1986

Interrupted Time Series Analysis: Computer Programs and Accuracy of Intervention Parameter Estimates

John W. Harrop
University of Rhode Island

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Recommended Citation

Harrop, John W., "Interrupted Time Series Analysis: Computer Programs and Accuracy of Intervention Parameter Estimates" (1986). *Open Access Dissertations*. Paper 939.
https://digitalcommons.uri.edu/oa_diss/939

This Dissertation is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

INTERRUPTED TIME SERIES ANALYSIS:
COMPUTER PROGRAMS AND ACCURACY OF
INTERVENTION PARAMETER ESTIMATES

BY

JOHN W. HARROP

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

1986

ABSTRACT

Five computer programs that perform Box and Jenkins (1970) ARIMA model interrupted time series analysis, i.e., TSX, GENTS, SAS, BMDP, and ITSE, were evaluated both qualitatively and quantitatively. The qualitative evaluation reviewed their documentation and computational features. The quantitative evaluation was performed by comparing analysis results (estimates of the minimum residual error variance and the intervention parameters) for 10 replications each of 44 simulated time series. The 44 series include 11 different ARIMA models and degrees of dependency; by two forms of intervention, i.e., 1) level (L) and change in level (DL) alone, and 2) level (L), change in level (DL), slope (S), and change in slope (DS); and by two series lengths (40 and 100 points). Major findings are 1) differencing and some autoregressive model solutions are highly inaccurate even with the correct model identification, 2) GENTS, which does not require that the model be identified, performed comparably (i.e., it provided equivalent solutions, both good and bad) to the "true model" solutions, and 3) model identification appears still to be necessary. Recommendations are made for current analyses, improvements to computer programs, and future research.

ACKNOWLEDGEMENTS

Many thanks to the members of my doctoral committee - Larry Grebstein and Ed Carney - and especially to Wayne Velicer. Wayne not only provided many helpful suggestions for this study, but has also assisted, supported, cajoled, and taught me over the past eight and one half years as my major professor, teacher, and friend.

I wish also to acknowledge assistance from BMDP. They provided Fortran source code for a number of subroutines along with information necessary to find and output solution values to a disk file. This significantly reduced the amount of paper that I had to deal with.

Finally, I would like to thank my family and friends who have provided moral support over the years. In particular, many thanks to Mary Liz Twiss for hours of searching the periodicals stacks at the URI library so that I could copy journal articles; to Charlotte Denton for two hectic days of keying in number after number as I read them aloud; and to Janet DiRuzzo and Doreen Kranz for last minute, emergency typing as I verged on the edge of panic.

Thank you all and thank goodness it's over!

TABLE OF CONTENTS

	page
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vi
Introduction	1
Method	11
Simulation and Design Characteristics	11
Program Versions and Analysis Characteristics	17
Results	20
Qualitative	20
Computational Features	20
Documentation Review	30
Quantitative	33
Minimum Residual Error Variance (EMIN)	35
With only L and DL estimated	35
With L, DL, S, and DS estimated	38
Forty versus one hundred point series	41
Estimate of Preintervention Level (L)	41
With only L and DL estimated	42
With L, DL, S, and DS estimated	45
Forty versus one hundred point series	45
Estimates of L with slope versus without slope.	48

	page
Estimate of Postintervention Change in Level (DL)	48
With only L and DL estimated	49
With L, DL, S, and DS estimated	52
Forty versus one hundred point series	55
Estimates of DL with slope versus without slope	55
Estimate of Preintervention Slope(S)	56
With L, DL, S, and DS estimated	56
Forty versus one hundred point series	61
Estimate of Postintervention Change in Slope (DS)	62
With L, DL, S, and DS estimated	62
Forty versus one hundred point series	67
Computer Time Requirements	67
Discussion	71
Model Identification	71
Computer Programs	77
Recommendations	83
References	91
Appendix A: Harrop and Velicer (1985)	98
Appendix B: Simulation Program	116

LIST OF TABLES

	page
Table 1 <u>Specifications for Time Series to be Generated and</u> <u>Analyzed by TSX, GENTS, SAS, BMDP, and ITSE</u>	12
Table 2 <u>Summary of Computational Features</u>	21
Table 3 <u>Mean and Standard Deviation of Minimum Residual Error</u> <u>Variance (EMIN) for N=40, with only Level (L) and Change</u> <u>in Level (DL) Estimated</u>	36
Table 4 <u>Mean and Standard Deviation of Minimum Residual Error</u> <u>Variance (EMIN) for N=100, with only Level (L) and Change</u> <u>in Level (DL) Estimated</u>	37
Table 5 <u>Mean and Standard Deviation of Minimum Residual Error</u> <u>Variance (EMIN) for N=40, with Level (L), Change in Level</u> <u>(DL), Slope (S), and Change in Slope (DS) Estimated . .</u>	39
Table 6 <u>Mean and Standard Deviation of Minimum Residual Error</u> <u>Variance (EMIN) for N=100, with Level (L), Change in Level</u> <u>(DL), Slope (S), and Change in Slope (DS) Estimated . .</u>	40
Table 7 <u>Mean and Standard Deviation of Estimate of Level(L)</u> <u>for N=40, with only Level (L) and Change in Level (DL)</u> <u>Estimated</u>	43
Table 8 <u>Mean and Standard Deviation of Estimate of Level(L)</u> <u>for N=100, with only Level (L) and Change in Level (DL)</u> <u>Estimated</u>	44
Table 9 <u>Mean and Standard Deviation of Estimate of Level(L)</u> <u>for N=40, with Level (L), Change in Level (DL),</u> <u>Slope (S), and Change in Slope (DS) Estimated</u>	46
Table 10 <u>Mean and Standard Deviation of Estimate of Level(L)</u> <u>for N=100, with Level (L), Change in Level (DL),</u> <u>Slope (S), and Change in Slope (DS) Estimated</u>	47

Table 11	
<u>Mean and Standard Deviation of Estimate of Change in Level (DL) for N=40, with only Level (L) and Change in Level (DL) Estimated</u>	50

Table 12	
<u>Mean and Standard Deviation of Estimate of Change in Level (DL) for N=100, with only Level (L) and Change in Level (DL) Estimated</u>	51

Table 13	
<u>Mean and Standard Deviation of Estimate of Change in Level (DL) for N=40, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated . .</u>	53

Table 14	
<u>Mean and Standard Deviation of Estimate of Change in Level (DL) for N=100, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated . .</u>	54

Table 15	
<u>Mean and Standard Deviation of Estimate of Slope (S) for N=40, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated</u>	57

Table 16	
<u>Mean and Plus and Minus One Standard Deviation Values of Slope (S) for N=40, in Degrees</u>	58

Table 17	
<u>Mean and Standard Deviation of Estimate of Slope (S) for N=100, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated</u>	59

Table 18	
<u>Mean and Plus and Minus One Standard Deviation Values of Slope (S) for N=100, in Degrees</u>	60

Table 19	
<u>Mean and Standard Deviation of Estimate of Change in Slope (DS) for N=40, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated . .</u>	63

Table 20	
<u>Mean and Plus and Minus One Standard Deviation Values of Change in Slope (DS) for N=40, in Degrees</u>	64

Table 21	
<u>Mean and Standard Deviation of Estimate of Change in Slope (DS) for N=100, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated . .</u>	65
Table 22	
<u>Mean and Plus and Minus One Standard Deviation Values of Change in Slope (DS) for N=100, in Degrees</u>	66
Table 23	
<u>Representative Central Processing Unit (CPU) Times . .</u>	68

Interrupted Time Series Analysis: Computer Programs and Accuracy of Intervention Parameter Estimates

Time series, i.e., measurements made at equal time intervals on a single unit or subject, are ideally suited for answering many research questions in the social and behavioral sciences. In particular, an interrupted time series which has observations occurring both before and after an intervention allows one to make inferential statements about the effect of the intervention on the dependent measure. Some examples might be the effect(s) of a medication on blood pressure, the effect(s) of contingent reinforcement on social responses from a schizophrenic patient, and the effect(s) of timeout on the daily frequency of hitting other children. Much has been written regarding the correct design for interrupted time series experiments which will not be duplicated here (see for example, Campbell, 1969; Campbell & Stanley, 1979; Cook & Campbell, 1979; Hersen & Barlow, 1976; Kratochwill, 1978). Although appropriate statistical methods for analyzing interrupted time series have been available for some time, this methodology is not broadly used in the social and behavioral sciences. This can be attributed to 1) the complexity of performing model identification (which eliminates many researchers), 2) the minimum recommended numbers of 50 preintervention and 50 postintervention data points for

model identification (which eliminates many data sets), and 3) until recently, the lack of generally available computer programs to perform the analysis.

In the past, and to a great extent in the present, analysis of time series research data consisted of "eyeballing" (graphical analysis) and the use of traditional parametric tests such as t-tests and regression analysis. The presence of dependency or autocorrelation in the data, however, invalidates these techniques to varying degrees. Although there are still advocates of the graphical technique (Baer, 1977; Parsonson & Baer, 1978) there is often poor agreement between qualified judges when using it and the presence of dependency can affect the judgement (De Prospero & Cohen, 1979; Jones, Weinrott, & Vaught, 1978). The presence of dependency in the data results in the violation of the independence assumptions of parametric tests resulting in biased tests of significance (Padia, 1973).

In order to deal with autocorrelated time series, the Autoregressive Integrated Moving Averages (ARIMA) model was introduced (Box & Jenkins, 1970; Box & Tiao, 1965). Glass, Willson, and Gottman (1975) introduced a technique to ARIMA model analysis which involves applying a transformation to the series based on the ARIMA model in order to transform the data to an independent series. The independent series is then analyzed by the General Linear Model solving for

both dependency and intervention parameters. Box and Tiao (1975) introduced the use of transfer function representations of the intervention parameters and suggested a nonlinear numerical solution to the resulting equations. McCain and McCleary (1979), McCleary and Hay (1980), and McDowall, McCleary, Meidinger, and Hay (1980) elaborated on the transfer function approach whereas Gottman (1981) recommended the exclusive use of autoregressive models and the easier linear solution to the resulting equations. All of these approaches provide a statistical analysis that 1) accounts for the dependency in the data, 2) estimates the degree of dependency, 3) estimates the values of various forms of interventions, as specified by the analyst (e.g. a constant level prior to intervention and an abrupt and constant change in level after the intervention or, similarly, slope and change in slope), and 4) provides tests of significance for the estimated parameters.

The most difficult part of performing an ARIMA analysis is identifying the underlying (p,d,q) model where p is the order (or number) of autoregressive coefficients, d is the order of differencing required to achieve stationarity, and q is the order of the moving averages coefficients. Velicer and Harrop (1983) demonstrated that even trained analysts performed disappointingly in this regard, and the published differences of opinion on the identification of a series by Deutsch (1979), Deutsch and Alt (1977), and Hay and McCleary

(1979) further demonstrated this. These demonstrations were based on the original method proposed by Box and Jenkins (1970). That model identification still remains a problem is demonstrated by the plethora of new techniques that have been proffered (Akaike, 1974; Beguin, Grieroux & Monfort, 1980; Bhansali & Downham, 1977; Gray, Kelley & McIntire, 1978; Hannan & Rissanen, 1982; Kashyap, 1977; Parzen, 1974; Pukkila, 1982; Schwartz, 1978; Tsay, 1984; Tsay & Tiao, 1984).

In addition to being complicated, controversial, and subjective, it is recommended that model identification be performed on series with a minimum number of 50 pre-intervention (pre-I) and 50 postintervention (post-I) points (Box & Jenkins, 1970; Glass, Willson, & Gottman, 1975; Padia, 1975). These two factors preclude model identification by a large number of potential researchers and of a large number of series respectively. It is of little wonder that as late as this past year, Miller (1985) was suggesting the use of graphical techniques; i.e., split middle, two sigma, and frequency histograms in order to provide "the ability for counselors to conduct and analyze single subject (N=1) research."

As a result, Harrop and Velicer (1985, see Appendix A) evaluated two suggestions that eliminate the model identification step entirely. The AUTO-1 approach (Simonton, 1977; see also Marmor & Marmor, 1978) assumes a (1,0,0)

model. The General Transformation Approach (GENERAL; Velicer & McDonald, 1984) assumes a (5,0,0) model. Harrop and Velicer generated simulated time series data manipulating the model, degree of dependency, and length of series; and analyzed the series using the true model (p,d,q) used to generate the series, a (1,0,0), and a (3,0,0) model identification as an approximation to Velicer and McDonald's (5,0,0). A tentative recommendation to eliminate the model identification step entirely was made based on the fact that the GENERAL approach provided estimates of the minimum residual error variance (EMIN), level (L), and change in level (DL) across the models, dependencies, and lengths of series that were generally as accurate as those from the true model analysis. The AUTO-1 approach fared almost equally as well, but did have some difficulty with (2,0,0) series. Of further note is that results for shorter series of 40 points were generally as good as the results for series of 100 points and that results for differencing and the (2,0,0) .8/.2 (i.e., $\phi_1 = .8$ and $\phi_2 = .2$) models were extremely poor. By way of explanation for the latter it was shown that the equations for the (2,0,0) .8/.2 and (0,1,0) models are quite similar.

Harrop and Velicer's (1985) conclusion to eliminate the model identification step entirely was tentative because they felt that additional factors needed to be studied. Among these were 1) the effects of changes in level both

smaller and larger than the standard deviation of the random component of the model; 2) the accuracy of the techniques in estimating slope and change in slope intervention parameters; 3) the accuracy of the GENERAL approach using the recommended (5,0,0) identification; and 4) the accuracy of results using a nonlinear numerical algorithm to determine EMIN as opposed to the grid search algorithm used by the program in their study.

Finally, in addition to the problems of model identification, another drawback to more widespread utilization of the statistical analysis of interrupted time series has been the lack of generally available computer programs. In the past six years or so, a large number of time series programs have become available to varying degrees. These programs may be categorized variously as mainframe computer (or minicomputer) versus microcomputer versions, run in batch mode versus run interactively, available for free (or for a minimum handling charge) versus requiring a purchase or rental fee, and including interrupted time series analysis versus not including the analysis of interrupted time series. Five representative programs from those currently available are TSX (Bower & Glass, 1974), GENTS (Velicer, Fraser, McDonald, & Harrop, 1980), BMDP (the 2T program; Dixon, 1985), SAS (Proc ARIMA; SAS Institute, 1984), and ITSE (from the Gottman programs; Williams & Gottman, 1982). Each of these programs run in batch mode on a

mainframe computer and perform analysis of interrupted time series.

TSX (Bower & Glass, 1974) was used by Harrop and Velicer (1985) and, although it has been available for some years, users have had to request a copy from the authors and have had to be at least partly familiar with Fortran programming in order to implement it. Additionally, unless analysts were familiar with Glass, Willson, and Gottman (1975), they would be unlikely to be aware of it's existence. GENTS (Velicer et al., 1980) is more recent but is in the same category as TSX except that analysts would need to be familiar with Harrop and Velicer (1985) or Velicer and McDonald (1984) to be aware of it's existence. BMDP (Dixon, 1985) and SAS (SAS Institute, 1984) are both commonly available at most Academic Computer Centers (ACC) along with manuals and support from ACC personnel, but they are not oriented towards the analysis of interrupted time series (at least not by most social and behavioral scientists). ITSE (Williams & Gottman, 1982) is available for a fee from the authors and also requires a partial familiarity with Fortran. Analysts would be unlikely to be aware of its existence without familiarity with Gottman (1981).

This brings us to the purposes of this study, i.e., first, to further investigate the necessity for model identification, and second, to evaluate a number of computer programs used for interrupted time series analysis. Regarding

the first purpose, the intent was to build on the results of Harrop and Velicer (1985), and either to provide further support for eliminating model identification or to refute the conclusion to do so. Elimination of model identification would dramatically increase the number of researchers capable of evaluating their research with ARIMA time series analysis, thereby complementing graphical analysis with a statistical "yardstick". Additionally, if results so indicated, many more "shorter" series would become appropriate for analysis. Regarding the second purpose, the intent was to evaluate several programs from various perspectives in order to provide some guidance to analysts in choosing among them.

In order to further investigate the necessity for model identification, simulated time series were input to TSX along with the true model identification. The results of these analyses were compared to the results of solutions by GENTS using the same simulated time series and a (5,0,0) identification as recommended by the Velicer and McDonald (1984) GENERAL approach. A total of 44 series, 10 samples of each, were generated representing various models, degrees of dependency, lengths of series, and types of intervention. Following the suggestions of Harrop and Velicer (1985), changes in level both smaller and larger than the standard deviation of the random component were included, as well as interventions including slope and change in slope. Addi-

tionally, the use of GENTS satisfied two more of their suggestions, i.e., the use of the recommended (5,0,0) identification and the use of a nonlinear numerical algorithm to determine the minimum residual error variance (EMIN) solution. The estimates of EMIN and the intervention parameters provided by TSX were compared to the simulation values (which are synonymous with true or criterion values) in order to determine whether or not (and if not under what conditions) these estimates were accurate. The estimates of the intervention parameters provided by GENTS were compared to the simulation values as well as to the TSX estimates in order to determine whether or not these estimates were sufficiently accurate to allow elimination of model identification. Regarding this latter, the length of series manipulation, i.e., 40 and 100 point series, was of particular interest.

Finally, five computer programs were evaluated quantitatively and qualitatively in order to provide guidance to analysts in choosing among them. The five programs are TSX, GENTS, SAS, BMDP, and ITSE. TSX was chosen because it is the tried and true program used in many studies. Its primary drawback is that it is slow (uses a lot of computer time) due to its use of a grid search algorithm as explained later. BMDP and SAS were selected because of their common availability, support, and their speed. They both use speedy, nonlinear numerical algorithms for their

solutions. GENTS was selected because it directly implements the GENERAL approach and because it uses a different nonlinear numerical algorithm than either SAS or BMDP. ITSE was selected because the estimation is for autoregressive models and, therefore, uses a linear algorithm different than any of the other four. The quantitative evaluation of these five programs consisted of comparing their solutions for the previously mentioned 44 series to the simulation values and to each other, as well as determining the amount of computer central processing unit (CPU) time required to obtain solutions. In essence, the accuracy or relative accuracy of their algorithms were evaluated along with their computational efficiency, across the various conditions of the 44 series. The qualitative evaluation investigated the documentation and output of the programs in order to address such issues as complexity, useability, inclusiveness (computational features) and flexibility.

