQUATION
C      ISUB  -  SUBSCRIPTS OF THE LEFT (W) AND RIGHT (A) SIDE TERMS
C       IMA  -  FOR EACH TERM OF THE LEFT HAND SIDE OF THE EQUATION,
C                 IMA(TERM) =  1 IF THE MINIMUM SUBSCRIPT IS ATTAINED
C                 AT THIS TERM, OTHERWISE IMA(TERM) = 0
C
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION A(N),W(N),PHIN(2),THETN(2),CPHIN(2),CTHETN(2)
      DIMENSION COEL(9),COER(9),ISUB(9),IMA(9)
C
C
C
      COEL(1)=1
      COEL(2)=-1.*PHIN(1)
      COEL(3)=-1.*PHIN(2)
      COEL(4)=-1.*CPHIN(1)
      COEL(5)=PHIN(1)*CPHIN(1)
      COEL(6)=PHIN(2)*CPHIN(1)
      COEL(7)=-1.*CPHIN(2)
      COEL(8)=PHIN(1)*CPHIN(2)
      COEL(9)=PHIN(2)*CPHIN(2)
C
      COER(1)=1
```

```
      COER(2)=-1.*THETN(1)
      COER(3)=-1.*THETN(2)
      COER(4)=-1.*CTHETN(1)
      COER(5)=THETN(1)*CTHETN(1)
      COER(6)=THETN(2)*CTHETN(1)
      COER(7)=-1.*CTHETN(2)
      COER(8)=THETN(1)*CTHETN(2)
      COER(9)=THETN(2)*CTHETN(2)
C
      ISUB(1)=0
      ISUB(2)=1
      ISUB(3)=2
      ISUB(4)=IS
      ISUB(5)=IS + 1
      ISUB(6)=IS + 2
      ISUB(7)=2*IS
      ISUB(8)=2*IS + 1
      ISUB(9)=2*IS + 2
C
      ISMAX=0
      DO 50 I=2,9
         IF (COEL(I) .NE. 0.) ISMAX=MAX0(ISMAX,ISUB(I))
   50 CONTINUE
C
      DO 60 I=1,9
         IMA(I)=0
         IF (ISUB(I) .EQ. ISMAX) IMA(I)=1
   60 CONTINUE
C
      DO 70 J=1,IV
C        WORK BACKWARDS
         I=IV - J + 1
         W(I)=0.
         SUM=0.
         DIV=0.
C        COMPUTE TERMS - LHS FIRST
         DO 65 K=1,9
C           IGNORE TERMS WHERE COEF IS 0
            IF (COEL(K) .EQ. 0.) GOTO 65
            INDX=I + ISMAX - ISUB(K)
C           CHECK FOR VALID TERM
            IF (INDX .LT. 1 .OR. INDX .GT. N) GO TO 70
C           USE IMA TO DETERMINE IF MOVE TO RHS (SUBTRACT) OR DIVIDE
            IF (IMA(K) .EQ. 0) SUM=SUM - COEL(K)*W(INDX)
            IF (IMA(K) .EQ. 1) DIV=DIV + COEL(K)
   65    CONTINUE
C
C        COMPUTE TERMS OM RHS NOW
         DO 67 K=1,9
C           IGNORE TERMS WHERE COEF IS 0
            IF (COER(K) .EQ. 0.) GO TO 67
            INDX=I + ISMAX - ISUB(K)
C           CHECK FOR VALID TERM
            IF (INDX .LT. 1 .OR. INDX .GT. N) GO TO 70
C           ADD TERM TO SUM
            SUM=SUM + COER(K)*A(INDX)
   67    CONTINUE
C
         W(I)=SUM/DIV
   70 CONTINUE
C
C
C
      RETURN
      END
```