Method

This section is divided into two subsections. The Simulation and Design Characteristics subsection presents the characteristics of the general simulation and the specific simulated series analyzed by the five computer programs. The Program Version and Analysis Characteristics subsection specifies the programs used and analysis options specific to each program.

Simulation and Design Characteristics

Forty-four simulated time series were input to TSX, GENTS, SAS, BMDP, and ITSE for analysis. All series followed the Autoregressive Integrated Moving Averages (ARIMA) model and were generated by a Fortran computer program (see Appendix B) adapted and revised by the author from a program by Padia (1975). The random number generator used was the GGNML subroutine of the International Mathematical and Statistical Library (IMSL), Revision 9.2. The mean of the random component of all series was zero and the variance was one.

The specific characteristics of the simulated series are listed in Table 1. Forty-four types of series representing the 11 combinations of models and dependency presented, by the two types of intervention presented, by two series lengths (N), i.e., 40 and 100 points, were used. All interventions occurred at the midpoint, i.e., at time 21

Table 1

Specifications for Time Series to be Generated and
Analyzed by TSX, GENTS, SAS, BMDP, and ITSE

ARIMA Model (p,d,q)	Phi(s) or Theta(s)	Level	Change in Level	Slope	Change in Slope
0,0,0	none	0	.5s	0	0
0,0,0	none	0	1.5s	15deg	-15deg
0,1,0	none	0	.5s	0	0
0,1,0	none	0	1.5s	15deg	-15deg
0,0,1	.8	0	.5s	0	0
0,0,1	.8	0	1.5s	15deg	-15deg
0,0,1	.4	0	.5s	0	0
0,0,1	.4	0	1.5s	15deg	-15deg
1,0,0	.8	0	.5s	0	0
1,0,0	.8	0	1.5s	15deg	-15deg
1,0,0	.4	0	.5s	0	0
1,0,0	.4	0	1.5s	15deg	-15deg
0,1,1	.4	0	.5s	0	0
0,1,1	.4	0	1.5s	15deg	-15deg
2,0,0	.4/.4	0	.5s	0	0
2,0,0	.4/.4	0	1.5s	15deg	-15deg
2,0,0	.8/.2	0	.5s	0	0
2,0,0	.8/.2	0	1.5s	15deg	-15deg
2,0,0	-.8/-.4	0	.5s	0	0
2,0,0	-.8/-.4	0	1.5s	15deg	-15deg
2,0,0	.8/-.4	0	.5s	0	0
2,0,0	.8/-.4	0	1.5s	15deg	-15deg

Note: For each of the 11 ARIMA (p,d,q) models and dependencies (phis or thetas) by 2 types of interventions shown, series with 40 and 100 points were simulated. Ten replications of each of the resulting 44 series were generated.

for N=40 series and at time 51 for N=100 series. Ten samples or replications of each of the resulting 44 series were generated and then analyzed by each of the five computer programs. Ten replications were chosen based on the results of some preliminary runs with 5, 10, and 20 replications. Results indicated that estimates were more accurate and stable with 10 and 20 replications than with 5, and that there was little or no improvement with 20 replications as opposed to 10.

The current study generally utilized the same models and dependencies for generating the data as Harrop and Velicer (1985). For this study, however, interventions with slope and change in slope were added in addition to level and change in level. Additionally, two different sizes of change in level were added representing .5 and 1.5 times the standard deviation (s) of the random error component of the models. Harrop and Velicer (1985) investigated only a single change in level, i.e., one times s.

The ARIMA (p,d,q) models chosen were reported by Glass et al. in 80 of 95 cases in an empirical study (1975, pp. 115-118) and thus are representative of models that one may expect to encounter in the social and behavioral sciences. Additionally, McDowall, McCleary, Meidinger, and Hay state "If our experiences are typical, most of the time series that social scientists encounter will be well represented by

zero- or first-order models such as ARIMA (0,0,0), ARIMA (1,0,0), ARIMA (0,0,1) and ARIMA (0,1,1)" (1980, p. 54). Of 98 series from a study of couple interaction during marital counseling, Revenstorf, Kessler, Schindler, Hahlweg, & Bluemner (1980) found that all of them could be represented by either (0,0,0) or (1,0,0) models. Finally, in a reanalysis of 70 series from "five practice journals that publish research on the effects of interventions monitored over time," Marsh and Shibano (1984, p. 9) identified 40% as (0,0,0), 34% as (1,0,0), and 14% as (2,0,0). They also identified 11% as (3,0,0) and one series as (5,0,0), which are not investigated in this study. These (3,0,0) and (5,0,0) series were probably identified because, following Gottmann (1981), Marsh and Shibano use only autoregressive representations of time series. There is a strong possibility that these series would be more parsimoniously represented by low order moving averages models of the type that are investigated in this study.

The dependency values (phis or thetas) were chosen to be representative of high and low autocorrelations within the invertibility/stationarity limits. These are a subset of the values used by Harrop and Velicer (1985). Negative dependency values for the (0,0,1), (1,0,0), and (0,1,1) models were deleted along with eight of the (2,0,0) dependency values. The current (2,0,0) values were chosen to

represent values that both increase variance (negative dependency) and decrease variance (positive dependency); and that have an additive effect (both values are the same sign) and a destructive or canceling effect (opposite signs). Harrop and Velicer found relatively little difference in solutions for their 12 values of dependency for the (2,0,0) models with the exception of .8/.2 which were retained for this study. It should be noted that the convention adapted when identifying a particular model and dependency in the text is to specify the (p,d,q) of the model first, and to follow it with the degree of dependency. In the case of first order autoregressive and moving averages models, the single dependency value is for ϕ_1 and θ_1 respectively. In the case of second order autoregressive models, ϕ_1 and ϕ_2 are presented in that order and separated by a slash, e.g., (2,0,0) .8/.2.

Two sets of intervention parameters were chosen, i.e., level and change in level alone and level and change in level in combination with slope and change in slope. Interventions of these types, are those that are most typically found, or at least, hypothesized and tested for in the behavioral sciences. Harrop and Velicer investigated series with an intervention effect of one times $\underline{s}$ in order to test "the sensitivity of the three model identification approaches for detecting a moderate intervention effect"

(1985, p. 33). The current study investigated the effects of smaller (.5s) and larger (1.5s) change in level interventions.

The values 15 and -15 degrees were selected for slope and change in slope. These values were selected because they represent relatively small values that would probably be difficult to discriminate by eye, an instance in which statistical analysis would be most important. Additionally, a more pronounced slope would be problematic with the 100 point series as compared to the change in level, i.e., 1.5. The chosen value represents a rise in the preintervention series of approximately 5.4 for 40 point series, and 13.5 for 100 point series. The chosen combination of "level" and "slope" , i.e., 0.0s/1.5s with 15/-15 degrees, was selected, therefore, as being a reasonable first look at the interaction effect of level and slope interventions.

Finally, two lengths of series (N) were generated, i.e., 40 and 100 points (20 and 50 points pre- and post-intervention respectively). These are the same series lengths as those used by Harrop and Velicer (1985). They were selected to represent long (for the behavioral sciences) and short series. The long series represent the shortest series recommended by various authors (Box & Jenkins, 1970; Glass, Willson, & Gottman, 1975; Gottman & Glass, 1978; Padia, 1975) as being necessary for accurate

identification of the model. The short series is representative of a more realistically attainable goal for studies in the behavioral sciences.

Program Versions and Analysis Characteristics

The resultant 44 series were analyzed by the five computer programs using either the true or (5,0,0) model identifications. TSX (Bower & Glass, 1974) was available from the authors of the Glass, Willson, and Gottman (1975) text as Fortran source code. Header information notes that it was copyrighted in 1974. TSX uses a grid search approach (see explanation under Computational Features) to the solution of the ARIMA model equations. For this study, the true model identification was used for all series, and an iteration increment of .02 was used for the grid search for all except the (2,0,0) models. The (2,0,0) models used an iteration increment of .10 in order to reduce computer time. Solutions estimating two or four parameters were specified as appropriate to the input series.

GENTS (Velicer, Fraser, McDonald, & Harrop, 1980) is available from the first author as Fortran source code. It uses a nonlinear "quasi-Newton minimization" algorithm to find a solution to a general transformation matrix. This is conceptually similar to applying a (5,0,0) model identification to all cases. The default value of 100 function iterations, maximum, was used in this study as were default values for starting estimates and the stopping criterion.

SAS (SAS Institute, 1984) is available as an applications package at many University Academic Computer Centers. Specifically, Proc ARIMA, version 5, was used in this study. SAS uses "Marquardt's method with step halving" to solve the nonlinear set of equations (SAS Institute, 1984; p. 139). Analyses for this study used the true model identification, default values for the starting estimates and stopping criterion, and the conditional least squares (CLS) method of estimation. The maximum number of iterations to be performed was increased to 50 because estimations for some models terminated before reaching the stopping criterion when the default value was used.

BMDP (Dixon, 1985) is also available as an applications package at many University Academic Computer Centers. Specifically, the BMDP-2T program, program version May 31, 1986 was used for this study. As with SAS, the true model identification, default values for the starting estimates and stopping criterion, the CLS estimation, and a maximum number of iterations to be performed equal to 50 were used.

ITSE (Williams & Gottman, 1982) is available from the authors as Fortran source code. ITSE solves only for autoregressive models and, therefore, performs a linear least squares solution. The model identifications used for ITSE were the true model for the (0,0,0) or white noise model and the autoregressive models. Differencing models were not used with ITSE because this would have necessitated the use

of another program to difference the series before the analysis. The remaining moving averages series were analyzed by ITSE using a (5,0,0) model identification as an approximation to the infinite series autoregressive model by which a moving averages series may be represented.

Results

The Qualitative review is presented first and is followed by the Quantitative results. The Qualitative section is comprised of two subsections, i.e., Computational Features and Documentation Review. The Quantitative section has a subsection for each of the five parameters estimated and a subsection for CPU requirements.

Qualitative

It should be noted that the comments that follow deal only with the mainframe computer version of the five computer programs, run in batch mode. Further, this study only addressed intervention analysis of interrupted time series with no seasonal components.

Computational Features

Each of the five programs were reviewed for the presence or absence of features appropriate to the analysis of interrupted time series. A summary of the review is presented in Table 2. First, terms necessary for an understanding of Table 2 and the following presentation of each program's features are explained.

Input Model specifies the type of ARIMA model that may be input to the programs. Intervention Parameters: Input refers to the types of intervention and how they may be input. Two (2) parameters signifies an L and DL intervention, four (4) parameters signifies an L, DL, S, and DS intervention, and X refers to the use of a design matrix to

Table 2

Summary of Computational Features

COMPUTATIONAL FEATURES	TSX (p,d,q)	GENTS (5,0,0) fixed	SAS (p,d,q)	BMDP (p,d,q)	ITSE (p,0,0)
Input Model					
Intervention Parameters					
Input	2,4, or $\underline{X}$	2,4, or $\underline{X}$	$\underline{X}$	$\underline{X}$	fixed
t-Statistics	yes	no	yes	yes	yes
Dependency Parameters					
Output	yes	yes	yes	yes	yes
t-Statistics	no	no	yes	yes	yes
Output EMIN	yes	no	yes	yes	yes
Save Residuals	yes	no	yes	yes	no
Tests for Residuals	no	no	yes	yes	no
Type of Solution	grid search	nonlinear	nonlinear	nonlinear	linear
Initial Values	yes	yes	yes	yes	NA
Stopping Criterion	NA	yes	yes	yes	NA
Max. N of Iterations	NA	yes	yes	yes	NA
Nonlinear Diagnostics	NA	yes	yes	yes	NA
Observe Stat./Invert.	yes	no	optional	no	unknown

specify various intervention forms (including any combination of L, DL, S, and DS occurring at times specified by the user). Intervention Parameters: t-Statistics specifies whether or not tests of significance are provided for the estimates of intervention parameters. Dependency Parameters: Output and Dependency Parameters: t-Statistics refer similarly to whether or not a program provides a printout of the estimates of the dependency values, and provides a test of statistical significance for these values, respectively.

Output EMIN specifies whether or not the program outputs the residual error variance of the solution obtained. All solutions of the ARIMA model equations attempt to minimize this variance in various ways, and this value is helpful in comparing solutions. Save Residuals and Tests for Residuals refer to whether or not a program allows the residual series (the series left over after the dependency effect and intervention parameters are removed) to be saved for diagnostic checking, and whether or not there are tests of the residuals that are performed without having to specify them, respectively.

Type of Solution distinguishes between a grid search solution, a linear solution, and a nonlinear numerical solution to the ARIMA equation. The grid search method is described below with the presentation of TSX's features. ITSE's linear solution requires neither a grid search

approach nor a nonlinear minimization algorithm because it does not solve for ARIMA models with a moving averages process.

The nonlinear solutions use various numerical minimization methods in order to arrive quickly at a solution for the ARIMA equations that minimizes a numerical criterion (which is intended ultimately to minimize the residual error variance). Nonlinear techniques require Initial Values (also called starting values) of the parameters to be estimated, and a Stopping Criterion in terms of some value that is calculated by the algorithm. Two of the difficulties with nonlinear algorithms are that they may not converge, i.e., they may oscillate around a correct solution or around various solutions; or they might find a relative or local minimum for the residual error variance. Initial values typically have default values built into the program that have been chosen to optimize the possibility of finding the correct solution in the least time possible. However, it is helpful to be able to provide selected values in the event that one feels that the analysis is converging on an incorrect solution.

Stopping criterion typically have default values as well. In essence a stopping criterion provides a way to specify the resolution or granularity to which the minimization process is to be taken. Max. N of Iterations refers to whether or not the user may specify a number of

iterations of the nonlinear algorithm that will not be exceeded. Nonlinear estimations are stopped when either the maximum number of iterations or the stopping criterion has been reached. They are normally used in combination with each other in order to prevent computer time from being wasted. For example, if the stopping criterion desired can be achieved with 10 iterations, it is wasteful to compute 20. Further, if 20 iterations have not resulted in achieving the stopping criterion, more should be performed. Additionally, however, an oscillating solution would continue forever (or until the computer was shut off) if a maximum number of iterations to be performed check was not included in the program. Finally, as can be seen from the complexity of nonlinear solutions, it is helpful if Nonlinear Diagnostics are provided for evaluating this process.

The last term to be explained is Observe Stat./Invert., i.e., observes stationarity/invertibility conditions. This refers to whether or not the program restricts a solution to values of dependency that meet these conditions. Technically, ARIMA models are only valid when this is done.

Generally speaking, all of the features contained in Table 2 are desirable except those indicated as NA, i.e., not applicable. Minimal input capabilities required are for the general ARIMA (p,d,q) model (or a specific subset as for GENTS and ITSE), and for various forms of intervention.

Minimal outputs required are the intervention parameter estimates along with test statistics, the minimum residual error variance (EMIN), and the residual series. These inputs and outputs allow an analyst to specify the desired analysis, test the significance of the intervention statistically, and evaluate the suitability of the model specified. For the nonlinear estimations, the ability to optionally specify initial values, a stopping criterion, and the maximum number of iterations are highly desirable, as is program output of nonlinear diagnostics.

As regards observing the stationarity/invertibility conditions, it is preferable if the program provides the option to observe or not observe them. The program's default, however, should be to observe them, given that the ARIMA equations are valid only when the stationarity/invertibility conditions are met. Not observing them, therefore, should be a conscious choice. Further, if the program does not provide the option to observe or not observe them, it is preferable, in the opinion of the author, that the program observe them.

It is desirable for a program to have optional and/or nonoptional tests of the residual series included in order to facilitate the analysis process (this reduces the number of separate computer runs required). This is not a requirement, however, given that the residual series, if saved, may be evaluated with one of the other programs that

provides the appropriate features. To date, the author has found little use for the estimates of dependency parameters and their test statistics. A suggestion for possibly using these values made in the Recommendations section, however, would make them necessary outputs.

TSX (Bower & Glass, 1974) uses a grid search approach to a least squares solution of the intervention parameters. In the grid search approach, the TSX program computes a solution to the ARIMA model equations for each of a number of dependency (ϕ and/or θ) values. The dependency value(s) are calculated by incrementing or decrementing the previous value by an iteration increment specified by the user (or by the default value of .02). The dependency values used are kept within the stationarity/ invertibility region. For a (1,0,0) model with an iteration increment of .02, for example, TSX would calculate 99 solutions with values of $\phi = .98, .96, .94, \dots, .02, 0.0, -.02, \dots, -.98$, corresponding to the stationarity/invertibility condition $-1.0 < \phi < 1.0$. The solution with the smallest residual error variance is then selected as being the correct one. The ARIMA model (p,d,q) is provided as an input to TSX by the user. Further, the user may specify either of the two forms of intervention used in this study and an appropriate design matrix will be generated, or alternately, may provide a design matrix which tests other forms of intervention. The minimum residual error variance

is computed and output along with the value(s) of dependency at which it was found. Also computed and output are the intervention parameters and their respective t-statistics. It should be noted that there is no test of significance for the dependency. Additionally, since the dependency is not estimated per se, the degrees of freedom for the tests of significance of the intervention parameters are not reduced by the number of dependency parameters. Finally, although the residual series may be output for diagnostic testing, TSX does not provide any diagnostic tests itself.

GENTS (Velicer, Fraser, McDonald, & Harrop, 1980) uses a nonlinear "quasi-Newton minimization" algorithm to find a solution to a (5,0,0) model. Solutions are not restricted to dependency values that meet the stationarity/invertibility conditions. The maximum number of function evaluations to be performed and starting values for the five dependency parameters may be provided to the program. Additionally, the user may specify that intermediate solutions be output for use in evaluating the minimization process. Estimates of the dependency and intervention parameters are computed and output. GENTS does not currently provide the minimum residual error variance, t-statistics for the estimated parameters, the residual series, nor diagnostic tests of the residual series. The stopping criterion for the nonlinear solution is fixed, i.e., it can not be specified by the user without modifying

the value in the Fortran program.

SAS uses "Marquardt's method with step halving" (SAS Institute, 1984; p. 139) to solve the nonlinear system of equations and to estimate dependency and intervention parameters of user specified (p,d,q) models. The user may specify whether or not dependency estimates are to be restricted to meeting the stationarity/invertibility conditions. The maximum number of iterations to be performed, starting values for the intervention parameters, and a stopping or convergence criterion in terms of changes in the estimates may be specified. Additionally, a grid printout of the error sum of squares (or log-likelihood surface) in the parameter space around the parameter estimates may be requested and used in evaluating the nonlinear solution. Conditional least squares (CLS), unconditional least squares (ULS), and maximum likelihood (ML) solutions are available. Except for a constant (level for nondifferenced models or slope for first differenced models), the user must specify a design matrix along with appropriate transfer function polynomials for the intervention parameters. Estimates of the dependency and intervention parameters are computed and output along with t-statistics. Two information criteria, i.e., Akaike's information criterion (AIC; Akaike, 1974) and Schwartz's Bayesian criterion (SBC; Schwartz, 1978) are also computed and output for use in evaluating the adequacy of the model. Additionally, autocorrelations, partial auto-

correlations, and inverse autocorrelations of the residual series may be printed and/or plotted for the same purpose along with the Ljung and Box (1978) test for white noise.

BMDP uses a "Gauss-Marquardt" algorithm (Dixon, 1985; p. 691) and there are no stationarity/invertibility restrictions on the solution. The maximum number of iterations to be performed, starting values for the dependency and/or intervention parameters, and stopping values in terms of the change in the residual sum of squares and/or parameter estimates may all be specified. For use in evaluating the nonlinear estimation, parameter estimates may optionally be printed at each iteration. Conditional least squares (CLS) and unconditional least squares (UCS, also known as backcasting) estimations may be performed. As in SAS, a user must specify a design matrix and transfer function polynomials for all intervention parameters except for a constant. Dependency and intervention parameter estimates along with t-statistics are computed and output for the specified (p,d,q) model and independent variables (intervention parameters). The residual series may be evaluated by the use of printed and plotted autocorrelations and partial autocorrelations, and the Ljung and Box (1978) test for white noise.

Because it estimates only autoregressive (AR) models, ITSE (Williams & Gottman, 1982) uses a linear solution. The user must only specify the AR order. A separate program

ARFIT (Williams & Gottman, 1982), is provided for determining the appropriate order. Differencing, if necessary, must also be accomplished by a separate program prior to analysis by ITSE. ITSE performs several estimates, none of which may be selected or deleted by the user, but the ones of interest here are the estimates of the dependency parameters and pre- and postintervention intercepts and slopes. These estimates are provided along with their t-statistics. The residual series is not available for diagnosing the adequacy of the model.

Documentation Review

The documentation for all of the programs leaves much to be desired, for the knowledgeable computer user as well as the novice. The minimal requirement for documentation is a manual that clearly explains how to specify the input required by the program, options that are available, the output provided, and warning and error messages. Illustrative examples are a necessary part of this exposition. Beyond this minimal requirement, it is most desirable for this manual to be reasonably self-contained, i.e., to require little or no reference to other literature (either textbooks or periodicals) in order to perform analyses. This, of course, assumes that the user understands the foundations, concepts, and suitability of the analytic technique. Finally, to the extent feasible, the equations/algorithms implemented by the computer program and

broad strategies for complicated procedures or processes would be most helpful. The latter applies, for example, to strategies for model identification, for deciding on which intervention type to be estimated, and for using a nonlinear estimation program (this last item applies to GENTS, SAS, and BMDP only).

As regards the above criteria, both TSX and GENTS fall far short. Neither TSX nor GENTS have a manual nor any supporting documentation per se. A user must interpret the "comments" at the beginning of the Fortran programs in order to determine the proper inputs and there is no explanation anywhere of the output. Although someone familiar with Glass, Willson, and Gottman (1975) or Velicer and McDonald (1984), respectively, would probably have no difficulty in interpreting their output, becoming familiar is no easy task. These programs are not difficult to provide input to, however, with a brief explanation, given that the user only wishes to test for either of the two forms of intervention used in this study. Both programs require that the user only specify which of these two is to be tested for, and the programs then automatically generate the necessary design matrix. Specifying the design matrix for other forms of intervention is quite difficult without guidance, however. Further, there is no guidance available regarding specifying initial values, specifying the stopping criterion for GENTS, diagnosing GENTS nonlinear estimation process, or diagnosing

the adequacy of the model.

SAS and BMDP both have extensive manuals but their major failing is that examples and explanations are geared more toward econometricians than toward behavioral scientists. Intervention analysis, for example, requires that the user provide a design matrix for the intervention parameters (for all forms of intervention) and this is neither explained nor demonstrated. This is further complicated by the fact that the transfer function representation of intervention parameters is used in these programs. Although the transfer function representation is very powerful in allowing a broad range of complicated interventions to be specified, it requires specifying upper and lower polynomials for each intervention parameter which adds greatly to the confusion. Further, this method of representation is totally unnecessary for the simple interventions of the type used in this study, i.e., those of most interest to behavioral scientists. This area, in particular, most requires simplification and further explication in the SAS and BMDP manuals. As regards their nonlinear estimation, SAS provides some warnings of potential pitfalls but BMDP does not. Neither provide a strategy for diagnosing this process nor do they explain how one might use the diagnostic output that is optionally available for this purpose. SAS provides insufficient elucidation of the equations solved. BMDP comes the closest

to providing a reasonable and proper presentation of the equations used but falls down specifically when dealing with the intervention estimates. The reader is referred at this point in their documentation to Box and Jenkins (1970), which does not address interventions. Finally, with the exceptions noted above, both program manuals adequately explain the inputs required, other options such as saving residuals, and the program output to be expected, quite well.

ITSE's user's guide is the clearest and most complete. It provides an explanation of the input and output along with a complete sample input and output. Additionally, the equations used by the program are presented which provides a more detailed explanation or definition of the inputs and outputs. It should be noted that the documentation for ITSE's companion program ARFIT, which is used for identifying the autoregressive model to be used by ITSE, was not reviewed.

Quantitative

All of the tables presenting the quantitative review of analysis estimates were prepared using the identical format, i.e., a matrix of results for model and dependency by computer program. In order to make comparison between tables easier for the reader, this format was followed even when a row or column might have been deleted because it was not applicable. Missing cells are explained as they are

encountered.

Individual tables are presented for each parameter's estimate, i.e., for estimates of the minimum residual error variance (EMIN), level (L), change in level (DL), slope (S), and change in slope (DS); by series length, i.e., for length of series (N) equals 40 and 100 points; and by the number of parameters estimated by the analysis, i.e., with only L and DL estimated and with L, DL, S, and DS estimated. Each entry in the tables presents the mean and standard deviation (S.D.) across the 10 replications for the model and degree of dependency (DEP.) indicated. In addition, the first row of each table presents the results of analyses of the "Random Numbers". The random numbers are the random error component of the simulated series before the ARIMA model is applied and the intervention effects are added.

As stated earlier, although the Gottman programs provide a standalone program for differencing series, ITSE does not provide for differencing. It was not used, therefore, with the differencing models (0,1,0) and (0,1,1), and the corresponding locations in the following tables are blank.

Both the tables and the text refer to either the "population" or the "criterion" value. These terms apply to the value used in the simulation program to generate the time series, and are, therefore, the "true" values. The term population value was used for the minimum residual

error variance because 10 samples were drawn for the random error component from a random number generator that provided numbers that were normal and independently distributed with a mean of zero and a variance of one, i.e., NID (0,1²). The term criterion value was used for the remaining estimates because these were the actual numbers used, i.e., they were not random numbers.

Minimum Residual Error Variance (EMIN)

Tables 3 through 6 present the means and standard deviations (S.D.) across 10 replications (except as noted) of the minimum residual error variance found by TSX, SAS, BMDP, and ITSE. GENTS currently does not provide this output. In all cases EMIN represents the residual sum of squares divided by the number of residual data points, i.e., 40 for 40 point series and 100 for 100 point series except for the differencing models where 39 and 99 were used respectively. The population value of EMIN is 1.00.

With only L and DL estimated. Tables 3 and 4 present the estimates of EMIN from the analyses with only L and DL estimated for series with 40 and 100 points respectively. ITSE estimates were consistently lower than estimates by TSX, SAS, and BMDP, and the population value of 1.00. This may have been because ITSE always fits four intervention parameters whereas the other programs fit two. ITSE, therefore, may be capitalizing on chance. TSX, SAS, and BMDP recovered the simulated value reasonably well. The

Table 3

Mean and Standard Deviation of Minimum Residual Error
Variance (EMIN) for N=40, with only Level (L) and Change
in Level (DL) Estimated

Population Value = 1.00

MODEL	DEP.		TSX	GENTS ^a	SAS	BMDP	ITSE
Random Numbers		mean	.99	-	.99	.99	.95
		S.D.	.22	-	.22	.22	.22
000	none	mean	.99	-	.99	.99	.95
		S.D.	.22	-	.22	.22	.22
010	none	mean	1.03	-	1.03	1.03	-
		S.D.	.23	-	.23	.23	-
001	.8	mean	.96	-	.97	5.08	.77
		S.D.	.22	-	.23	13.16	.14
001	.4	mean	.95	-	.95	1.82	.68
		S.D.	.19	-	.19	1.82	.16
100	.8	mean	1.01	-	1.01	.95	.83
		S.D.	.24	-	.24	.23	.22
100	.4	mean	.98	-	.98	.96	.90
		S.D.	.21	-	.21	.21	.20
011	.4	mean	1.01	-	1.02	1.02	-
		S.D.	.23	-	.23	.23	-
200	.4/.4	mean	.96	-	.96	.89	.75
		S.D.	.24	-	.24	.24	.23
200	.8/.2	mean	1.03	-	.97	.86	.66
		S.D.	.24	-	.23	.24	.18
200	-.8/-.4	mean	.97	-	.97	.92	.88
		S.D.	.21	-	.21	.22	.23
200	.8/-.4	mean	.96	-	.96	.90	.83
		S.D.	.24	-	.23	.22	.22

^aCurrently does not provide an estimate.

Table 4

Mean and Standard Deviation of Minimum Residual Error
Variance (EMIN) for N=100, with only Level (L) and Change
in Level (DL) Estimated

Population Value = 1.00

MODEL	DEP.		TSX	GENTS ^a	SAS	BMDP	ITSE
Random Numbers		mean	.94	-	.94	.94	.91
		S.D.	.22	-	.22	.22	.23
000	none	mean	.94	-	.94	.94	.91
		S.D.	.22	-	.22	.22	.23
010	none	mean	.96	-	.96	.96	-
		S.D.	.23	-	.23	.23	-
001	.8	mean	.94	-	.94	.95	.86
		S.D.	.23	-	.23	.22	.21
001	.4	mean	.94	-	1.01	.94	.81
		S.D.	.23	-	.35	.23	.19
100	.8	mean	.95	-	.95	.93	.90
		S.D.	.22	-	.22	.22	.22
100	.4	mean	.94	-	.94	.93	.91
		S.D.	.22	-	.22	.22	.22
011	.4	mean	.96	-	.96	.96	-
		S.D.	.22	-	.22	.22	-
200	.4/.4	mean	.93	-	.93	.90	.86
		S.D.	.22	-	.22	.21	.21
200	.8/.2	mean	1.06	-	.94	.87 ^b	.82
		S.D.	.26	-	.22	.16	.20
200	-.8/-.4	mean	.93	-	.93	.90	.88
		S.D.	.22	-	.22	.20	.21
200	.8/-.4	mean	.93	-	.93	.91	.88
		S.D.	.23	-	.23	.22	.21

^aCurrently does not provide an estimate.

^bBased on six solutions.

exception was the mean estimates for the 40 point series (0,0,1) .8 and (0,0,1) .4 solutions by BMDP. The former was caused by one outlier value of 42 and the latter by two outlier values of 6 and 3. All three provided identical results for the white noise model and BMDP tended to find a slightly smaller value than TSX and SAS, and the population value. This was more pronounced for 40 point series. No particular tendencies regarding model were evident. In general, variability was about the same across programs and models.

With L, DL, S, and DS estimated. Tables 5 and 6 present the estimates of EMIN from the analyses with L, DL, S, and DS estimated for series with 40 and 100 points respectively. As noted, BMDP failed to find a solution in one case and found only one solution in three other cases. ITSE's mean estimates were generally slightly better than for series with only L and DL estimated, but they were still notably underestimates of EMIN in some cases. BMDP also notably underestimated the population value of 1.00, in some cases more than ITSE did. TSX, SAS, and BMDP mean estimates were notably less than for series with only L and DL estimated, and they were generally significantly less than the population value. No particular tendencies regarding model were evident. The aberrant BMDP (0,1,1) solution was caused by one outlier value of 10. In general variability was about the same across programs and models.

Table 5

Mean and Standard Deviation of Minimum Residual Error
Variance (EMIN) for N=40, with Level (L), Change in Level
(DL), Slope (S), and Change in Slope (DS) Estimated

Population Value = 1.00

MODEL	DEP.		TSX	GENTS ^a	SAS	BMDP	ITSE
Random Numbers		mean	.95	-	.95	.95	.95
		S.D.	.22	-	.22	.22	.22
000	none	mean	.95	-	.95	.95	.95
		S.D.	.22	-	.22	.22	.22
010	none	mean	.98	-	.98	.98	-
		S.D.	.22	-	.22	.22	-
001	.8	mean	.85	-	.89	- ^b	.93
		S.D.	.20	-	.23	-	.21
001	.4	mean	.81	-	.82	- ^c	.75
		S.D.	.23	-	.24	-	.19
100	.8	mean	.90	-	.90	.84	.83
		S.D.	.22	-	.22	.21	.22
100	.4	mean	.93	-	.93	.90	.90
		S.D.	.21	-	.21	.20	.20
011	.4	mean	.86	-	.92	2.12	-
		S.D.	.22	-	.21	3.12	-
200	.4/.4	mean	.84	-	.83	.76	.75
		S.D.	.24	-	.24	.22	.22
200	.8/.2	mean	.76	-	.76	.66	.67
		S.D.	.22	-	.22	.19	.18
200	-.8/-.4	mean	.92	-	.92	.87	.92
		S.D.	.22	-	.22	.22	.25
200	.8/-.4	mean	.91	-	.91	.83	.84
		S.D.	.23	-	.23	.21	.21

^aCurrently does not provide an estimate.

^bNo solutions completed.

^cOne solution completed.

Table 6

Mean and Standard Deviation of Minimum Residual Error
Variance (EMIN) for N=100, with Level (L), Change in Level
(DL), Slope (S), and Change in Slope (DS) Estimated

Population Value = 1.00

MODEL	DEP.		TSX	GENTS ^a	SAS	BMDP	ITSE
Random Numbers		mean	.91	-	.91	.91	.91
		S.D.	.23	-	.23	.23	.23
000	none	mean	.91	-	.91	.91	.91
		S.D.	.23	-	.23	.23	.23
010	none	mean	.94	-	.94	.94	-
		S.D.	.23	-	.23	.23	-
001	.8	mean	.89	-	.88	- ^b	.95
		S.D.	.21	-	.21	-	.21
001	.4	mean	.91	-	.91	.91	.83
		S.D.	.23	-	.23	.23	.20
100	.8	mean	.91	-	.91	- ^b	.91
		S.D.	.22	-	.22	-	.22
100	.4	mean	.91	-	.91	.91	.91
		S.D.	.23	-	.23	.22	.22
011	.4	mean	.94	-	.94	.94	-
		S.D.	.22	-	.23	.23	-
200	.4/.4	mean	.90	-	.90	.88	.87
		S.D.	.22	-	.22	.22	.21
200	.8/.2	mean	.88	-	.87	.87 ^c	.83
		S.D.	.20	-	.20	.18	.20
200	-.8/-.4	mean	.91	-	.90	.88	.90
		S.D.	.22	-	.22	.20	.20
200	.8/-.4	mean	.90	-	.89	.88	.89
		S.D.	.22	-	.22	.21	.22

^aCurrently does not provide an estimate.

^bOne solution completed.

^cBased on nine solutions.

Forty versus one hundred point series. Mean estimates were generally slightly lower for 100 point series with only L and DL estimated than for 40 point series with the same two parameters. Most of the TSX, SAS, and BMDP mean estimates for these two parameter analyses were slightly less than the population value, whereas ITSE significantly underestimated the population value. Mean estimates for series with L, DL, S, and DS estimated were mixed between higher and lower for the two series lengths. Mean estimates for all programs except ITSE were significantly lower than for series with only L and DL estimated, and they were significantly lower than the population value. ITSE's estimates were generally the same for two and four parameter series. In general variability was about the same across programs, models, number of parameters estimated, and series length.

Estimate of Preintervention Level (L)

Tables 7 through 10 present the mean and standard deviation across 10 replications (except as noted) of the estimates of preintervention level (with the exception of ITSE which estimates the preintervention intercept). SAS and BMDP do not allow estimation of L with first differenced models. TSX and GENTS provide for estimation of the value of the series at $t = 0$, but these values are not included because there is no single criterion value to compare them to. The criterion value for L is zero. There is no direct

criterion value for ITSE's preintervention intercept.

With only L and DL estimated. Tables 7 and 8 present the results from analyses with L and DL estimated for series with 40 and 100 points respectively. The mean of the estimates were generally all accurate. The exception is the mean estimates of TSX, GENTS, SAS and BMDP for the (2,0,0) .8/.2 model which are all poor. Additionally, the mean of the estimates of the BMDP (1,0,0) .8 model and the 100 point BMDP (2,0,0) .4/.4 series appear to be aberrant. In general the white noise, moving averages, and (2,0,0) -.8/-.4 models are significantly better estimated than the first order and three of the second order auto-regressive models. The variability was roughly comparable across the programs. It was significantly greater for the (1,0,0) .8, the (2,0,0) .4/.4, the (2,0,0) .8/.2, and the 40 point (1,0,0) .4 and (2,0,0) .8/-.4 series.

There is no direct criterion value for the pre-intervention intercept estimates provided by ITSE. In this case, however, since there is no slope in the generated series, the intercept should be close to the value of L. With this in mind, ITSE fared poorly with the same models as the other programs. Additionally, the mean estimates tended to be slightly lower than those for the others. The variability of these estimates, however, was significantly greater in general than for the other programs.