```
      SUBROUTINE AMISS (N,IV,A,IAMS,IS,IDIF,PHI,THETA,CPHI,CTHETA,
     *                  KP,KT,KCP,KCT)
C
C
C     ROUTINE TO DETERMINE MISSING VALUES OF A
C
C     PARAMETERS:
C
C          N - NUMBER OF TIME POINT VALUES
C         IV - NUMBER OF INPUT VALUES OF A
C                (SAME AS THE INTERVENTION POINT)
C          A - INPUT ARRAY OF RESIDUALS
C       IAMS - IAMS(I) = 1 IF A(I) IS UNDEFINED, OTHERWISE IAMS(I) = 0
C         IS - ORDER OF SEASONALITY
C       IDIF - TOTAL DEGREE OF DIFFERENCING
C       PHI, THETA, CPHI, CTHETA - PARAMETERS OF THE MODEL
C        KP,    KT,  KCP,    KCT - SET = 1 IF THE CORRESPONDING
C                                  MODEL PARAMETERS ARE NON-ZERO
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION A(N),IAMS(N)
      DIMENSION PHI(2),THETA(2),CPHI(2),CTHETA(2)
      DIMENSION KP(2),KT(2),KCP(2),KCT(2)
C
C
      DO 5 I=1,2
         KP(I)=0
         IF (PHI(I) .NE. 0.) KP(I)=1
         KT(I)=0
         IF (THETA(I) .NE. 0.) KT(I)=1
         KCP(I)=0
         IF (CPHI(I) .NE. 0.) KCP(I)=1
         KCT(I)=0
         IF (CTHETA(I) .NE. 0.) KCT(I)=1
    5 CONTINUE
C
C
      NMISS=IDIF
      DO 10 I=1,2
         NMISS=MAX0(NMISS,(IDIF+I)*KP(I))
         NMISS=MAX0(NMISS,I*KT(I))
         NMISS=MAX0(NMISS,(IDIF+I*IS)*KCP(I))
         NMISS=MAX0(NMISS,I*IS*KCT(I))
C
         DO 8 J=1,2
             NMISS=MAX0(NMISS,(IDIF+J*IS+I)*KCP(J)*KP(I))
             NMISS=MAX0(NMISS,(J*IS+I)*KCT(J)*KT(I))
    8    CONTINUE
   10 CONTINUE
C
C
      DO 15 I=1,N
         IAMS(I)=0
         IF (I .LE. NMISS) A(I)=0.
         IF (I .LE. NMISS) IAMS(I)=1
   15 CONTINUE
C
C
RETURN
END
```

```
      SUBROUTINE BINEXP (N,DCOEF,ID)
C
C
C     ROUTINE TO COMPUTE BINOMIAL EXPANSION COEFFICIENTS OF
C        (1-X)**ID
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION DCOEF(N)
C
C
C
      IDP1=ID + 1
C
      DO 25 I=1,IDP1
         DCOEF(I)=0
   25 CONTINUE
C
      DCOEF(1)=1
      IF (ID .EQ. 0) GOTO 35
      DO 32 I=2,IDP1
         DCOEF(I)=DCOEF(I-1)*(ID-I+2)/(I-1)
   32 CONTINUE
C
      DO 33 I=2,IDP1
         DCOEF(I)=(-1.*(I-1))*DCOEF(I)
   33 CONTINUE
C
C
   35 CONTINUE
C
C
      RETURN
      END
```

```
      SUBROUTINE DIFFEQ (W,N,IB,PHI,THETA,CPHI,CTHETA,IS,A,IAMS)
C
C
C     ROUTINE TO CALCULATE A(T) FOR T = IB,...N USING
C
C  (1-PHI(1)*B-PHI(2)*B**2)(1-CPHI(1)*B**IS-CPHI(2)*B**2IS)W(T)  =
C  (1-THETA(1)*B-THETA(2)*B**2)(1-CTHETA(1)*B**IS-CTHETA(2)*B**2IS)A(T)
C
C     OTHER PARAMETERS:
C
C       IAMS - IAMS(I) = 1 IF A(I) IS UNDEFINED, OTHERWISE IAMS(I) = 0
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION W(N),A(N),IAMS(N)
      DIMENSION PHI(2),THETA(2),CPHI(2),CTHETA(2)
C
C
C
      DO 50 I=IB,N
         IF (IAMS(I) .EQ. 1) GO TO 50
C
         A(I)=W(I) - PHI(1)*W(I-1) - PHI(2)*W(I-2) + THETA(1)*A(I-1)
     *        + THETA(2)*A(I-2)
C
         A(I)=A(I)
     *        + CPHI(1)*W(I-IS) + CTHETA(1)*A(I-IS)
     *        + PHI(1)*CPHI(1)*W(I-(IS+1))
     *        - THETA(1)*CTHETA(1)*A(I-(IS+1))
     *        + PHI(2)*CPHI(1)*W(I-(IS+2))
     *        - THETA(2)*CTHETA(1)*A(I-(IS+2))
     *        - CPHI(2)*W(I-2*IS) + CTHETA(2)*A(I-2*IS)
     *        + PHI(1)*CPHI(2)*W(I-(2*IS+1))
     *        - THETA(1)*CTHETA(2)*A(I-(2*IS+1))
     *        + PHI(2)*CPHI(2)*W(I-(2*IS+2))
     *        - THETA(2)*CTHETA(2)*A(I-(2*IS+2))
C
 50   CONTINUE
C
C
      RETURN
      END
```

```
      SUBROUTINE DIFFZ (NMAX,N,Z,ID,ICD,IS,DCOEF,CDCOEF,WCOEF,W)
C
C
C     ROUTINE TO COMPUTE W(T) = ((1-B)**ID)*((1-B**IS)**ICD)*Z(T)
C
C     OTHER PARAMETERS:
C
C       NMAX - MAXIMUM DIMENSION OF ARRAYS
C          N - NUMBER OF POINTS
C      DCOEF - ARRAY OF COEFFICIENTS FOR (1-B)**ID
C     CDCOEF - ARRAY OF COEFFICIENTS FOR (1-B**IS)**ICD
C      WCOEF - ARRAY OF COEFFICIENTS FOR ((1-B)**ID)*((1-B**IS)**ICD)
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION Z(N),DCOEF(N),CDCOEF(N),WCOEF(NMAX,N),W(N)
C
C
C
C     CALCULATE POLYNOMIAL COEFFICIENTS
      CALL BINEXP (N,DCOEF,ID)
      CALL BINEXP (N,CDCOEF,ICD)
C
      IDP1=ID + 1
      ICDP1=ICD + 1
C
      DO 40 I=1,IDP1
      DO 40 J=1,ICDP1
         WCOEF(I,J)=DCOEF(I)*CDCOEF(J)
   40 CONTINUE
C
C     NMISS = NUMBER OF UNDEFINED W(T)
      NMISS=ID + ICD*IS
C
      DO 45 K=1,N
         W(K)=0
         IF (K .LE. NMISS) GO TO 45
         DO 43 I=1,IDP1
         DO 43 J=1,ICDP1
            KNDX=(I-1) + (J-1)*IS
            W(K)=W(K) + WCOEF(I,J)*Z(K-KNDX)
   43    CONTINUE
   45 CONTINUE
C
C
      RETURN
      END
```

```
      SUBROUTINE GAMCUM(NOPRAM,PARAM,X,P)
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION PARAM(1)
C
C
C     GAMCUM COMPUTES THE CUMULATIVE (0, X) PROBABILITY
C     P FOR THE GAMMA DISTRIBUTION WITH PARAMETERS ALPHA AND BETA.
C     FORM OF GAMMA IS
C                    U**(ALPHA-1) * EXP(BETA*U).
C
C    ON ENTRY:
C            NOPRAM = 0  DEFAULT VALUES FOR ALPHA (=1) AND
C                        BETA (=1) ASSUMED
C                   = 2  USER SPECIFIES PARAM(1) = ALPHA
C                                       PARAM(2) = BETA
C
C            PARAM  - ARRAY OF USER DEFINED VALUES.
C                     MAY BE DUMMY ARGUMENT IF NOPRAM = 0.
C
C                 X - VALUE CUMULATIVE PROBABILITY DESIRED FOR
C
C    ON RETURN:
C                 P - CUMULATIVE PROBABILITY OF GAMMA(ALPHA,BETA)
C                      (0, X)
C    PRINTING: NONE
C
C    COMMON BLOCKS: NONE
C
C    EXTERNAL REFERENCES:
C            FUNCTION - GAMLN
C
C---- IF X.GE.(ALPHA/2 + 4) THE ASYMPTOTIC EXPANSION GIVEN BY
C     EQ. 6.5.32 IN ABRAMOWITZ AND STEGUN IS USED.
C     OTHERWISE, A CONFLUENT HYPERGEOMETRIC FUNCTION REPRESENTATION
C     FOR THE INCOMPLETE GAMMA FUNCTION IS USED.  SEE EQUATIONS
C     6.5.12 AND 13.1.2 IN ABRAMOWITZ AND STEGUN.
C---- THE RESULTS OF THE ROUTINE WERE CHECKED AGAINST TABLE 26.7
C     IN ABRAMOWITZ AND STEGUN
C
C
      DATA ERR/1.E-4/
C
      IF(NOPRAM) 2,1,2
C
C---- DEFAULT VALUES
    1 ALPHA=1.
      BETA =1.
      GO TO 3
C
    2 ALPHA=PARAM(1)
      BETA =PARAM(2)
C
    3 A=ALPHA
      IF(X.GT.0.) GO TO 4
      P=0.
      RETURN
C
    4 Y=BETA*X
```