Table 7

Mean and Standard Deviation of Estimate of Level(L)
for N=40, with only Level (L) and Change in Level (DL)
Estimated

Criterion Value = .00

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	.05	.03	.05	.05	-.02
		S.D.	.27	.28	.27	.27	.50
000	none	mean	.05	.03	.05	.05	-.02
		S.D.	.27	.28	.27	.27	.50
010	none	mean	-	-	-	-	-
		S.D.	-	-	-	-	-
001	.8	mean	.00	.00	.00	.01	-.15
		S.D.	.06	.06	.07	.08	.89
001	.4	mean	.02	.02	.02	.03	-.13
		S.D.	.16	.16	.16	.17	.92
100	.8	mean	.14	.03	.14	.36	-.08
		S.D.	1.17	1.28	1.16	1.10	.97
100	.4	mean	.10	.04	.10	.12	.02
		S.D.	.44	.48	.44	.43	.64
011	.4	mean	-	-	-	-	-
		S.D.	-	-	-	-	-
200	.4/.4	mean	.06	-.17	.05	.16	-.19
		S.D.	1.08	1.14	1.08	1.15	1.24
200	.8/.2	mean	-1.10	-1.44	-1.36	.78	-.71
		S.D.	4.72	4.54	4.35	7.71	2.47
200	-.8/-.4	mean	.01	.01	.01	.01	-.07
		S.D.	.12	.12	.12	.11	.51
200	.8/-.4	mean	.07	.05	.07	.07	-.03
		S.D.	.48	.51	.48	.46	.65

^aThis is the preintervention intercept.

Table 8

Mean and Standard Deviation of Estimate of Level(L)
for N=100, with only Level (L) and Change in Level (DL)
Estimated

Criterion Value = .00

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	.05	.06	.05	.06	-.03
		S.D.	.11	.12	.11	.11	.29
000	none	mean	.05	.06	.05	.06	-.03
		S.D.	.11	.12	.11	.11	.29
010	none	mean	-	-	-	-	-
		S.D.	-	-	-	-	-
001	.8	mean	.01	.01	.01	.01	.02
		S.D.	.02	.03	.02	.03	.21
001	.4	mean	.03	.04	.03	.03	.04
		S.D.	.07	.07	.07	.07	.36
100	.8	mean	.11	.13	.10	.34	-.01
		S.D.	.66	.65	.65	.57	.36
100	.4	mean	.08	.11	.08	.11	.01
		S.D.	.19	.21	.19	.17	.29
011	.4	mean	-	-.1	-	-	-
		S.D.	-	-	-	-	-
200	.4/.4	mean	.07	.04	.10	.43	-.04
		S.D.	.57	.69	.57	.58	.52
200	.8/.2	mean	-.74	-1.93	-1.86	.60 ^b	-.45
		S.D.	5.56	4.39	4.38	10.97	1.40
200	-.8/-.4	mean	.03	.03	.03	.03	.02
		S.D.	.05	.05	.05	.05	.34
200	.8/-.4	mean	.08	.10	.08	.11	.03
		S.D.	.21	.22	.21	.21	.36

^aThis is the preintervention intercept.

^bBased on six solutions.

With L, DL, S, and DS estimated. Tables 9 and 10 present the results from analyses with L, DL, S, and DS estimated for series with 40 and 100 points respectively. Mean estimates of L for the (1,0,0) .8, (2,0,0) .4/.4, and (2,0,0) .8/.2 models were clearly lower than the criterion value of zero, and exhibit great variability. Additionally, the variability of the (1,0,0) .4 and (2,0,0) .8/-.4 models is poor as is the variability of the 40 point series white noise solutions. The four programs' solutions are otherwise equivalent and reasonable except for the four cases where BMDP completed zero or one solution.

Since a 15 degree slope was used in the generated series, no meaningful statement can be made regarding ITSE's preintervention intercept, however, it may be noted that ITSE consistently returned a higher mean estimate than the other programs. These mean values were a closer estimate of L for the first order autoregressive and three of the four second order autoregressive models than the others provided. Except for the white noise model, the estimates for the remaining models appear to be curiously high.

Forty versus one hundred point series. The same pattern of results was found with 40 and 100 point series. Mean values for the pathological cases, i.e., the (1,0,0) .8, (2,0,0) .4/.4, and the (2,0,0) .8/.2 models, were generally worse for 100 point series. Variability was generally lower for 100 point series.

Table 9

Mean and Standard Deviation of Estimate of Level (L)
for N=40, with Level (L), Change in Level (DL), Slope (S),
and Change in Slope (DS) Estimated

Criterion Value = .00

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	-.02	.00	-.02	-.02	-.02
		S.D.	.50	.48	.50	.50	.50
000	none	mean	-.02	.00	-.02	-.02	-.02
		S.D.	.50	.48	.50	.50	.50
010	none	mean	-	-	-	-	-
		S.D.	-	-	-	-	-
001	.8	mean	-.02	-.02	-.01	- ^b	2.20
		S.D.	.20	.19	.21	-	.87
001	.4	mean	-.01	-.00	-.02	- ^c	1.81
		S.D.	.33	.33	.34	-	.89
100	.8	mean	-.62	-.73	-.62	-1.55	-.15
		S.D.	1.68	1.69	1.67	2.96	.94
100	.4	mean	-.21	-.16	-.21	-.27	.17
		S.D.	.71	.73	.71	1.02	.62
011	.4	mean	-	-	-	-	-
		S.D.	-	-	-	-	-
200	.4/.4	mean	-.71	-.79	-.71	-1.78	-.28
		S.D.	1.60	1.58	1.61	2.81	1.18
200	.8/.2	mean	-1.34	-1.46	-1.31	-1.83	-.78
		S.D.	4.30	4.15	4.31	5.72	2.56
200	-.8/-.4	mean	.18	.18	.18	.16	1.04
		S.D.	.27	.25	.27	.28	.50
200	.8/-.4	mean	-.15	-.07	-.13	-.17	.29
		S.D.	.71	.76	.72	.95	.66

^aThis is the preintervention intercept.

^bNo solutions completed.

^cOne solution completed.

Table 10

Mean and Standard Deviation of Estimate of Level(L)
for N=100, with Level (L), Change in Level (DL), Slope (S),
and Change in Slope (DS) Estimated

Criterion Value = .00

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	-.03	-.01	-.03	-.03	-.03
		S.D.	.29	.29	.29	.29	.29
000	none	mean	-.03	-.01	-.03	-.03	-.03
		S.D.	.29	.29	.29	.29	.29
010	none	mean	-	-	-	-	-
		S.D.	-	-	-	-	-
001	.8	mean	.01	.00	.00	- ^b	2.40
		S.D.	.06	.06	.06	-	.21
001	.4	mean	.00	.00	-.00	-.00	1.94
		S.D.	.17	.17	.17	.17	.34
100	.8	mean	-.78	-.71	-.78	- ^b	.00
		S.D.	1.31	1.32	1.31	-	.37
100	.4	mean	-.22	-.18	-.22	-.16	.17
		S.D.	.48	.49	.48	.46	.29
011	.4	mean	-	-	-	-	-
		S.D.	-	-	-	-	-
200	.4/.4	mean	-.98	-.95	-.97	-1.84	-.02
		S.D.	1.18	1.22	1.19	2.10	.56
200	.8/.2	mean	-1.71	-1.68	-1.74	-2.23 ^c	-.43
		S.D.	3.99	4.05	3.99	7.20	1.39
200	-.8/-.4	mean	.21	.21	.21	.21	1.18
		S.D.	.11	.11	.11	.15	.34
200	.8/-.4	mean	-.02	-.02	-.03	.05	.36
		S.D.	.56	.56	.56	.59	.36

^aThis is the preintervention intercept.

^bOne solution completed.

^cBased on nine solutions.

Estimates of L with slope versus without slope. The mean estimates of L are significantly deteriorated, and the variability of estimates of L is significantly increased, by the introduction of a four parameter analysis. That the introduction of the four parameter analysis caused this, as opposed to the introduction of the actual slope and change in slope effects themselves, may be deduced by comparing the random number solutions to the white noise solutions. The former are identical to the latter within each table, and thus are deteriorated in the L, DL, S, and DS solutions. Since the random numbers did not contain any L, DL, S, and DS effects, however, as did the white noise model, the deterioration must have been introduced by the mechanics of the analysis.

As regards the mean estimate of L, the accuracy of all 40 point series solutions except for the white noise and (0,0,1) .4 models was deteriorated, sometimes significantly so. The same observation holds true in 100 point series except that the mean estimate for the (0,0,1) .8 model was essentially the same for with and without slope solutions, and the solutions with slope for the (2,0,0) .8/-.4 model were better than the ones for those without slope.

Estimate of Postintervention Change in Level (DL)

Tables 11 through 14 present the mean and standard deviation across 10 replications (except as noted) of the estimates of postintervention change in level except for

ITSE which estimates the postintervention intercept.

With only L and DL estimated. Tables 11 and 12 present the results from analyses with only L and DL estimated for series with 40 and 100 points respectively. The criterion value for DL was .50. There is no criterion value for ITSE's postintervention intercept.

TSX, GENTS, SAS, and BMDP mean estimates for the (0,1,1) .4 and (2,0,0) .8/.2 models were generally lower than the criterion value (.50) as were the mean estimates for the 40 point (0,1,0) model. BMDP noticeably underestimated DL for the (1,0,0) .8 model as did GENTS for the 100 point (0,1,0) series. The remaining model solutions provided reasonable and generally equivalent mean estimations of DL although TSX, SAS, and BMDP tended to underestimate DL with the 40 point first order autoregressive models as compared to GENTS, and as compared to their own first order moving averages solutions. Additionally, the variability was significant for the differencing models and all of the autoregressive models except the (2,0,0) -.8/-.4 model.

ITSE's postintervention intercept has no direct criterion value, however, given that there is no slope and change in slope in the generated series one may perform an approximate comparison of the difference between the post-intervention intercepts given in Tables 11 and 12 and the preintervention intercepts given in Tables 7 and 8 to the

Table 11

Mean and Standard Deviation of Estimate of Change in Level (DL) for N=40, with only Level (L) and Change in Level (DL) Estimated

Criterion Value = .50

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	-.02	.01	-.02	-.02	-.10
		S.D.	.33	.36	.33	.33	.40
000	none	mean	.48	.51	.48	.47	.40
		S.D.	.33	.36	.33	.33	.40
010	none	mean	.08	-.34	.08	.08	-
		S.D.	.93	1.06	.93	.93	-
001	.8	mean	.50	.51	.50	.48	1.05
		S.D.	.07	.07	.08	.06	.45
001	.4	mean	.50	.51	.50	.54	.77
		S.D.	.20	.21	.20	.26	.76
100	.8	mean	.33	.47	.33	.15	.16
		S.D.	1.19	1.40	1.18	1.07	.89
100	.4	mean	.41	.51	.41	.39	.26
		S.D.	.55	.61	.55	.54	.53
011	.4	mean	-.00	.10	.03	.03	-
		S.D.	.99	.97	.93	.93	-
200	.4/.4	mean	.52	.39	.51	.40	.32
		S.D.	1.12	1.40	1.12	1.00	1.06
200	.8/.2	mean	.26	-.35	-.03	.06	.12
		S.D.	1.20	1.03	.85	.82	3.72
200	-.8/-.4	mean	.51	.51	.51	.51	.83
		S.D.	.15	.16	.15	.14	.43
200	.8/-.4	mean	.47	.52	.47	.47	.27
		S.D.	.58	.63	.58	.55	.65

^aThis is the postintervention intercept.

Table 12

Mean and Standard Deviation of Estimate of Change in Level (DL) for N=100, with only Level (L) and Change in Level (DL) Estimated

Criterion Value = .50

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	-.01	-.03	-.01	-.01	-.02
		S.D.	.20	.22	.20	.20	.33
000	none	mean	.49	.47	.49	.49	.48
		S.D.	.20	.22	.20	.20	.33
010	none	mean	.46	.30	.46	.46	-
		S.D.	.74	.77	.74	.75	-
001	.8	mean	.50	.49	.50	.50	1.36
		S.D.	.04	.04	.04	.04	.26
001	.4	mean	.49	.49	.49	.49	.84
		S.D.	.12	.12	.12	.12	.39
100	.8	mean	.53	.49	.53	.36	.13
		S.D.	.92	.93	.92	.84	.44
100	.4	mean	.48	.43	.49	.46	.32
		S.D.	.34	.40	.34	.32	.35
011	.4	mean	.29	.09	.30	.29	-
		S.D.	.77	.81	.77	.77	-
200	.4/.4	mean	.49	.36	.47	.25	.14
		S.D.	.90	1.03	.89	.84	.52
200	.8/.2	mean	1.40	.18	.61	.28 ^b	-.41
		S.D.	1.93	.79	1.11	.83	2.43
200	-.8/-.4	mean	.49	.49	.49	.49	1.01
		S.D.	.09	.09	.09	.09	.30
200	.8/-.4	mean	.49	.46	.49	.47	.32
		S.D.	.35	.38	.35	.34	.40

^aThis is the postintervention intercept.

^bBased on six solutions.

criterion value of DL, i.e., .50. When this is done, the resulting mean estimates are not accurate.

With L, DL, S, and DS estimated. Tables 13 and 14 present the results from analyses with L, DL, S, and DS estimated for series with 40 and 100 points respectively. The criterion value for DL was 1.50. There is no direct criterion value for ITSE's postintervention intercept.

The TSX, GENTS, SAS, and BMDP mean estimates of DL for the (0,1,0), (0,1,1) .4 and (2,0,0) .8/.2 models were significantly lower than the criterion value (1.50) as were their mean estimates for the 40 point series (1,0,0) .8, (1,0,0) .4, (2,0,0) .4/.4, and (2,0,0) .8/-.4 models (with the exception of GENTS mean estimates for the latter three). Note also that BMDP completed zero or one solution for four of the models. Estimations of DL for the remaining models was essentially the same for the four programs and reasonably recovered the criterion value. The variability is excessive for the white noise model, the differencing models, all autoregressive models except the (2,0,0) -.8/-.4 model, and the 40 point (0,0,1) .4 model.

ITSE's postintervention intercept estimates have no direct criterion value, however, it was noted that in some cases, the estimate considerably exceeded the largest value in the postintervention series. Since the postintervention series is essentially flat, i.e., slope = 0, these appear to be inaccuracies in the ITSE solutions.

Table 13

Mean and Standard Deviation of Estimate of Change in Level (DL) for N=40, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated

Criterion Value = 1.50

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	-.21	-.14	-.21	-.21	-.10
		S.D.	.62	.83	.62	.62	.40
000	none	mean	1.29	1.36	1.29	1.29	6.76
		S.D.	.62	.83	.62	.62	.40
010	none	mean	.76	.24	.83	.80	-
		S.D.	.96	1.66	.94	.97	-
001	.8	mean	1.46	1.44	1.46	- ^b	13.49
		S.D.	.30	.30	.35	-	2.82
001	.4	mean	1.40	1.41	1.36	- ^c	11.66
		S.D.	.57	.54	.48	-	3.14
100	.8	mean	.88	1.10	.87	.89	2.95
		S.D.	1.23	1.62	1.23	1.28	1.52
100	.4	mean	1.18	1.51	1.18	1.18	4.75
		S.D.	.92	1.21	.92	.98	1.25
011	.4	mean	.65	.91	.58	.76	-
		S.D.	1.08	1.54	1.05	1.61	-
200	.4/.4	mean	1.18	1.38	1.21	1.07	4.02
		S.D.	1.26	1.55	1.26	1.28	2.24
200	.8/.2	mean	.45	.44	.44	.47	.55
		S.D.	1.23	1.58	1.24	1.22	3.76
200	-.8/-.4	mean	1.25	1.25	1.25	1.20	13.54
		S.D.	.35	.39	.36	.33	1.82
200	.8/-.4	mean	1.14	1.33	1.13	1.15	5.15
		S.D.	1.00	1.17	1.00	1.09	1.28

^aThis is the postintervention intercept.

^bNo solutions completed.

^cOne solution completed.

Table 14

Mean and Standard Deviation of Estimate of Change in Level (DL) for N=100, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated

Criterion Value = 1.50

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE ^a
Random Numbers		mean	-.15	-.12	-.15	-.15	-.02
		S.D.	.47	.49	.47	.47	.33
000	none	mean	1.35	1.38	1.35	1.35	14.88
		S.D.	.47	.49	.47	.47	.33
010	none	mean	1.15	.87	1.15	1.15	-
		S.D.	.75	.86	.75	.75	-
001	.8	mean	1.49	1.48	1.48	- ^b	31.08
		S.D.	.09	.12	.10	-	3.52
001	.4	mean	1.43	1.43	1.43	1.31	22.75
		S.D.	.29	.30	.29	.51	2.74
100	.8	mean	1.54	1.46	1.54	- ^b	4.03
		S.D.	1.07	1.08	1.08	-	.86
100	.4	mean	1.44	1.46	1.44	1.46	9.73
		S.D.	.70	.78	.70	.73	.52
011	.4	mean	.97	.63	.97	.97	-
		S.D.	.76	1.02	.75	.75	-
200	.4/.4	mean	1.50	1.48	1.52	1.46	4.84
		S.D.	.97	.98	1.00	1.06	1.50
200	.8/.2	mean	.87	.72	.89	1.08 ^c	-.21
		S.D.	.93	.96	.93	.93	2.40
200	-.8/-.4	mean	1.26	1.27	1.26	1.26	31.93
		S.D.	.21	.22	.21	.23	2.40
200	.8/-.4	mean	1.37	1.36	1.37	1.40	9.79
		S.D.	.76	.80	.77	.81	1.26

^aThis is the postintervention intercept.

^bOne solution completed.

^cBased on nine solutions.

Forty versus one hundred point series. The mean estimates for all of the $(1,0,0)$.8 and $(1,0,0)$.4, and the with slope $(2,0,0)$.4/.4 and $(2,0,0)$.8/-.4 models improved sufficiently enough to become acceptable estimates of DL for series with 100 points. The high variability of the estimates from all of these series, however, was still sufficient to make the estimates unacceptable overall. In general, variability of estimates for 100 point series was significantly less than for 40 point series.

Estimates of DL with slope versus without slope. As with the estimates of L, the estimates of DL are generally considerably deteriorated by the introduction of a four parameter analysis. The same rationale as previously explained regarding the deterioration of estimates of L serves here, except it must be noted that the random number and white noise mean estimates are not the same. In this case, the white noise estimate is larger than the random number estimate by the appropriate criterion value, i.e., by .50 when two parameters are estimated and by 1.50 when four parameters are estimated.