```
      IF(Y.LT.(A/2.+4.)) GO TO 30
      SUM=1.
      R=1.
      L=INT(A)
      DO 10 I=1,L
      AI=I
      R=R*(A-AI)/Y
      IF(R.LT.ERR) GO TO 20
   10 SUM=SUM+R
   20 P=1.-SUM*EXP((A-1.)*ALOG(Y)-Y-GAMLN(A))
      RETURN
C
   30 SUM=1.
      R=1.
      DO 40 I=1,50
      AI=I
      R=R*Y/(A+AI)
      IF(R.LT.ERR) GO TO 50
   40 SUM=SUM+R
   50 P=SUM/A*EXP(A*ALOG(Y)-Y-GAMLN(A))
      RETURN
      END
      FUNCTION GAMLN(ALPHA)
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
C---- COMPUTES LN(GAMMA(A))
C
C---- FOR A.LT.4 USES RATIONAL FUNCTION EXPANSION GIVEN BY
C     WILK,GNANADESIKAN,HUYETT(1962) TECHNOMETRICS 4_1-18
C     ATTRIBUTED TO HASTINGS(1955) APPROXIMATIONS FOR
C     DIGITAL COMPUTERS.
C
C     FOR A.GE.4 THE ASYMPTOTIC EXPANSION GIVEN BY EQ.6.1.41
C     IN ABRAMOWITZ AND STEGUN IS USED.
C
C---- THE RESULTS WERE CHECKED AGAINST THE VALUES TABLED IN
C     ABRAMOWITZ AND STEGUN_ PP. 267 AND 274
C
C    ON ENTRY:
C            ALPHA = VALUE FOR GAMMA FUNCTION ARGUMENT
C
C    ON RETURN:
C            GAMLN = LN(GAMMA(ALPHA))
C
C    PRINTING: NONE
C    COMMON BLOCKS: NONE
C    EXTERNAL REFERENCES: NONE
C
C
      DIMENSION B(8),C(4)
      DATA B/.577191652,.988205891,.897056937,.918206857,
     1       .756704078,.482199394,.193527818,.035868343/
      DATA C/12.,-360.,1260.,-1680./
C
      IF(ALPHA.GE.4.) GO TO 100
C
C---- RATIONAL FUNCTION APPROXIMATION ALPHA.LT.4
C
      IA=INT(ALPHA)-1
```

```
      AF=ALPHA-AINT(ALPHA)
      G=B(8)*AF
      DO 10 J=1,7
      I=8-J
   10 G=(B(I)-G)*AF
      G=1.-G
      IF(IA) 20,50,30
   20 G=G/ALPHA
      GO TO 50
   30 DO 40 I=1,IA
      AI=I
   40 G=G*(AF+AI)
   50 GAMLN=ALOG(G)
      RETURN
C
C---- ASYMPTOTIC APPROXIMATION FOR ALPHA.GE.4
C
  100 G=0.
      DO 110 I=1,4
  110 G=G+1./(C(I)*ALPHA**(2*I-1))
      GAMLN=ALPHA*(ALOG(ALPHA)-1.)+ALOG(6.2831853/ALPHA)/2.+G
      RETURN
      END
```

```
      SUBROUTINE INDAT (IU,NMAX,IPT,IPRT,N,IV,IS,ID,ICD,PHI,THETA,
     *                 CPHI,CTHETA,Z,A,PHIN,THETN,CPHIN,CTHETN,VP)
C
C
C     ROUTINE TO READ INPUT DATA
C
C     PARAMETERS:
C
C         IU - DATA INPUT UNIT NUMBER
C       NMAX - MAXIMUM NUMBER OF ALLOWABLE TIME POINTS
C        IPT - TYPE OF RUN
C               (1=INTERVENTION ANALYSIS, 2=INTERVENTION ADJUSTMENT
C                WITHOUT MSE, 3=INTERVENTION ADJUSTMENT WITH MSE)
C       IPRT - IF IPRT = 1 THEN THE INPUT VP MATRIX IS PRINTED,
C               OTHERWISE THE VP MATRIX IS NOT PRINTED
C          N - NUMBER OF TIME POINTS
C         IV - TIME POINT AFTER WHICH INTERVENTION OCCURRED
C         IS - ORDER OF SEASONALITY
C         ID - DEGREE OF NON-SEASONAL DIFFERENCING
C        ICD - DEGREE OF SEASONAL DIFFERENCING
C        ****  PRE-INTERVENTION MODEL PARAMETERS
C        PHI - COEFFICIENTS (2) OF THE PHI POLYNOMIAL
C      THETA - COEFFICIENTS (2) OF THE THETA POLYNOMIAL
C       CPHI - COEFFICIENTS (2) OF THE CAPITAL PHI POLYNOMIAL
C     CTHETA - COEFFICIENTS (2) OF THE CAPITAL THETA POLYNOMIAL
C        ****
C          Z - ORIGINAL TIME POINT VALUES
C          A - RESIDUALS (PRE-INTERVENTION MODEL)
C
C      ***** FOLLOWING INPUT ONLY WHEN IPT = 2 OR 3
C
C        ****  POST-INTERVENTION MODEL PARAMETERS
C       PHIN - COEFFICIENTS (2) OF THE PHI POLYNOMIAL
C      THETN - COEFFICIENTS (2) OF THE THETA POLYNOMIAL
C      CPHIN - COEFFICIENTS (2) OF THE CAPITAL PHI POLYNOMIAL
C     CTHETN - COEFFICIENTS (2) OF THE CAPITAL THETA POLYNOMIAL
C        ****
C
C      ***** FOLLOWING INPUT ONLY WHEN IPT = 3
C
C         VP - COVARIANCE MATRIX OF THE ESTIMATED PI WEIGHTS
C               OF THE PRE-INTERVENTION MODEL
C
C
C     LOCAL VARIABLES:
C
C        FMT - FORMAT SPECIFICATION FOR Z, A, AND VP ARRAYS
C      TITLE - TITLE OF RUN
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION PHI(2),THETA(2),CPHI(2),CTHETA(2)
      DIMENSION PHIN(2),THETN(2),CPHIN(2),CTHETN(2)
      DIMENSION Z(NMAX),A(NMAX),VP(NMAX,NMAX)
      DIMENSION TITLE(20),FMT(20)
C
C
C
C     ITERM = EARLY TERMINATION FLAG
```