Specifically, except for the 100 point $(0,0,1)$.8, $(1,0,0)$.8, $(1,0,0)$.4, and $(2,0,0)$.4/.4 models which essentially had mean estimates equivalent to the nonslope models, the mean estimate for series with slope estimated was not as accurate as for nonslope series. Additionally, variability was increased consistently.

Estimate of Preintervention Slope (S)

Table 15 contains the mean and standard deviation of the estimates of preintervention slope for series with 40 points. The criterion value for S was .27. Table 16 presents the angles in degrees represented by the mean estimates of these slopes and plus and minus one standard deviation from the mean estimate. This table is provided to give the reader a more meaningful sense of the range of inaccuracy of any particular estimate (because degrees are more meaningful in general than tangents and because the same variability represents a different angle depending on the mean estimate). The criterion value for S in degrees was 15. Similarly Tables 17 and 18 present the mean and standard deviation of the slope, and the angles, respectively, for series with 100 points. The criterion values for S for these series were also .27 and 15 degrees respectively.

With L, DL, S, and DS estimated. TSX, GENTS, SAS, and BMDP all seriously underestimated S for the two differencing models and the pseudo differencing (2,0,0) .8/.2 model. Reasonable mean estimates were provided by all four programs for the remaining models, except for the four models for which BMDP completed either zero or one solution. ITSE fared poorly with all models except the white noise model. It overestimated with moving averages and the (2,0,0) -.8/-.4 models and underestimated with the remaining

Table 15

Mean and Standard Deviation of Estimate of Slope (S)
for N=40, with Level (L), Change in Level (DL), Slope (S),
and Change in Slope (DS) Estimated

Criterion Value = .27

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
Random Numbers		mean	.01	.00	.01	.01	.01
		S.D.	.04	.04	.04	.04	.04
000	none	mean	.27	.27	.27	.27	.27
		S.D.	.04	.04	.04	.04	.04
010	none	mean	.05	.05	.05	.05	-
		S.D.	.26	.25	.26	.26	-
001	.8	mean	.27	.27	.27	- ^a	.57
		S.D.	.02	.02	.02	-	.11
001	.4	mean	.27	.27	.27	- ^b	.51
		S.D.	.03	.03	.03	-	.18
100	.8	mean	.27	.27	.27	.29	.14
		S.D.	.12	.13	.12	.28	.12
100	.4	mean	.28	.27	.28	.28	.20
		S.D.	.06	.07	.06	.08	.07
011	.4	mean	.04	.03	.04	.04	-
		S.D.	.17	.18	.16	.18	-
200	.4/.4	mean	.24	.24	.24	.30	.20
		S.D.	.11	.11	.11	.19	.18
200	.8/.2	mean	.04	.05	.04	.07	.06
		S.D.	.22	.23	.22	.33	.19
200	-.8/-.4	mean	.27	.27	.27	.27	.56
		S.D.	.02	.02	.02	.03	.09
200	.8/-.4	mean	.29	.28	.29	.29	.22
		S.D.	.06	.06	.06	.08	.08

^aNo solutions completed.

^bOne solution completed.

Table 16

Mean and Plus and Minus One Standard Deviation Values of
Slope (S) for N=40, in Degrees

Criterion Value = 15.0 degrees

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
000	none		17.6	17.5	17.6	17.6	17.6
		mean	15.4	15.2	15.3	15.3	15.4
			13.0	12.7	13.0	13.0	13.0
010	none		17.7	16.4	17.7	17.7	-
		mean	3.1	2.6	3.1	3.1	-
			-11.8	-11.5	-11.8	-11.8	-
001	.8		16.1	16.0	16.1	-	34.3
		mean	15.1	15.1	15.0	-	29.6
			14.1	14.2	14.0	-	24.4
001	.4		16.8	16.7	16.7	-	34.6
		mean	15.2	15.1	15.2	-	27.0
			13.5	13.5	13.6	-	18.2
100	.8		21.3	21.9	21.3	29.7	14.5
		mean	15.1	15.1	15.1	16.1	7.9
			8.5	7.9	8.5	.4	1.1
100	.4		18.8	18.6	18.8	20.2	14.9
		mean	15.7	15.1	15.7	15.9	11.1
			12.4	11.5	12.4	11.3	7.1
011	.4		12.1	11.8	11.6	12.6	-
		mean	2.4	1.9	2.5	2.4	-
			-7.3	-8.1	-6.8	-8.0	-
200	.4/.4		19.4	19.4	19.4	26.1	20.7
		mean	13.4	13.4	13.4	16.8	11.2
			7.2	7.0	7.0	6.5	1.0
200	.8/.2		14.8	15.9	14.7	21.8	14.0
		mean	2.4	3.0	2.4	4.0	3.4
			-10.2	-10.2	-10.2	-14.5	-7.4
200	-.8/-.4		16.4	16.3	16.4	16.6	32.7
		mean	15.2	15.1	15.2	15.2	29.1
			13.9	13.9	13.9	13.9	25.2
200	.8/-.4		19.1	18.9	19.1	20.3	16.5
		mean	16.0	15.6	16.0	16.1	12.4
			12.9	12.1	12.8	11.7	8.1

Table 17

Mean and Standard Deviation of Estimate of Slope (S)
for N=100, with Level (L), Change in Level (DL), Slope (S),
and Change in Slope (DS) Estimated

Criterion Value = .27

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
Random Numbers		mean	.00	.00	.00	.00	.00
		S.D.	.01	.01	.01	.01	.01
000	none	mean	.27	.27	.27	.27	.27
		S.D.	.01	.01	.01	.01	.01
010	none	mean	.06	.05	.06	.06	-
		S.D.	.10	.12	.10	.10	-
001	.8	mean	.27	.27	.27	— ^a	.57
		S.D.	.00	.00	.00	-	.06
001	.4	mean	.27	.27	.27	.27	.41
		S.D.	.01	.01	.01	.01	.04
100	.8	mean	.27	.27	.27	— ^a	.07
		S.D.	.03	.03	.03	-	.02
100	.4	mean	.27	.27	.26	.27	.18
		S.D.	.02	.02	.05	.02	.02
011	.4	mean	.04	.04	.04	.04	-
		S.D.	.06	.07	.06	.06	-
200	.4/.4	mean	.27	.27	.27	.29	.03
		S.D.	.02	.02	.02	.06	.03
200	.8/.2	mean	.04	.04	.04	.05 ^b	.00
		S.D.	.10	.10	.10	.12	.02
200	-.8/-.4	mean	.27	.27	.27	.27	.58
		S.D.	.00	.00	.00	.01	.05
200	.8/-.4	mean	.27	.27	.27	.27	.18
		S.D.	.02	.02	.02	.02	.02

^aOne solution completed.

^bBased on nine solutions.

Table 18

Mean and Plus and Minus One Standard Deviation Values of
Slope ($\hat{S}$) for N=100, in Degrees

Criterion Value = 15.0 degrees

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
000	none		15.7	15.7	15.7	15.7	15.7
		mean	15.2	15.1	15.2	15.2	15.2
			14.6	14.6	14.6	14.6	14.6
010	none		9.4	9.5	9.4	9.4	-
		mean	3.7	2.8	3.7	3.7	-
			-2.1	-3.9	-2.1	-2.1	-
001	.8		15.1	15.1	15.1	-	32.1
		mean	15.0	15.0	15.0	-	29.5
			14.9	14.9	14.9	-	26.7
001	.4		15.4	15.4	15.4	15.4	24.5
		mean	15.1	15.1	15.1	15.1	22.4
			14.7	14.7	14.7	14.7	20.2
100	.8		16.9	16.8	16.9	-	5.1
		mean	15.3	15.2	15.3	-	4.2
			13.6	13.5	13.6	-	3.3
100	.4		16.1	16.1	17.1	16.1	10.9
		mean	15.3	15.2	14.5	15.2	10.0
			14.4	14.3	11.8	14.3	9.2
011	.4		5.9	5.9	5.8	5.8	-
		mean	2.3	2.0	2.4	2.4	-
			-1.2	-1.9	-1.1	-1.1	-
200	.4/.4		16.2	16.1	16.2	19.0	3.4
		mean	14.9	14.9	14.9	16.0	1.7
			13.7	13.6	13.7	12.9	0.0
200	.8/.2		8.0	8.1	8.1	9.3	1.2
		mean	2.5	2.6	2.6	2.6	.2
			-3.1	-3.1	-3.0	-4.1	-.8
200	-.8/-.4		15.3	15.2	15.3	15.3	32.2
		mean	15.0	15.0	15.0	15.0	30.1
			14.8	14.8	14.8	14.8	27.9
200	.8/-.4		16.2	16.2	16.2	16.2	11.2
		mean	15.2	15.2	15.2	15.1	10.1
			14.2	14.2	14.2	14.0	9.0

autoregressive models.

The variability of estimates produced by TSX, GENTS, SAS, and BMDP was essentially the same with the exceptions being 1) the increased variability of the BMDP (1,0,0) .8, (2,0,0) .4/.4, and (2,0,0) .8/.2 estimates for 40 point series, and 2) the increased variability of the SAS (1,0,0) .4 and BMDP (2,0,0) .4/.4 estimates for 100 point series. Of the models with reasonable mean estimates, variability was noticeably less with these four programs for the white noise, moving averages and (2,0,0) -.8/-.4 models than for the remaining autoregressive models. The variability of ITSE's estimates was considerably larger than the other programs' for the two moving averages and (2,0,0) -.8/-.4 models, and generally the same for the remaining models.

Forty versus one hundred point series. For TSX, GENTS, SAS, and BMDP the mean estimates of S are essentially the same for series with 40 and 100 points. As is to be expected, however, the variability is considerably reduced with 100 point series. Note for example that the GENTS (1,0,0) .8 estimate has a plus and minus one s range of 8 to 22 degrees for the 40 point series and 13 to 17 degrees for the 100 point series. Some of the ITSE mean estimates changed and the variability was also considerably reduced by the larger number of points. All of the ITSE results are poor, however, as stated previously.

Estimate of Postintervention Change in Slope (DS)

Table 19 contains the mean and standard deviation of the estimates of postintervention change in slope for series with 40 points. The criterion value was $-.27$. Table 20 presents the angles in degrees represented by the mean estimate and plus and minus one standard deviation. The criterion value for the mean estimate was -15 degrees. Similarly Tables 21 and 22 contain the mean and standard deviation, and angles, for series with 100 points. The criterion values were also $-.27$ and -15 degrees respectively.

With L, DL, S, and DS estimated. TSX, GENTS, SAS, and BMDP seriously underestimated DS for the differencing, $(1,0,0) .8$, $(2,0,0) .4/.4$, and $(2,0,0) .8/.2$ models. The remaining estimates are essentially the same for these four programs with the exception of the four models for which BMDP completed zero or one solution (as noted) and BMDP's 40 point series $(1,0,0) .8$ solution. Of those mean estimates remaining, the criterion value of $-.27$ was reasonably recovered. For the same four programs the variability of the five models with inaccurate mean estimates cited above was excessive. Additionally, the variability of the 40 point series $(1,0,0) .4$ and $(2,0,0) .8/-.4$ estimates is borderline, representing approximately plus and minus five degrees. ITSE estimated well for the white noise model and, as with the slope estimates, overestimated with moving

Table 19

Mean and Standard Deviation of Estimate of Change in Slope (DS) for N=40, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated

Criterion Value = $-.27$

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
Random Numbers		mean	.01	.01	.01	.01	.01
		S.D.	.05	.06	.05	.05	.05
000	none	mean	-.26	-.26	-.26	-.26	-.26
		S.D.	.05	.06	.05	.05	.05
010	none	mean	-.01	.03	-.01	-.01	-
		S.D.	.34	.30	.34	.34	-
001	.8	mean	-.27	-.27	-.26	- ^a	-.52
		S.D.	.02	.02	.02	-	.09
001	.4	mean	-.26	-.26	-.26	- ^b	-.46
		S.D.	.04	.04	.04	-	.14
100	.8	mean	-.18	-.18	-.17	-.24	-.11
		S.D.	.17	.20	.17	.25	.12
100	.4	mean	-.25	-.25	-.25	-.26	-.18
		S.D.	.08	.09	.08	.10	.08
011	.4	mean	-.00	.00	.04	-.03	-
		S.D.	.22	.21	.20	.23	-
200	.4/.4	mean	-.13	-.14	-.13	-.20	-.14
		S.D.	.16	.18	.16	.22	.17
200	.8/.2	mean	.02	.00	.02	-.02	-.03
		S.D.	.28	.29	.27	.35	.21
200	-.8/-.4	mean	-.27	-.27	-.27	-.27	-.52
		S.D.	.03	.03	.03	.02	.08
200	.8/-.4	mean	-.27	-.27	-.27	-.27	-.20
		S.D.	.08	.09	.08	.10	.09

^aNo solutions completed.

^bOne solution completed.

Table 20

Mean and Plus and Minus One Standard Deviation Values of
Change in Slope (DS) for N=40, in Degrees

Criterion Value = -15.0 degrees

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
000	none		-11.9	-11.4	-11.9	-11.9	-11.9
		mean	-14.7	-14.6	-14.7	-14.7	-14.7
			-17.5	-17.7	-17.5	-17.5	-17.5
010	none		18.1	18.0	18.1	18.1	-
		mean	-.5	1.6	-.5	-.5	-
			-19.0	-15.0	-19.0	-19.0	-
001	.8		-14.0	-13.9	-13.5	-	-23.1
		mean	-15.0	-14.9	-14.7	-	-27.4
			-15.9	-15.8	-15.9	-	-31.3
001	.4		-12.7	-12.7	-12.5	-	-17.7
		mean	-14.8	-14.7	-14.6	-	-24.8
			-16.8	-16.7	-16.7	-	-31.2
100	.8		-.5	.8	-.5	1.0	.6
		mean	-9.9	-10.4	-9.9	-13.3	-6.4
			-18.8	-20.8	-18.8	-26.2	-13.3
100	.4		-9.6	-9.3	-9.6	-8.8	-5.8
		mean	-14.2	-14.3	-14.2	-14.4	-10.4
			-18.6	-19.1	-18.6	-19.7	-14.9
011	.4		12.2	12.1	13.3	11.1	-
		mean	-.1	.1	2.3	-1.9	-
			-12.4	-11.9	-8.9	-14.8	-
200	.4/.4		1.6	2.5	1.6	1.3	1.5
		mean	-7.5	-7.7	-7.5	-11.4	-8.0
			-16.2	-17.5	-16.2	-23.1	-17.1
200	.8/.2		16.3	16.3	16.1	18.5	10.5
		mean	.9	.0	1.0	-1.0	-1.6
			-14.6	-16.2	-14.3	-20.3	-13.6
200	-.8/-.4		-13.5	-13.4	-13.5	-13.8	-23.8
		mean	-15.0	-14.9	-15.0	-15.1	-27.6
			-16.4	-16.4	-16.4	-16.4	-31.2
200	.8/-.4		-10.6	-10.0	-10.5	-9.7	-6.2
		mean	-15.1	-15.0	-15.0	-15.2	-11.4
			-19.4	-19.8	-19.4	-20.4	-16.4

Table 21

Mean and Standard Deviation of Estimate of Change in Slope (DS) for N=100, with Level (L), Change in Level (DL), Slope (S), and Change in Slope (DS) Estimated

Criterion Value = $-.27$

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
Random Numbers		mean	-.00	-.00	-.00	-.00	-.00
		S.D.	.01	.01	.01	.01	.01
000	none	mean	-.27	-.27	-.27	-.27	-.27
		S.D.	.01	.01	.01	.01	.01
010	none	mean	-.02	.01	-.02	-.02	-
		S.D.	.20	.23	.20	.20	-
001	.8	mean	-.27	-.27	-.27	- ^a	-.56
		S.D.	.00	.00	.00	-	.06
001	.4	mean	-.27	-.27	-.27	-.27	-.41
		S.D.	.01	.01	.01	.01	.04
100	.8	mean	-.25	-.25	-.25	- ^a	-.07
		S.D.	.04	.04	.04	-	.02
100	.4	mean	-.27	-.27	-.27	-.27	-.18
		S.D.	.02	.02	.02	.02	.02
011	.4	mean	-.01	.01	-.01	-.01	-
		S.D.	.12	.14	.12	.12	-
200	.4/.4	mean	-.23	-.23	-.23	-.25	-.08
		S.D.	.03	.03	.03	.07	.02
200	.8/.2	mean	.01	.01	.00	.00 ^b	.01
		S.D.	.19	.19	.19	.17	.04
200	-.8/-.4	mean	-.27	-.27	-.27	-.27	-.57
		S.D.	.01	.01	.01	.01	.04
200	.8/-.4	mean	-.27	-.27	-.27	-.27	-.18
		S.D.	.02	.02	.02	.02	.03

^aOne solution completed.

^bBased on nine solutions.

Table 22

Mean and Plus and Minus One Standard Deviation Values of
Change in Slope (DS) for N=100, in Degrees

Criterion Value = -15.0 degrees

MODEL	DEP.		TSX	GENTS	SAS	BMDP	ITSE
000	none		-14.4	-14.4	-14.4	-14.4	-14.4
		mean	-15.0	-15.0	-15.0	-15.1	-15.0
			-15.7	-15.7	-15.7	-15.7	-15.7
010	none		9.8	13.2	9.8	9.8	-
		mean	-1.3	.5	-1.3	-1.3	-
			-12.3	-12.3	-12.3	-12.3	-
001	.8		-14.8	-14.8	-14.8	-	-26.5
		mean	-15.0	-15.0	-15.0	-	-29.2
			-15.2	-15.2	-15.2	-	-31.7
001	.4		-14.6	-14.6	-14.6	-14.6	-20.1
		mean	-15.0	-15.0	-15.0	-15.0	-22.2
			-15.4	-15.4	-15.4	-15.4	-24.3
100	.8		-11.9	-11.8	-11.9	-	-3.2
		mean	-14.1	-13.8	-14.1	-	-4.1
			-16.1	-15.8	-16.1	-	-5.0
100	.4		-14.0	-13.9	-13.9	-14.0	-9.1
		mean	-15.1	-15.0	-14.9	-15.0	-10.0
			-16.1	-16.0	-15.9	-16.0	-10.8
011	.4		6.2	8.1	6.1	6.1	-
		mean	-.6	.3	-.7	-.7	-
			-7.5	-7.6	-7.4	-7.4	-
200	.4/.4		-11.2	-11.1	-11.2	-10.6	-3.4
		mean	-13.0	-12.8	-13.0	-14.2	-4.8
			-14.7	-14.5	-14.7	-17.6	-6.1
200	.8/.2		11.2	11.1	11.0	9.7	2.5
		mean	.5	.4	.2	.3	.4
			-10.2	-10.5	-10.6	-9.1	-1.7
200	-.8/-.4		-14.7	-14.7	-14.7	-14.7	-27.9
		mean	-15.0	-15.0	-15.0	-15.0	-29.8
			-15.3	-15.3	-15.3	-15.3	-31.6
200	.8/-.4		-14.0	-14.0	-14.0	-13.9	-8.6
		mean	-15.1	-15.1	-15.1	-15.0	-10.0
			-16.2	-16.2	-16.2	-16.1	-11.4

averages and the (2,0,0) $-.8/.4$ models and underestimated with the remaining autoregressive models.