```
      ITERM=0
C
C
C
      READ (IU,120) TITLE
      READ (IU,100) IPT,IPRT
      READ (IU,100) N,IV
      READ (IU,100) IS,ID,ICD
      READ (IU,110) PHI
      READ (IU,110) THETA
      READ (IU,110) CPHI
      READ (IU,110) CTHETA
C
      WRITE (6,200) TITLE
C
C     ASSUME IPT = 1 IF IT IS NOT VALID
      IF (IPT .LT. 1 .OR. IPT .GT. 3) IPT=1
C     SET ID, ICD, IS = ZERO IF NEGATIVE
      IF (ID .LT. 0) ID=0
      IF (ICD .LT. 0) ICD=0
      IF (IS .LT. 0) IS=0
C
      IF (IPT .EQ. 1) WRITE (6,205)
      IF (IPT .EQ. 2) WRITE (6,210)
      IF (IPT .EQ. 3) WRITE (6,211)
C
C     INSURE THAT IF THERE IS NO SEASONALITY, THEN THE CORRESPONDING
C     POLYNOMIAL AND DIFFERENCING PARAMETERS ARE ZERO
C
      IF (IS .GT. 0) GO TO 2
C
      ICD=0
      DO 1 I=1,2
         CPHI(I)=0.
         CTHETA(I)=0.
    1 CONTINUE
C
    2 CONTINUE
C
C
      WRITE (6,213)
      WRITE (6,215)
      WRITE (6,220)
      WRITE (6,225) IS,ID,ICD
      WRITE (6,227)
      WRITE (6,230) PHI
      WRITE (6,235) CPHI
      WRITE (6,240) THETA
      WRITE (6,245) CTHETA
C
C
C     CHECK THAT THE INPUT VALUES ARE VALID
C     IF NOT, THEN STOP
C
      IF (N .LE. NMAX) GO TO 5
      WRITE (6,300) NMAX
      ITERM=1
C
    5 CONTINUE
C
```

```
      IF (ID + ICD*IS .LT. IV) GO TO 6
      WRITE (6,305)
      ITERM=1
C
    6 CONTINUE
C
      IF (IV .GE. 1 .AND. IV .LT. N) GO TO 7
      WRITE (6,310) N
      ITERM=1
C
    7 CONTINUE
C
      IF (ITERM .EQ. 0) GO TO 10
C     DATA INPUT NOT VALID
      WRITE (6,315)
      STOP
C
C
   10 CONTINUE
C
C
C     READ TIME POINT VALUES OF Z AND A
      READ (IU,120) FMT
      READ (IU,FMT) (Z(I),I=1,N)
C
      READ (IU,120) FMT
      READ (IU,FMT) (A(I),I=1,IV)
C
C
      IF (IPT .EQ. 1) GO TO 50
C
C     ADDITIONAL DATA FOR INTERVENTION ADJUSTMENT RUNS
C     READ POST-INTERVENTION MODEL PARAMETERS
C
      READ (IU,110) PHIN
      READ (IU,110) THETN
      READ (IU,110) CPHIN
      READ (IU,110) CTHETN
C
C     IF THERE IS NO SEASONALITY, INSURE PARAMETER CONSISTENCY
      IF (IS .NE. 0) GO TO 25
      DO 20 I=1,2
         CPHIN(I)=0.
         CTHETN(I)=0.
   20 CONTINUE
C
   25 CONTINUE
      IF (IPT .EQ. 2) GO TO 40
C
C     ADDITIONAL DATA FOR TYPE = 3 INTERVENTION ADJUSTMENT RUN
C     READ VP MATRIX
C
      READ (IU,120) FMT
C     ASSUME ONLY UPPER TRIANGLE OF VP IS INPUT
      DO 30 I=1,IV
      READ (IU,FMT) (VP(I,J),J=I,IV)
   30 CONTINUE
C
C     FORM REST OF SYMMETRIC VP MATRIX
      DO 35 I=1,IV
```

```
      DO 35 J=1,I
         VP(I,J)=VP(J,I)
   35 CONTINUE
C
C
   40 CONTINUE
C
      WRITE (6,250)
      WRITE (6,230) PHIN
      WRITE (6,235) CPHIN
      WRITE (6,240) THETN
      WRITE (6,245) CTHETN
C
C
C
   50 CONTINUE
C
      WRITE (6,255) N,IV
C
C
      RETURN
C
C
  100 FORMAT (3I5)
  110 FORMAT (2F10.0)
  120 FORMAT (20A4)
  200 FORMAT ('0TITLE OF RUN:  ',20A4)
  205 FORMAT ('0 TYPE OF RUN:  INTERVENTION ANALYSIS')
  210 FORMAT ('0 TYPE OF RUN:  INTERVENTION ADJUSTMENT WITHOUT MEAN',
     1        ' SQUARE ERROR COMPUTATION')
  211 FORMAT ('0 TYPE OF RUN:  INTERVENTION ADJUSTMENT WITH MEAN',
     1        ' SQUARE ERROR COMPUTATION')
  213 FORMAT ('0MODEL:')
  215 FORMAT ('    ((1-B)**D)*((1-B**S)**SD)*(1-AR(1)B-AR(2)B**2)',
     *         '*(1-SAR(1)B**S-SAR(2)B**2S)*Z(T) = ',/,
     *        28X,' (1-MA(1)B-MA(2)B**2)*(1-SMA(1)B**S-SMA(2)B**2S)',
     *         '*A(T)')
  220 FORMAT ('0WHERE')
  225 FORMAT ('0  S = ',I3,'  D = ',I3,'  SD = ',I3)
  227 FORMAT ('0  PRE-INTERVENTION PARAMETERS:')
  230 FORMAT ('             AR(1)  = ',E13.5, '   AR(2)  = ',E13.5)
  235 FORMAT ('            SAR(1)  = ',E13.5, '  SAR(2)  = ',E13.5)
  240 FORMAT ('             MA(1)  = ',E13.5, '   MA(2)  = ',E13.5)
  245 FORMAT ('            SMA(1)  = ',E13.5, '  SMA(2)  = ',E13.5)
  250 FORMAT ('0  POST-INTERVENTION PARAMETERS:')
  255 FORMAT ('0NUMBER OF TIME POINT VALUES = ',I3,/,
     *        '    INTERVENTION OCCURRENCE AFTER POINT = ',I3)
  300 FORMAT ('0****  THE NUMBER OF TIME POINTS',
     *         ' EXCEEDS THE PROGRAM MAXIMUM OF ',I4)
  305 FORMAT ('0****  AFTER DIFFERENCING, ALL TIME POINT VALUES',
     *         ' PRIOR TO THE INTERVENTION WILL BE UNDEFINED')
  310 FORMAT ('0****  THE INTERVENTION MUST HAVE OCCURRED AFTER',
     *         ' TIME POINT 1 AND BEFORE TIME POINT ',I4)
  315 FORMAT ('0****  PROGRAM TERMINATED  ****')
C
      END
```

```
      SUBROUTINE PARAMC (NMAX,N,IV,IS,A,PHI,THETA,CPHI,CTHETA,KP,KT,
     *                   KCP,KCT,CWPHI,CWTHE,CWCPHI,CWCTHE)
C
C
C     ROUTINE TO COMPUTE THE PREDICTED ERRORS DUE TO CHANGES IN THE
C       PARAMETERS OF THE EQUATION:
C
C     A(T) =
C
C      (1-PHI(1)*B-PHI(2)*B**2)(1-CPHI(1)*B**IS-CPHI(2)*B**2IS)  Z(T)
C    -------------------------------------------------------------------
C     (1-THETA(1)*B-THETA(2)*B**2)(1-CTHETA(1)*B**IS-CTHETA(2)*B**2IS)
C
C
C     OTHER VARIABLES:
C
C          NMAX   -   MAXIMUM NUMBER OF ALLOWABLE TIME POINTS
C             N   -   NUMBER OF TIME POINTS
C            IV   -   TIME POINT AFTER WHICH INTERVENTION OCCURRED
C          KP, KT, KCP, KCT   -   SET = 1 IF THE CORRESPONDING PARAMETERS
C                                 (PHI, THETA, CPHI, CTHETA) ARE NON ZERO
C         CWPHI   -   ERRORS WITH RESPECT TO PHI(1), PHI(2)
C         CWTHE   -   ERRORS WITH RESPECT TO THETA(1), THETA(2)
C        CWCPHI   -   ERRORS WITH RESPECT TO CPHI(1), CPHI(2)
C        CWCTHE   -   ERRORS WITH RESPECT TO CTHETA(1), CTHETA(2)
C
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION A(N),PHI(2),THETA(2),CPHI(2),CTHETA(2)
      DIMENSION KP(2),KT(2),KCP(2),KCT(2)
      DIMENSION CWPHI(NMAX,2),CWTHE(NMAX,2)
      DIMENSION CWCPHI(NMAX,2),CWCTHE(NMAX,2)
C
C
C
C     WITH RESPECT TO PHI
C
      DO 35 I=1,2
C
      NMISS=MAX0(KP(1),2*KP(2),I)
      NMISS=MAX0(NMISS,IV)
C
      DO 33 J=1,N
         CWPHI(J,I)=0.
         IF (J .LE. NMISS) GO TO 33
         CWPHI(J,I)=PHI(1)*CWPHI(J-1,I) + PHI(2)*CWPHI(J-2,I)
     *              + A(J-I)
   33 CONTINUE
   35 CONTINUE
C
C
C
C     WITH RESPECT TO CPHI
C
      DO 45 I=1,2
C
      NMISS=MAX0(IS*KCP(1),2*IS*KCP(2),I*IS)
      NMISS=MAX0(NMISS,IV)
```