Forty versus one hundred point series. The mean estimates of DS from TSX, GENTS, SAS, and BMDP are essentially the same for 40 and 100 point series with models that provide reasonable solutions. As with the slope estimates, however, the variability with 100 point series estimates is considerably reduced. The largest variability for 100 point series is approximately plus and minus one degree. Only the least variable of the 40 point series, i.e, the (0,0,1) $.8$ model, achieves this accuracy, while the (1,0,0) $.4$, and (2,0,0) $.8/- .4$ estimates vary approximately plus and minus five degrees. ITSE estimates are unreasonable for both series lengths.

Computer Time Requirements

Table 23 presents representative central processing unit (CPU) times required by each of the five programs. The CPU times presented are the times for one solution or analysis. It should be noted that a number of factors that might increase time requirements were not controlled for, such as Fortran underflow or overflow processing and the number of function iterations required for a particular solution. Since SAS and BMDP are executed from an object code library which adds a negligible amount of time to the processing, TSX, GENTS, and ITSE times are provided for the execution step only. If a user were to compile any of these

Table 23

Representative Central Processing Unit (CPU) Times

N	n of parms.	MODEL	DEP.	TSX ^a	GENTS ^b	SAS	BMDP	ITSEC ^c
		000	none	.0	12.3	.4	1.0	.0
40	2	100	.4	2.0	11.7	.5	1.1	.0
		200	.4/.4	7.5	14.1	.7	1.1	.1
		000	none	.1	13.0	.5	1.0	.0
40	4	100	.4	3.7	14.1	.8	1.1	.0
		200	.4/.4	13.7	13.2	1.1	1.2	.1
		000	none	.1	155.3	.5	1.7	.1
100	2	100	.4	9.9	173.9	.6	1.7	.1
		200	.4/.4	36.3	176.7	.9	1.9	.1
		000	none	.2	161.5	.7	1.8	.1
100	4	100	.4	17.5	166.1	1.2	1.9	.1
		200	.4/.4	65.0	171.9	1.6	2.1	.1

^aAdd 2.4 seconds for compilation.^bAdd 2.9 seconds for compilation.^cAdd 1.5 seconds for compilation.

three programs prior to an analysis, the times provided in the Table 23 footnotes should be added to the times in the body of the table. Times presented as .0 represent times less than .05 seconds. Finally, the times presented were obtained from computer analyses performed on an IBM 4381-3 mainframe computer using the MVS/SP1-3-4 operating system with VS Fortran. The CPU time required at different installations will of course vary, but the values for each program relative to each other should remain approximately constant.

CPU times for TSX varied by the length of the series, number of intervention parameters estimated, and the number of dependency parameters. The amount of time required varies from less than .05 seconds ($N = 40$; only L and DL estimated; no dependency) to approximately 65 seconds ($N = 100$; L, DL, S, and DS estimated; second order model). Greater CPU times were required for a greater number of points, more intervention parameters, and more dependency parameters. It should be noted that the (2,0,0) models were estimated using an iteration increment of .10 as opposed to .02 for the remaining models. Even with this advantage (2,0,0) models required more time.

GENTS CPU times were increased generally by the number of points in the series alone, partly because it always estimates five dependency parameters. The remaining differences are negligible.

The amount of CPU time required by SAS was increased by the order of the model, the number of intervention parameters estimated, and the series length. Times varied between .4 seconds and 1.6 seconds. These times also include computing an autocorrelation, inverse autocorrelation, and partial autocorrelation of the series residuals.

BMDP CPU times also increased with the order of the model, number of intervention parameters estimated, and the series length. Its times varied between 1.0 and 2.1 seconds.

ITSE times were considerably less than all of the other programs. They varied between less than .05 seconds to approximately .1 seconds.

Discussion

This section is divided into three subsections. The Model Identification subsection summarizes and discusses the results of this study as they pertain to the question of whether or not the problematic model identification stage of interrupted time series analysis is required. The Computer Programs subsection summarizes and discusses the results of the evaluation of the five selected computer programs. And, finally, the Recommendations subsection presents recommendations for current analyses, improvements to computer programs, and future research.

Model Identification

As stated earlier, the approach for addressing the question of whether or not it is necessary to perform model identification was to 1) compare the TSX solutions to the true or simulated values, and 2) compare the GENTS solutions to the TSX values as well as to the true values. In this manner, information was obtained regarding the accuracy or inaccuracy of the "best solutions" that could be achieved (i.e., given that the true model identification is known or can be determined), and regarding the accuracy or inaccuracy of the General Transformation Approach (Velicer and McDonald, 1984) which does not require that the model be known or identified.

As regards the best solutions, it was seen that, generally, estimates of EMIN and the intervention parameters were more accurate and consistent for series with 100 points than with 40 points, as is to be expected. Excluding the specific cases that are noted below, however, results for 40 point series were reasonable.

Estimates of EMIN were generally lower than the population value, as is to be expected due to shrinkage. The four parameter solutions, however, underestimated EMIN to a greater degree than the two parameter solutions. This would probably result in a notably greater number of Type 1 errors for four parameter solutions.

Additionally, it was seen that the introduction of a four parameter analysis, i.e., estimating for L, DL, S, and DS, significantly deteriorated the consistency of estimates of level and change in level. It was shown in the Results section that this was not due to the introduction of the slope and change in slope effects themselves, but was specifically due to the mechanics of the analysis. It remains unclear, however, just exactly how this occurs.

More specifically, in a number of cases intervention parameter estimates were seen to be either moderately to highly inaccurate (the mean estimate differed from the criterion value), or inconsistent (there was considerable variability), or both. A summary of the cases for which an inaccurate estimate may be obtained due to either the

inaccuracy of the mean estimate or inconsistency is as follows:

- 1) Estimates of all intervention parameters for the differencing and (2,0,0) .8/.2 models; across all conditions (i.e., across series length and number of parameters estimated).
- 2) Estimates of level (L) for all autoregressive models except the (2,0,0) -.8/-.4 model; across all conditions.
- 3) Estimates of change in level (DL) for all autoregressive models except the (2,0,0) -.8/-.4 model; across all conditions.
- 4) Estimates of change in level (DL) for the white noise model with L, DL, S, and DS estimated.
- 5) Estimates of change in slope (DS) for all autoregressive models except the (2,0,0) -.8/-.4; for 40 point series.

These cases represent a significant portion of the cases investigated in this study. Specifically, estimates for 20 of 36 cases for L, 30 of 44 cases for DL, and 10 of 22 cases for DS are considered by the author to be notably inaccurate or inconsistent. Of further note is that the two most important parameters, i.e., DL and DS, are both affected. These are usually considered to be the more important parameters because they specifically represent the effect or change due to the intervention as opposed to the baseline or

preintervention parameters L and S. Researchers are typically more interested in the change due to an intervention than in the preexisting conditions.

The difficulties with the differencing models were discussed in Harrop and Velicer (1985) where it was shown that the equation for the (2,0,0) .8/.2 model was similar to that for the (0,1,0) model. The current model has ruled out several causal hypotheses, i.e., first, the difficulties were not caused by differencing not being accounted for. A differencing model was identified to TSX, SAS, and BMDP, as appropriate to the series. Second, it was not caused by an artifact of the TSX solution since both SAS and BMDP experienced similar difficulties. Third, a preliminary study by the author considered and rejected the possibility that the outlier created at the intervention point when the series is differenced first, adversely affects the subsequent analysis. This hypothesis was investigated by differencing a limited number of simulated first order differencing series first, removing the outlier, and then analyzing the resultant series. The results were equally problematic.

No further insight into the cause of difficulties with differencing and pseudo-differencing models has been gained, therefore, beyond Harrop and Velicer's (1985) statement that perhaps "once instability has been introduced into a series, it will be difficult, if not impossible, to remove," (p. 42). It has been demonstrated in the current study,

however, that the use of the General Transformation Approach's recommended (5,0,0) model in lieu of the (3,0,0) approximation used by Harrop and Velicer (1985) is not the solution to the problem.

As regards the difficulties with estimates of L, DL, and DS for all of the autoregressive models except the (2,0,0) $-.8/-.4$ model, it appears that this is caused by the nature of a high positive autoregressive dependency. First, the equations for the problematic (2,0,0) cases can be shown to be similar to the (1,0,0) case, particularly the (1,0,0) $.8$ case. Second, the effect of positive autoregressive dependency is to amplify the "positiveness" or "negativeness" of a series and to provide inertia for that direction. Third, it takes a rather large random error to change this direction. And finally, once this direction is changed, it will be sustained until another rather large random error occurs. The net effect is to introduce a type of instability that is quite similar to a second order differencing process.

Harrop and Velicer (1985) found similar results for their comparable analyses (see Appendix A, note that their study did not include interventions with slope and change in slope). It must be noted, however, that their results for the (0,0,1) $.4$ and (2,0,0) $.8/-.4$ models were not as inaccurate as those for the current study, and that their results for autoregressive models with dependencies not

included in this study were more accurate in general than for the aberrant cases. This latter fact keeps us from generalizing the results of this study to all autoregressive models.

The GENTS solutions were essentially the same as the TSX or best solutions, with the exception that GENTS did not provide an estimate of the minimum residual error variance. This makes it impossible to make comments regarding statistical tests of significance when they are implemented for GENTS. However, it may be noted that the EMINS found by two different solutions of the models, i.e., by TSX, and SAS, did not differ significantly. It is highly likely, therefore, given that the GENTS intervention parameter estimates were so similar to those of TSX and SAS, that the GENTS implementation will be as accurate.

The information regarding the necessity of model identification is still mixed. One statement that may be made is that, given that one is comfortable with the accuracy of the solution that would be obtained with the true model identification, an equivalent solution may be obtained without identifying the model, i.e., by using the General Transformation Approach as implemented by the GENTS program. This statement applies to the conditions included in this study. One must be wary, however, of the solutions obtained by either method under the conditions specified above. Unfortunately, this last statement requires that at

least a rough model identification be performed in order to determine whether or not one's data falls into one of the inaccurate solution categories, e.g., that it is best represented by a differencing model. The results of this study indicate that distinguishing differencing from non-differencing models is of greater import than distinguishing autoregressive from moving average models. It is possible that the former distinction may be easier to make.

The author's final conclusion, therefore, is that it is best to perform the model identification step. GENTS, however, is a useful tool in evaluating this stage as discussed in the Recommendations section along with other recommendations for current analyses.

Computer Programs

Regarding the accuracy and/or consistency of estimates of EMIN and the intervention parameters (L, DL,S, and DS), all of the statements made under Model Identification about TSX apply to both SAS and BMDP. There are, however, some notable additions regarding BMDP. BMDP's estimates of EMIN were often lower than for the other programs which might result in a greater number of Type 1 errors. Generally speaking, of the intervention parameter estimates deemed adequate by the author, in a number of cases the BMDP solutions were not quite as accurate or consistent as those of TSX, GENTS, and SAS. Additionally, four outliers were identified, and in a number of cases solutions were not

provided (i.e., no solutions were found for 40 point (0,0,1) .8 series with L, DL, and DS estimated; 1 of 10 solutions were found for the 40 point (0,0,1) .4 series with L, DL, S, and DS estimated; 1 of 10 solutions were found for each of the 100 point (0,0,1) .8 and (1,0,0) .8 series with L, DL, S, and DS estimated; 6 of 10 solutions were found for the 100 point (2,0,0) .8/.2 series with L and DL estimated; and 9 of 10 solutions were found for the 100 point (2,0,0) .8/.2 series with L, DL, S, and DS estimated). The four outliers occurred when BMDP completed its solution after only two or three iterations because of a small relative change in the parameter estimates. This indicates that the default value for this stopping criterion may be too large. All of the non-solutions stopped with the message "HIGH CORRELATION AMONG PARAMETER ESTIMATES, ESTIMATION IS TERMINATED," which is not explained in the BMDP manual. It would seem, however, that there is some test in the program that is too sensitive. This conclusion is arrived at by observing that, in most cases, one or more of the 10 series were evaluated; all of the series were valid ARIMA series; and all of the other programs completed their solutions with the same data.

The ITSE program is in a category all by itself since only three of its five estimates included herein have direct criterion values. As observed, its estimates of EMIN were consistently less than for any of the other three programs.

In addition, there were some values that were sufficiently lower than the criterion to cause concern as regards finding significance in statistical tests of the parameters. This would be of greatest concern with shorter series where primarily a change in level is anticipated, since the degrees of freedom would be further reduced by two for the unnecessary estimates of pre- and postintervention slope. ITSE's estimates of S and DS were consistently and highly inaccurate except for the white noise model. Additionally, although there were no direct criterion values to compare ITSE's estimates of the pre- and postintervention intercepts to, some approximate comparisons noted in the Results section indicated that they were probably inaccurate as well. This is not surprising given that slope estimates affect intercept estimates in a linear regression computation and that ITSE's S and DS estimates were inaccurate. Finally, it should be noted that ITSE was provided with a (5,0,0) identification for the two purely moving averages models. This may have affected the accuracy of those solutions. It would not account for the inaccuracy of estimates for the remaining six autoregressive models, however.

Regarding CPU time requirements, it is clear that GENTS and TSX, in that order, use considerably more time than the other three programs. GENTS uses almost three minutes of CPU time, e.g., for 100 point series. In the opinion of the

author, however, none of the programs requires a sufficient amount of CPU time to discount it, and he would not select or exclude any of these programs based on CPU time.

ITSE has the best documentation of the five programs from the point of view of providing input, knowing what the solution calculates, and reading the output. TSX and GENTS require reading a textbook and journal article, respectively, in order to know what is being calculated, and to read the commented text at the beginning of each program's Fortran source listing to determine the input required, a formidable task without assistance. (Neither ITSE, TSX, nor GENTS explains how to compile and execute the program, however.)

SAS and BMDP both have extensive manuals. As regards providing input for the simple type of analyses presented herein, however, they are both sorely lacking. Documentation on the equations solved is either unclear or nonexistent for both programs. However, both clearly define the output that will be obtained.

Additionally, the documentation for the three nonlinear estimation programs, i.e, GENTS, SAS, and BMDP, does not provide explanations of the causes of various error or warning messages and suggestions on how to avoid them. General information regarding the pitfalls of nonlinear estimation techniques, which would be most helpful, is also not furnished. SAS comes the closest to providing this

latter information currently.

All five of the programs desperately need a primer describing how to specify a simple interrupted time series of the type used in this study. Of further note is that Academic Computer Center personnel generally can not provide assistance in this area because they are not familiar with interrupted time series analysis.

The computational features contained or not contained in the programs best discriminates between them. SAS and BMDP provide the best data handling and the only general services for identifying models and diagnosing the residuals. Specifically, for any time series, they allow one to plot it; and to compute and plot autocorrelations, inverse autocorrelations (SAS only), and partial autocorrelations. These series may be, for example, the input series, a differenced series, or the residual series. It should be noted that one could obtain the residual series from TSX, for example, and use these SAS or BMDP features. Additionally, one could difference an input series with SAS or BMDP and output it to analyze with ITSE.

TSX is best distinguished by its use of the grid search technique. The advantage of this technique over the nonlinear techniques is that it is much less likely to find a solution that is not in the correct region. Additionally, recalling that solutions are obtained across the entire

range of the invertibility/stationarity region(s), one may scan the output looking for other areas with a similarly low EMIN in order to make an assessment of whether or not a relative minimum has been found. One disadvantage of TSX's grid search technique is that the absolute minimum may actually occur between the values of dependency used; between ϕ_1 equals .90 and .80, for example, if an iteration increment of .10 is used. Another less consequential disadvantage is the significantly larger amount of CPU time required by TSX than is required by SAS, BMDP, or ITSE.

GENTS, SAS, and BMDP all implement numerical nonlinear estimations. All three also provide for outputting diagnostic information as regards the adequacy of the estimation. They do not, however, discuss how to use this information. The author, being unfamiliar with the details of nonlinear estimation, is not able to distinguish between them on this basis. GENTS is most notably distinguished by the fact that it does not require that the model be identified. And, as noted previously, it provides results comparable to TSX and SAS (both good and bad as also noted). Of the two most generally available programs, i.e., SAS and BMDP, SAS was more consistently accurate. Also, the author prefers the option offered by SAS which allows restricting the solutions to within the stationarity/invertibility

region(s). This option is not offered by BMDP and GENTS. Finally, the problems that BMDP has with outliers and with not providing solutions in a large number of cases, although not a computational feature per se, deserve to be mentioned here again.

ITSE uses an autoregressive model only and linear estimation of the pre- and postintervention intercepts and slopes. This is as opposed to the variety of intervention parameters (conventional and non-conventional) that the other programs allow the user to specify. The author strongly prefers the more conventional forms (as used in this study, for example) over ITSE's parameterization because these conventional forms generally fit his conceptualization of intervention effects more closely.

Recommendations

The suggestions presented here fall into three categories, i.e., recommendations for current analyses, recommendations for improvements to computer programs, and recommendations for future research. Recommendations result from this study, Harrop and Velicer (1985), and the author's previous experience with time series analysis.

The principal recommendation for current analyses is not to trust results for data that has differencing in the model, and to be wary of results for the autoregressive models with inaccurate results specified previously. Unfortunately, this requires performing the model identification

step. SAS and BMDP should assist in this process and there are guidelines provided in Glass, Willson, and Gottman (1975), McCleary and Hay (1980), and Gottman (1981).

Further, the author recommends that TSX, GENTS, and SAS analyses all be performed on the same data, for comparative purposes. The TSX analysis would serve as a check against SAS's nonlinear procedure finding a relative minimum. In the event that the TSX and SAS solutions were in different dependency regions, starting values could be provided to SAS from the TSX solution, for example, which might result in SAS converging in the TSX region. The most correct solution, of course, would be the one with the smallest residual error variance. The GENTS solution would serve as a check against an incorrect model identification being provided to TSX and SAS; i.e., if the GENTS solution is quite different from TSX and SAS solutions which are similar to each other, the model identification provided to TSX and SAS might be in error. This is based on the observation that TSX, GENTS, and SAS generally provided equivalent results in this study.

Additionally, it must be emphasized that one must be careful, regardless of the program or programs used, not to specify an intervention form with L, DL, S, and DS when only interested in or expecting L and DL. As noted, the L and DL estimates are deteriorated by the addition of estimates of S and DS to the analysis.

It has long been recommended that series be made as long as feasible, and this study does not contradict that. It has shown, however, that shorter series (approximately 40 points) are appropriate candidates for analysis, although estimations may not be quite as accurate as for longer series. Model identification would also be more difficult. More specifically, the choice of intervention form is more critical for shorter series than for longer ones because, in addition to the deterioration of estimates discussed above, the loss of degrees of freedom due to estimating unnecessary parameters may result in not achieving significance.

Once an analysis is complete, the residual series should be evaluated in order to determine that it has, in fact, been reduced to white noise (indicating that all of the dependency has been removed). This evaluation is performed using the same procedure as in identifying the model. This may be particularly helpful if the three proposed analyses are divergent.

Additionally, one may wish to consult the tables provided herein and in Harrop and Velicer (1985). If the analyst's solution is similar to one of the table entries; i.e., for model, degree of dependency, and form of intervention; the tables might provide an indication of the relative accuracy of the solution. This latter recommendation requires further study and would only be valid if it were known that solutions were estimating the degree of

dependency fairly closely.

Preliminary studies indicate that SAS analyses with a (5,0,0) identification might be a reasonable substitute for GENTS if it is not available. These same studies do not indicate that BMDP used in the same manner is a reasonable substitute. BMDP may be used as a substitute for SAS if SAS is not available. There is no guarantee that a solution will be found, however, and an early termination of the estimation should be watched for. Finally, the author does not recommend ITSE for reasons noted elsewhere.

Recommendations for improvements to computer programs are numerous. All, except ITSE, need to improve and/or provide documentation. Specifically required is a clear explanation of the inputs required to analyze interrupted time series, with examples, with more information about forms of interventions, and, as appropriate, with more information about the pitfalls and diagnosing of the nonlinear estimation technique used. It would also be most helpful if SAS and BMDP would include guidelines for model identification, diagnosing the nonlinear solution process, and evaluating the adequacy of the ARIMA model.

Again, all programs except ITSE, should make it easier to specify analyses. SAS and BMDP in particular require that a design matrix be provided for all forms of intervention whereas TSX and GENTS at least will generate this matrix for the two interventions used in this study. All of

these programs should allow for these types and others to be selected without having to provide a design matrix. The most immediate need, in addition to the two forms used in this study, would be for an S and DS alone type intervention, and for combinations of levels and slopes across multiple interventions, as in an ABA experimental design for example.

More specifically regarding individual programs, GENTS, of course, requires the addition of an estimate of the minimum residual error variance, tests of significance for the intervention parameters, white noise test(s) for the residual series, outputting of the residual series, and a printout of the number of iterations performed. BMDP requires that the problems specified earlier be fixed, and both SAS and BMDP require that the number of iterations performed be output for solutions when the nonlinear solution diagnostic options are not specified.

The following recommendations apply to a "system" program which would include many options for model identification, analysis, and diagnosis. Currently, only SAS and BMDP fit this definition. Therefore, these recommendations are directed primarily toward these programs. First, the author would like to see these programs add a "TSX" option, i.e., a grid search algorithm, and a "GENTS" option (including the different nonlinear estimation as well as the General Transformation Approach equations). These would be

available for comparative purposes as recommended earlier, but would be all in one package. Second, a simulation option might be helpful. Essentially, the user would specify the number of replications and the program would perform a simulation (as in the current study) using the model and intervention type previously specified for the solution, and the degree of dependency found by the solution. Basically, this would replace the table lookup recommendation made earlier and provide the analyst with an approximation to the degree of confidence he could place in his solution (subject to the previous caveat). Third, an output plot option which plots the series with a plot of the estimated intervention parameters superimposed would be quite helpful for interpreting one's output. Two superimposed solutions, one with all of the estimated parameters and one with the nonsignificant ones deleted, would also be helpful. The analyst would be required to specify the probability desired to reach statistical significance when specifying this option. Fourth, and last, the author would like to be able to specify a probability for significance and an intervention form; and have the program perform the analysis and test the intervention parameters for significance. The program would then automatically delete any nonsignificant parameters and reanalyze the data with the new "form." Both solutions would be presented for comparative purposes.

There are numerous possible recommendations for future research. Only the major ones are presented here.

- 1) The numerous techniques for model identification mentioned above should be evaluated and compared to each other toward the goal of establishing a unified and comprehensive strategy for identifying time series. Special consideration should be placed on identifying shorter series and on distinguishing series with differencing from those without.
- 2) The accuracy of estimation of the dependency parameters should be evaluated with the goal of determining the feasibility of using simulations based on a given solution to make probabilistic statements about the accuracy of estimation of intervention parameters.
- 3) Other available programs should be evaluated as were the five in this study.
- 4) The reasons for the inaccuracy of differencing solutions and the other inaccurate models specified earlier should be investigated. Possibly, alternative modeling techniques for these cases should be explored and evaluated. One possibility is to use a (0,0,5) model for analyzing the autoregressive models.

5) Finally, the author would like to see Type 1 and Type 2 Monte Carlo investigations performed with various forms and magnitude of intervention parameters. The ultimate goal would be tables of various intervention types and magnitudes (relative to the variance of the time series) that would indicate the Type 1 and Type 2 errors likely to be encountered. These tables would serve as guidance both before and after experimentation, e.g., before to suggest to an experimenter the size of intervention effect he should be attempting to create in order to find significance, and after, to investigate the possibility that a nonsignificant but reliable intervention effect was not significant because of a lack of power in the experimental design.

References

- Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-723.
- Baer, D.M. (1977). Perhaps it would be better not to know everything. Journal of Applied Behavior Analysis, 10, 167-1 .
- Beguin, J.M., Gourieroux, C., & Monfort, A. (1980). Identification of a mixed autoregressive - moving average process: The corner method. In O.D. Anderson (Ed.), Time Series: Proceedings of the International Conference held at Nottingham University, March 1979 (pp. 423-436). Amsterdam: North-Holland.
- Bhansali, R.J. & Downham, D.Y. (1977). Some properties of the order of an autoregressive model selected by a generalization of Akaike's FPE-Criterion. Biometrika, 64, 547-551.
- Bower, C. & Glass, G.V. (1974). TSX Computer program . University of Colorado, Boulder.
- Box, G.E.P., & Jenkins, G.M. (1970). Time-series analysis: Forecasting and control. San Francisco: Holden Day.
- Box, G.E.P., & Tiao, G.C. (1965). A change in level of a nonstationary time-series. Biometrika, 52, 181-192.

- Box, G.E.P., & Tiao, G.C. (1975). Intervention analysis with application to economic and environmental problems. Journal of the American Statistical Association, 70, 70-92.
- Campbell, D.T. (1969). Reforms as experiments. American Psychologist, 24, 409-429.
- Campbell, D.T., & Stanley, J.C. (1979). Experimental and Quasi-experimental designs for research. Chicago: Rand McNally, & Co.
- Cook, T.D., & Campbell, D.T. (Eds.). (1979). Quasi-experimentation: Design & analysis issues for field settings. Chicago: Rand McNally College Publishing.
- DeProspero, A. & Cohen, S. (1979). Inconsistent visual analyses of intrasubject data. Journal of Applied Behavioral Analysis, 12, 573-579.
- Deutsch, S.J. (1979). Lies, damn lies, and statistics: A rejoinder to the comment by Hay and McCleary. Evaluation Quarterly, 3, 315-328.
- Deutsch, S.J., & Alt, F.B. (1977). The effect of Massachusetts' gun control law on gun-related crimes in the city of Boston. Evaluation Quarterly, 1, 543-568.
- Dixon, W.J. (1985). BMDP statistical software. Computer program manual . Berkeley:University of California Press.

- Glass, G.V., Willson, V.L., & Gottman, J.M. (1975). Design and Analysis of Time Series Experiments. Boulder, Colorado: Colorado Associated University Press.
- Gottman, J.M. (1981). Time-series analysis: A comprehensive introduction for social scientists. New York: Cambridge University Press.
- Gottman, J.M., & Glass, G.V. (1978). Analysis of interrupted time-series experiments. In Kratochwill (Ed.), Single subject research: Strategies for evaluating change. New York: Academic Press.
- Gray, H.L., Kelley, G.D., & McIntire, D.D. (1978). A new approach to ARMA modeling. Communications in Statistics, B7, 1-77.
- Hannan, E.J. & Rissanen, J. (1982). Recursive estimation of mixed autoregressive moving average order. Biometrika, 69, 81-94.
- Harrop, J.W., & Velicer, W.F. (1985). A comparison of alternative approaches to the analysis of interrupted time-series. Multivariate Behavioral Research, 20, 27-44.
- Hay, R.A. Jr., & McCleary, R. (1979). Box-Tiao time series model for impact assessment: a comment on the recent work of Deutsch and Alt. Evaluation Quarterly, 3, 277-314.

- Hersen, M., & Barlow, D.H. (1976). Single case experimental designs: Strategies for studying behavior change. New York: Pergamon Press.
- Jones, R.R., Weinrott, M.R., & Vaught, R.S. (1978). Effects of serial dependency on the agreement between visual and statistical inference. Journal of Applied Behavior Analysis, 11, 277-283.
- Kashyap, R.L. (1977). A Bayesian comparison of different classes of dynamic models using empirical data. IEEE Transactions on Automatic Control, 22, 715-727.
- Kratochwill, T.R. (1978). (Ed.) Single subject research: Strategies for evaluating change. New York: Academic Press.
- Ljung, G.M., & Box, G.E.P. (1978). On a measure of lack of fit in time series models. Biometrika, 65, 297-303.
- Marmor, G.S., & Marmor, M. (1978). Comment on Simonton's "Cross-sectional time-series experiments: Some suggested statistical analyses." Psychological Bulletin, 85, 1102-1103.
- Marsh, J.C. & Shibano, M. (1984). Issues in the statistical analysis of clinical time-series data. Social Work Research and Abstracts, 20 (4), 7-12.

- McCain, C.J., & McCleary, R. (1979). The statistical analysis of the simple interrupted time-series quasi-experiment. In T.D. Cook & D.T. Campbell (Eds.), Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.
- McCleary, R., & Hay, R.A., Jr. (1980). Applied time series analysis for the social sciences. Beverly Hills, CA: Sage.
- McDowall, D., McCleary, R., Meidinger, E.E., & Hay, R.A., Jr. (1980). Interrupted time series analysis. Beverly Hills, CA: Sage.
- Miller, M.J. (1985). Analyzing client change graphically. Journal of Counseling and Development, 63, 491-494.
- Padia, W.L. (1973). Effect of autocorrelation on probability statements about the mean. Unpublished master's thesis, University of Colorado, Boulder.
- Padia, W.L. (1975). The consequences of model misidentification in the interrupted time-series experiment. Dissertation Abstracts International, 36, 4875A. (University Microfilms No. 76-3938).
- Parsonson, B.S., & Baer, D.M. (1978). The analysis and presentation of graphic data. In T.R. Kratochwill (Ed.), Single subject research: Strategies for evaluating change. New York: Academic Press.

- Parzen, E. (1974). Some recent advances in time series modelling. IEEE Transactions on Automatic Control, 19, 723-729.
- Pukkila, T.M. (1982). On the identification of ARMA (p,q) models. In O.D. Anderson (Ed.), Time Series Analysis: Theory and Practice 1: Proceedings of the International Conference held at Valencia, Spain, June 1981 (pp. 81-103). Amsterdam: North-Holland.
- Revenstorff, D., Kessler, A., Schindler, L., Hahlweg, K., & Bluemner (1980). Time series analysis: Clinical applications evaluating intervention effects in diaries. In O.D. Anderson (Ed.), Analysing Time Series: Proceedings of the International Conference held on Guernsey, Channel Islands, October, 1979 (pp. 291-312). Amsterdam: North-Holland.
- SAS Institute (1984). SAS/ETS user's guide, version 5 edition Computer program manual . Cary, NC:SAS Institute Inc.
- Schwartz, G. (1978). Estimating the dimension of a model. Annals of Statistics, 6, 461-464.
- Simonton, D.K. (1977). Cross-sectional time-series experiments: Some suggested statistical analyses. Psychological Bulletin, 84, 489-302.
- Tsay, R.S. (1984). Regression models with time series errors. Journal of the American Statistical Association, 79, 118-124.

- Tsay, R.S. & Tiao, G.C. (1984). Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary ARMA models. Journal of the American Statistical Association, 79, 84-96.
- Velicer, W.F., Fraser, C., McDonald, R.P., & Harrop, J.W. (1980). GENTS Computer program . University of Rhode Island, Kingston.
- Velicer, W.F., & Harrop, J.W. (1983). The reliability and accuracy of time series model identification. Evaluation Review, 7, 551-560.
- Velicer, W.F., & McDonald, R.P. (1984). Time series analysis without model identification. Multivariate Behavioral Research, 19, 33-47.
- Williams, E.A., & Gottmann, J.M. (1982). A user's guide to the Gottman-Williams times-series analysis computer programs for social scientists Computer program manual . Cambridge: Cambridge University Press.

Multivariate Behavioral Research, 1985, 20, 27-44

A Comparison of Alternative Approaches to the Analysis of Interrupted Time-Series

John W. Harrop
Wayne F. Velicer

University of Rhode Island

Computer generated data representative of 26 ARIMA models was used to compare the results of interrupted time-series analysis using: (1) the known model identification, (2) an assumed (1, 0, 0) model, and (3) an assumed (3, 0, 0) model as an approximation to the General Transformation approach. Ten samples each of 26 ARIMA processes were generated representing ARIMA (0, 0, 0), (0, 0, 1), (1, 0, 0), (0, 1, 0), and (0, 1, 1) models with varying degrees of dependency. Series were generated with both 40 points and 100 points, with an immediate and constant intervention effect of magnitude one (occurring at the midpoint), and with a random error component variance of magnitude one. The results of the three analyses are compared to the criterion values and to each other for the point estimates of: 1) the minimum residual error variance; 2) the pre-intervention level; and 3) the post-intervention change in level. The results for models which included a difference parameter were inaccurate for all three approaches. All three analyses provided reasonable and equivalent results for the remaining or non-differencing models for series with both 40 and 100 points. The findings strongly suggest that the problematic model identification process might be eliminated entirely and replaced with the assumed (1, 0, 0) approach or the General Transformation approach.

Observations made across time on a single observational unit (or subject) are quite common in the social and behavioral sciences. When an event or intervention occurs that may affect our dependent measure, we refer to the resultant as an interrupted time-series. Analyses of these interrupted time-series in the past have often consisted of either graphical methods or applications of traditional statistical procedures such as *t*-tests comparing the pre- and post-intervention means of the series. Both of these approaches are invalid (see for example Jones, Weinrott, & Vaught, 1978, and Padia, 1973) because they ignore the problem of dependency (i.e. autocorrelation or the correlation between the repeated measurements of the time-series). Statistical procedures have been recently developed that account for this dependency.

To date the most promising and widespread statistical analysis of interrupted time-series from the social and behavioral sciences is that proposed by Glass, Willson, and Gottman (1975). The Glass et al. technique is based on earlier work by Box and Tiao (1965) and Box and

This work was partially supported by Grant CA 27821 from the National Cancer Institute and computer time provided by the University of Rhode Island. A previous version of this paper was presented at the 1980 Annual Convocation of the Northeastern Educational Research Association. Reprint requests should be addressed to Wayne F. Velicer, Department of Psychology, University of Rhode Island, Kingston, RI 02881. The authors would like to thank David Kenny and Joseph Rossi for their comments on an earlier version of this paper.

Jenkins (1976) as are similar approaches discussed by McCleary and Hay (1980) and Gottman (1981). This technique employs the Auto Regressive Integrated Moving Averages (ARIMA) model and is an adaptation of spectral analysis used in the physical sciences with continuous time-series data to the case of discrete time-series data. A Glass, Willson, and Gottman (1975) analysis first requires that the analyst "identify the model" by determining the following: the proper values of p , the order of the autoregressive component(s); d , the order of differencing; and q , the order of the moving averages component(s). Their method of identifying the ARIMA (p, d, q) model requires inspection of autocorrelations and partial autocorrelations computed from the time-series at various lags. Simply stated, a researcher "fits" his data with a particular form of the ARIMA model which accounts for the observed dependency in the series. With the use of the identified ARIMA model, the data is then transformed to an independent series that meets the assumptions of the General Linear Model. A regression solution is then obtained, along with associated test statistics for the regression parameters. In effect, we are using the pre-intervention data to predict where we would be after intervention and then comparing this to where we actually were after intervention. The regression parameters tested commonly include pre-intervention level, pre-intervention slope, post-intervention change in level, and post-intervention change in slope. Other forms of post-intervention change may also be tested statistically.

The most problematic part of the above process is the model identification step. First, the requirement of a minimum of 50 observations before intervention and 50 observations after intervention (Box & Jenkins, 1976; Glass, Willson, & Gottman, 1975) is frequently prohibitive in applied settings. Applied settings, such as single subject clinical psychology designs or single unit program evaluation studies, represent some of the largest areas of potential applications for this method. Second, the model identification step requires the most mathematical sophistication on the part of the user. Mastery of complex mathematics is a practical hurdle to the widespread employment of any procedure. Third, even when an adequate number of data points is available and an adequate degree of mathematical sophistication has been achieved, the model is still likely to be incorrectly identified. Velicer and Harrop (1983) studied the accuracy of model identification employing 12 subjects who had been extensively trained in the Glass, Willson, and Gottman (1975) procedure for model identification. The subjects were only able to correctly identify 36% of 16 different series with 100 points each and 20% of 16 series with 40

points each. An example of the practical problems likely to occur is provided by the recent model identification controversy on analysis of the effects of the Massachusetts Gun Control Law (Deutsch, 1979; Deutsch & Alt, 1977; and Hay & McCleary, 1979).