```
C
      DO 43 J=1,N
         CWCPHI(J,I)=0.
         IF (J .LE. NMISS) GO TO 43
         CWCPHI(J,I)=CPHI(1)*CWCPHI(J-IS,I) + CPHI(2)*CWCPHI(J-2*IS,I)
     *               + A(J-I*IS)
   43 CONTINUE
   45 CONTINUE
C
C
C     WITH RESPECT TO THETA
C
      DO 55 I=1,2
C
      NMISS=MAX0(KT(1),2*KT(2),I)
      NMISS=MAX0(NMISS,IV)
C
      DO 53 J=1,N
         CWTHE(J,I)=0.
         IF (J .LE. NMISS) GO TO 53
         CWTHE(J,I)=THETA(1)*CWTHE(J-1,I) + THETA(2)*CWTHE(J-2,I)
     *              - A(J-I)
   53 CONTINUE
   55 CONTINUE
C
C
C
C     WITH RESPECT TO CTHETA
C
      DO 65 I=1,2
C
      NMISS=MAX0(IS*KCT(1),2*IS*KCT(2),I*IS)
      NMISS=MAX0(NMISS,IV)
C
      DO 63 J=1,N
         CWCTHE(J,I)=0.
         IF (J .LE. NMISS) GO TO 63
         CWCTHE(J,I)=CTHETA(1)*CWCTHE(J-IS,I)
     *               + CTHETA(2)*CWCTHE(J-2*IS,I) - A(J-I*IS)
   63 CONTINUE
   65 CONTINUE
C
C
C
      RETURN
      END
```

```
      SUBROUTINE POLYD (H,G,N,D)
C
C
C     ROUTINE TO PERFORM POLYNOMIAL DIVISION
C
C     PARAMETERS:
C
C        H - NUMERATOR
C        G - DENOMINATOR
C        D - QUOTIENT
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION H(N),G(N),D(N)
C
C
C
      DO 1 I=1,N
         D(I)=H(I)
    1 CONTINUE
C
      ALPHA=1./G(1)
C
      DO 5 J=1,N
         BETA=ALPHA*D(J)
         K=J+1
         DO 5 I=K,N
            D(I)=D(I) - (BETA*G(I-J+1))
    5 CONTINUE
C
      DO 10 J=1,N
         D(J)=D(J)*ALPHA
   10 CONTINUE
C
C
      RETURN
      END
```

```
      SUBROUTINE PRTORG (NMAX,N,IV,Z,A,IAMS,IPT,IPRT,VP)
C
C     ROUTINE TO PRINT OUT ORIGINAL INPUT ARRAYS
C
C     PARAMETERS:
C
C       NMAX  -  MAXIMUM DIMENSION OF ARRAYS
C          N  -  NUMBER OF TIME POINTS
C         IV  -  INTERVENTION INDEX
C          Z  -  TIME SERIES
C          A  -  RESIDUALS (PRE-INTERVENTION MODEL)
C       IAMS  -  IAMS(I) = 1 IF A(I) IS UNDEFINED, OTHERWISE IAMS(I)=0
C        IPT  -  IF IPT=1, THEN THIS IS AN INTERVENTION ANALYSIS RUN
C                IF IPT=2, THEN THIS IS AN INTERVENTION ADJUSTMENT RUN
C                          WITHOUT MSE CALCULATIONS
C                IF IPT=3, THEN THIS IS AN INTERVENTION ADJUSTMENT RUN
C                          WITH MSE CALCULATIONS
C       IPRT  -  IF IPRT=1 THEN THE INPUT VP MATRIX IS PRINTED OUT,
C                  OTHERWISE THE VP MATRIX IS NOT PRINTED
C         VP  -  COVARIANCE MATRIX OF THE ESTIMATED PI WEIGHTS
C                  OF THE PRE-INTERVENTION MODEL
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
      DIMENSION Z(N),A(N),IAMS(N),VP(NMAX,IV)
C
C
      WRITE (6,100)
      WRITE (6,110)
      WRITE (6,113)
C
      DO 5 I=1,N
         IF (IAMS(I) .EQ. 0) WRITE (6,115) I,Z(I),A(I)
         IF (IAMS(I) .EQ. 1) WRITE (6,120) I,Z(I)
   5  CONTINUE
C
      IF (IPT .EQ. 1) WRITE (6,125) IV
      IF (IPT .EQ. 2 .OR. IPT .EQ. 3) WRITE (6,126) IV
C
      IF (IPT .EQ. 1 .OR. IPT .EQ. 2) GO TO 20
      IF (IPRT .NE. 1) GO TO 20
C
      WRITE (6,130)
      WRITE (6,113)
C
      DO 15 I=1,IV
         WRITE (6,133) I
         WRITE (6,135) (VP(I,J),J=1,I)
  15  CONTINUE
C
  20  CONTINUE
C
C
      RETURN
C
C
 100  FORMAT ('0***    ORIGINAL DATA    ***')
 110  FORMAT ('0   T     Z(T)           A(T)')
 113  FORMAT (' ')
 115  FORMAT (1X,I4,1X,E13.6,2X,E13.6)
 120  FORMAT (1X,I4,1X,E13.6,6X,'****')
 125  FORMAT ('0NOTE:  AFTER T =',I4,', A(T) IS COMPUTED FROM THE ',
     *        'PRE-INTERVENTION MODEL')
 126  FORMAT ('0NOTE:  AFTER T =',I4,', A(T) IS COMPUTED FROM THE ',
     *        'POST-INTERVENTION MODEL')
 130  FORMAT ('0COVARIANCE MATRIX OF THE ESTIMATED PI WEIGHTS')
 133  FORMAT (' T =',I4)
 135  FORMAT (1X,10E13.5)
C
      END
```

```
      SUBROUTINE RATIOC (IDIM,IS,PHI,THETA,CPHI,CTHETA,ANUM,ADEN)
C
C
C     ROUTINE TO COMPUTE THE NUMERATOR AND DENOMINATOR COEFFICIENTS OF:
C
C     (1-THETA(1)*B-THETA(2)*B**2)(1-CTHETA(1)*B**IS-CTHETA(2)*B**2IS)
C   -----------------------------------------------------------------------
C         (1-PHI(1)*B-PHI(2)*B**2)(1-CPHI(1)*B**IS-CPHI(2)*B**2IS)
C
C
      IMPLICIT REAL*4 (A-H, O-Z)
C
C
      DIMENSION PHI(2),THETA(2),CPHI(2),CTHETA(2)
      DIMENSION ANUM(IDIM),ADEN(IDIM)
C
C
      DO 10 I=1,IDIM
         ANUM(I)=0.
         ADEN(I)=0.
  10  CONTINUE
C
      ANUM(1)=1.
      ADEN(1)=1.
      ANUM(2)=-1.*THETA(1)
      ADEN(2)=-1.*PHI(1)
      ANUM(3)=-1.*THETA(2)
      ADEN(3)=-1.*PHI(2)
C
      DO 20 I=1,IDIM
         IF (I .EQ. IS+1) ANUM(I)=ANUM(I)   -1.*CTHETA(1)
         IF (I .EQ. IS+1) ADEN(I)=ADEN(I)   -1.*CPHI(1)
         IF (I .EQ. IS+2) ANUM(I)=ANUM(I) + THETA(1)*CTHETA(1)
         IF (I .EQ. IS+2) ADEN(I)=ADEN(I) + PHI(1)*CPHI(1)
         IF (I .EQ. IS+3) ANUM(I)=ANUM(I) + THETA(2)*CTHETA(1)
         IF (I .EQ. IS+3) ADEN(I)=ADEN(I) + PHI(2)*CPHI(1)
         IF (I .EQ. 2*IS+1) ANUM(I)=ANUM(I)   -1.*CTHETA(2)
         IF (I .EQ. 2*IS+1) ADEN(I)=ADEN(I)   -1.*CPHI(2)
         IF (I .EQ. 2*IS+2) ANUM(I)=ANUM(I) + THETA(1)*CTHETA(2)
         IF (I .EQ. 2*IS+2) ADEN(I)=ADEN(I) + PHI(1)*CPHI(2)
         IF (I .EQ. 2*IS+3) ANUM(I)=ANUM(I) + THETA(2)*CTHETA(2)
         IF (I .EQ. 2*IS+3) ADEN(I)=ADEN(I) + PHI(2)*CPHI(2)
  20  CONTINUE
C
C
      RETURN
      END
```

# DISTRIBUTION

No. of
Copies

## OFFSITE

Dr. Yvonne M.M. Bishop
Director, Office of Statistical Standards
Energy Information Administration
Forrestal Building
1000 Independence Avenue, N.W.
Washington, D.C. 20585

Dr. Nancy J. Kirkendall
Office of Statistical Standards
Energy Information Administration
Forrestal Building
1000 Independence Avenue, N.W.
Washington, D.C. 20585

Dr. Turkan K. Gardenier
TKG Consultants, Ltd.
1309 Fourth Street, S.W.
Washington, D.C. 20024

27 DOE Technical Information Center

## ONSITE

### DOE Richland Operations Office

H.E. Ransom

### 12 Pacific Northwest Laboratory

D.L. Hall (3)
J.A. Mahaffey
A.L. Slavich
Publishing Coordination (2)
Technical Information (5)