Two proposals have been made recently to eliminate the model identification step entirely. Simonton (1977) has proposed analyses of cross-sectional time-series data which *assume* that the series are represented by an ARIMA (1, 0, 0), i.e. a first order autoregressive model. Velicer and McDonald (1984) have proposed fitting the time-series with a fifth order autoregressive model, (5, 0, 0), and have demonstrated theoretically that this solution will satisfy a large proportion of the time-series example commonly encountered.

Simonton's (1977) proposal to assume an ARIMA (1, 0, 0) model for all analyses is based first on necessity. Since he is dealing with cross-sectional time-series with a small number of points, i.e. typically "between 4 and 12 observations on 20 or more statistical cases" (1977, p. 490), he cannot use the Glass, Willson and Gottman (1975) approach. He has asserted that the (1, 0, 0) model is the one having the greatest a priori plausibility and "the first-order autoregressive process was chosen as the empirically most common and the substantively most reasonable for behavioral and social data (cf. Glass et al., 1975, pp. 116-118)." (Simonton, 1977, p. 499). Simonton also comments that the consequences of this a priori assumption being incorrect are unknown.

Simonton is in error on two counts, both regarding his reference to Glass, Willson, and Gottman (1975, pp. 116-118). First, the (1, 0, 0) model was found to be the second most common model (21 of 95). The most common model was the (0, 1, 1) found in 22 of 95 cases. Second, Glass et al., did not state that the autoregressive model was more reasonable than the moving averages model for the social sciences. What they did state was that generally *third or higher order* moving averages processes were rarely found in a number of social and behavioral studies that they examined.

Velicer and McDonald (1984) have examined the transformation necessary for changing the dependent series Z_t into the independent series Y_t for use with the General Linear Model. This transformation may be represented as

$$[1] \quad Y = TZ$$

where Y is the vector of transformed observations, Z is the vector of observations, and T is the transformation matrix.

If we let a design matrix X indicate the presence or absence of the

"form" parameters i.e. level and change in level, and if $\mathbf{b}$ is the vector of parameters to be estimated, then in terms of the GLM

$$[2] \quad Y = TXb + a$$

where $\mathbf{a}$ is a vector representing the random error component of the series which is NID $(0, \sigma^2)$. The transformation matrix T contains only functions of ϕ s and/or θ s which depend on the ARIMA model, whereas X and $\mathbf{b}$ are model independent (they are dependent on the form of the intervention only).

After comparing the equivalent elements of T for the $(0, 0, 0)$, $(0, 0, 1)$, $(0, 1, 1)$, and $(k, 0, 0)$ models, Velicer and McDonald concluded that the transformation matrix for the $(5, 0, 0)$ model would satisfactorily represent T for most applications. The use of an autoregressive model of high order is referred to here as the General Transformation Approach.

The present study compared the results of analyses performed on computer-generated interrupted time-series (1) using the correct model identification, (2) assuming a $(1, 0, 0)$ ARIMA model as suggested by Simonton, and (3) assuming a $(3, 0, 0)$ ARIMA model as a substitute for the General Transformation Approach. One purpose of the study was to determine if either or both of the "assumed model" proposals would perform as well as or superior to the correct model identification. This would justify the elimination of the model identification step, and therefore greatly increase the number of applications of the Glass et al. (1975) method of interrupted time-series analysis by reducing both the degree of statistical sophistication required and the number of observations required.

The present study can also be viewed as one of relatively few that provides empirical evidence about the effects of model misidentification. The simulation study by Padia (1973) can be viewed as assessing the effects of misidentifying time series models as $(0, 0, 0)$ models. The results are compelling that this is an inappropriate procedure. Zinkgraf & Willson (1981) investigated a limited number of models— $(1, 0, 0)$, $(0, 0, 1)$, and $(1, 0, 1)$ —and assessed the effect of analyzing data generated from each of the three models by the three other models and a white noise model. They concluded that the $(1, 0, 0)$ model was the "... most resistant to statistical cost due to misidentification" (Zinkgraf and Willson, 1981, p. 651). This provides some justification for the Simonton approach.

The present study differs from the Padia (1973) and Zinkgraf and Willson (1981) studies in a number of ways. First, the alternative models are not selected blindly from the class of possible models. The

two alternative models represent two approaches that are both potentially robust to model misidentification and represent two alternatives that are theoretically justified. Second, the number of models and the levels of dependency investigated are much broader than previously considered. Third, the study will focus on parameter estimates rather than the Type I error rates. In error rate simulation studies, if the alpha-level is inaccurate, the reader typically does not know the reason for the inaccuracy. Possible reasons include: (1) the error rate is systematically underestimated; (2) the intervention parameter is systematically under- or overestimated; (3) the parameter estimates are excessively variable.

Several different purposes will be served by the present study. All models will be analyzed with the correct model identification. This represents an idealized variation of the Glass, Willson, and Gottman (1975) approach, with all assumptions underlying the procedure met exactly. This represents the first broad empirical evaluation of this procedure. The study will also provide an empirical evaluation of the Simonton (1977) procedure. Some limited evidence in support of this procedure is provided by Zinkgraf and Willson (1981). An empirical evaluation of an approximation of the Velicer and McDonald (1984) General Transformation Approach is also provided. Program limitations did not permit an exact evaluation of the procedure. In addition, a broad based comparison of these three procedures will assist in selecting among them. Finally, some evidence will be provided indirectly relating to the more general issue of the effects of model misidentification.

Method

The simulated time-series used in this study were generated by a Fortran computer program developed by Padia (1975) and modified for use on the computer employed in this study. Additionally, the random number generator used was the GGNML subroutine of the International Mathematical and Statistical Libraries (IMSL) Edition 8 instead of the GGNOR subroutine of IMSL Edition 3 used by Padia.

Table 1 presents the characteristics of the simulated time-series. A total of 52 series were generated (10 samples each) and analyzed three times, once by each approach. Before discussing the rationale for selecting these particular series, we will discuss the analyses performed.

The analysis of each series was performed using the TSX computer program developed by Glass, Bower, and Padia (1974). TSX requires

Table 1
Characteristics of Simulated Time-Series

MODEL	β or θ values	N of Series
(0,0,0)	no β s or θ s in this model	1
(0,0,1)	$\theta_1 = .8, .4, -.4, \text{ or } -.8$	4
(1,0,0)	$\beta_1 = .8, .4, -.4, \text{ or } -.8$	4
(0,1,0)	no β s or θ s in this model	1
(0,1,1)	$\theta_1 = .8, .4, -.4, \text{ or } -.8$	4
(2,0,0)	$\beta_1 / \beta_2 = .8/.2, -.8/.2, .4/.4, -.4/.4, .8/-.4, -.8/-.4, .8/-.2, -.8/-.2, .4/-.4, -.4/-.4, .2/-.4, \text{ or } -1.2/-.4$	12

All series were generated using the following values:

- A. pre-intervention level (L) = 0.
- B. change in level post-intervention (D) = 1.
- C. mean of random component ($\bar{X}$) = 0.
- D. variance of random component (σ^2) = 1.

Series generated with 40 points ($n_1 = n_2 = 20$) and 100 points ($n_1 = n_2 = 50$).

Ten (10) samples generated for each process.

that the analyst provide it with the proposed form of the intervention effect and the model identification. For all three analyses (referred to as IDENT, AUTO-1, and GENERAL), the form identified was a constant pre-intervention level (L), and an immediate and constant post-intervention change in level (D). What differed across the three analyses was the model identification. For the IDENT approach, the model specified was the correct one (the one used to generate the series). Thus, the results obtained from this analysis were the best that could be obtained, i.e. they represent the results that would be obtained if analysts were 100% accurate in identifying the model. For the AUTO-1 approach the model specified was the (1, 0, 0) as suggested by Simonton (1977). For the GENERAL approach, the model specified was the (3, 0, 0) as a substitute for the (5, 0, 0) suggested by Velicer and

McDonald (1984). This substitution was required because TSX is not programmed to analyze a series with autoregressive order greater than three. Finally, all of the GENERAL approach analyses were performed using grid search iteration increments for ϕ s of .10, as were IDENT analyses of all (2, 0, 0) series with 100 points. All other analyses were performed with a .02 increment. The reason for this change is that the analyses involving higher order autoregressive terms would have required a prohibitive amount of computer time if performed using a .02 increment.

The rationale for the series generated, Table 1, will be presented in the following order: (1) the model parameters; (2) the ϕ or θ values; (3) the form of the series; and (4) the number of points in the series. There were ten samples generated for each series.

The models selected were found by Glass et al. in 80 of 95 cases in an empirical study (1975, pp. 115-118). Thus, they are representative of series that one can expect to encounter commonly in the social and behavioral sciences. Of the remaining 15 cases of 95 found by Glass et al. that were not simulated in this study, four were identified as (0, 1, 2); two each were identified as (0, 0, 2), (0, 2, 0), (0, 2, 1), (0, 2, 2), and (0, 1, 3); and one was identified as (1, 1, 0).

The values of the ϕ s and θ s were selected to be representative of the range of ϕ s and θ s that are valid for the ARIMA model. The valid ranges ($-1 < \phi$ or $\theta < 1$) for these coefficients were determined by Box and Jenkins (1976) in accordance with the ARIMA model restrictions that the series be finite and stationary.

The form of the series generated is specified by the pre-intervention level (L), the post-intervention change in level (D), the mean of the random component (0), and the variance of the random component (σ^2).

Random component refers to a_t in the ARIMA model. The form selected for this study was a constant pre-intervention level of zero and an immediate and constant intervention effect of 1. The choice of $L = 0$ was made because it simplified input to the generation program and the analysis of results. The intervention effect $D = 1$ was chosen for similar reasons. The variance of a_t ($\sigma^2 = 1$) was chosen relative to the value of D in order to generate series capable of testing the sensitivity of the three model identification approaches for detecting a moderate intervention effect. The values of σ^2 and D represent a relative intervention effect that is large enough to be detectable by a statistical analysis and yet small enough to be missed by a graphical analysis, the situation in which the ARIMA model analysis of interrupted time-series is most needed.

The next item to be discussed is the number of points generated for

each series. The value of 100 ($n_1 = n_2 = 50$) was chosen because this is the minimum number suggested as being necessary for the accurate identification of models (Box and Jenkins, 1976; Glass, Willson, and Gottman, 1975). The intent was to determine whether or not the simplified AUTO-1 and GENERAL approaches are as effective as the IDENT approach in the case where the latter is deemed to be accurate. The value 40 ($n_1 = n_2 = 20$) was chosen to be representative of the cases likely to be encountered in applied research. These are cases for which the GWG approach may not be viable, but are a type of case for which the AUTO-1 and GENERAL approaches were developed.

Results

Results are presented in tables 2 through 11. Tables 2 through 5 and 7 and 8 are identical in format and present the mean ($\bar{X}$) and standard deviation (s) across 10 samples for each ARIMA model and analysis approach i.e. IDENT, AUTO-1 and GENERAL. Tables 2 and 3 present the results for the minimum residual error variance (EMIN) for series with 40 and 100 points, respectively. In like manner, Tables 4 and 5 present results for the pre-intervention level of the series (L) and Tables 7 and 8 for the post-intervention change in level (D). Additionally, these six tables present the averages of $\bar{X}$ and s across the 26 models used in the simulation.

Tables 6 and 9 through 11 will be described below.

Minimum Residual Error Variance (EMIN)

Tables 2 and 3 present the EMINs of the analyses for series of 40 and 100 points respectively. The criterion or reference value for EMIN (this corresponds to the variance of the random component, i.e. a_t , used in generating the series) is 1. It is to be expected that a least squares analysis approach would recover a value that is slightly less than the actual values used because it may capitalize on chance. For the same reason, smaller values are to be expected for series with 40 points than with 100 points.

In general all three approaches recovered an error variance close to 1 with IDENT and GENERAL being generally less than 1 and with AUTO-1 generally being greater than 1 especially for the 12 (2, 0, 0) models. Although the average of the means for series with 40 points was lower across all three approaches than for series with 100 points, this is offset by the increased average variation of the estimates. None of the approaches appeared to have any relative difficulty with a "type" of model, i.e. autoregressive, differencing, or moving averages.

Some isolated cases of relative difficulty exist as in the AUTO-1

Table 2

Estimates of Minimum Residual Error Variance for N = 40

Model	Dependency		IDENT		AUTO-1		GENERAL	
	Lag 1	Lag 2	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
(0,0,0)	0	0	.97	.22	.94	.21	.90	.20
(0,0,1)	.8	0	.96	.24	1.27	.29	1.03	.22
(0,0,1)	.4	0	.93	.23	1.00	.22	.92	.20
(0,0,1)	-.4	0	.93	.19	.95	.20	.90	.19
(0,0,1)	-.8	0	.95	.18	1.15	.26	.95	.21
(1,0,0)	.8	0	.89	.20	.89	.20	.86	.21
(1,0,0)	.4	0	.92	.20	.92	.20	.89	.20
(1,0,0)	-.4	0	.97	.22	.97	.22	.91	.19
(1,0,0)	-.8	0	1.03	.26	1.03	.26	.98	.24
(0,1,0)	0	0	.92	.19	.88	.17	.90	.19
(0,1,1)	.8	0	.87	.16	.83	.15	.80	.14
(0,1,1)	.4	0	.87	.15	.88	.14	.81	.12
(0,1,1)	-.4	0	.39	.19	1.03	.28	1.13	.36
(0,1,1)	-.8	0	.88	.18	1.46	.48	1.65	.66
(2,0,0)	.8	.2	.85	.15	.90	.18	.86	.16
(2,0,0)	-.8	.2	1.25	.47	1.28	.49	1.55	.63
(2,0,0)	.4	.4	.86	.20	.98	.27	.86	.20
(2,0,0)	-.4	.4	.96	.22	1.08	.36	.94	.23
(2,0,0)	.8	-.4	.90	.20	1.11	.29	.91	.20
(2,0,0)	-.8	-.4	.98	.17	1.24	.29	.96	.17
(2,0,0)	.8	-.2	.88	.20	.96	.22	.88	.20
(2,0,0)	-.8	-.2	.97	.19	1.07	.23	.96	.19
(2,0,0)	.4	-.4	.93	.20	1.12	.27	.93	.20
(2,0,0)	-.4	-.4	.94	.19	1.18	.29	.92	.19
(2,0,0)	.2	-.4	.93	.20	1.13	.26	.93	.20
(2,0,0)	-1.2	-.4	1.10	.29	1.32	.32	1.09	.29
Averages			.94	.21	1.06	.26	.98	.24

and GENERAL results for the (0, 1, 1) $\theta_1 = -.8$ model and in the results from all approaches for the (2, 0, 0) $\phi_1 = -.8$, $\phi_2 = .2$. The AUTO-1 approach has relatively more difficulty with (2, 0, 0) models as is to be expected since it is only fitting one parameter to a two parameter model.

It should be noted here that the GENERAL approach for all cases and the IDENT approach for (2, 0, 0) models with 100 points would most likely have found solutions with smaller EMINs had the analysis been performed using iteration increments of .02 as opposed to .10.

Estimate of Pre-intervention Level (L)

Tables 4 and 5 present the estimates of L of the analyses for series of 40 and 100 points respectively. The criterion value for L (this

Table 3
Estimates of Minimum Residual Error Variance for N = 100

Model	Dependency		IDENT		AUTO-1		GENERAL	
	Lag 1	Lag 2	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
(0,0,0)	0	0	.96	.14	.94	.14	.92	.13
(0,0,1)	.8	0	.95	.13	1.22	.18	1.02	.14
(0,0,1)	.4	0	.94	.14	.98	.14	.92	.13
(0,0,1)	-.4	0	.95	.13	.98	.15	.93	.13
(0,0,1)	-.8	0	.97	.12	1.22	.22	1.01	.14
(1,0,0)	.8	0	.98	.14	.98	.14	.96	.14
(1,0,0)	.4	0	.95	.13	.95	.13	.94	.13
(1,0,0)	-.4	0	.94	.14	.94	.14	.92	.13
(1,0,0)	-.8	0	.95	.14	.95	.14	.93	.13
(0,1,0)	0	0	.94	.14	.95	.16	1.07	.28
(0,1,1)	.8	0	.94	.14	.96	.18	.92	.15
(0,1,1)	.4	0	.94	.14	1.05	.17	.96	.17
(0,1,1)	-.4	0	.93	.14	1.12	.23	1.42	.46
(0,1,1)	-.8	0	.93	.14	1.60	.35	2.10	.74
(2,0,0)	.8	.2	1.02	.23	.98	.16	1.01	.23
(2,0,0)	-.8	.2	1.37	.43	1.07	.18	1.37	.43
(2,0,0)	.4	.4	.95	.14	1.09	.16	.95	.14
(2,0,0)	-.4	.4	.95	.13	1.09	.22	.94	.13
(2,0,0)	.8	-.4	.96	.12	1.19	.22	.96	.12
(2,0,0)	-.8	-.4	.93	.11	1.13	.16	.92	.12
(2,0,0)	.8	-.2	.96	.12	1.03	.16	.96	.12
(2,0,0)	-.8	-.2	.94	.13	.99	.13	.93	.13
(2,0,0)	.4	-.4	.94	.13	1.18	.23	.93	.13
(2,0,0)	-.4	-.4	.92	.12	1.15	.20	.91	.12
(2,0,0)	.2	-.4	.93	.14	1.17	.23	.93	.13
(2,0,0)	-1.2	-.4	.98	.16	1.15	.16	.96	.16
Averages			.97	.15	1.08	.18	1.03	.19

corresponds to the value used in generating the series) is zero.

In general all three approaches consistently slightly underestimated L. Estimates for series with 100 points were further away on the average than for series with 40 points. All three approaches had considerable difficulty with differencing models, i.e. (0, 1, 0) and (0, 1, 1) models, and the (2, 0, 0) $\phi_1 = .8$, $\phi_2 = .2$ model. Table 6 presents the averages of $\bar{X}$ and s for each approach across (1) the 5 differencing models, (2) the remaining 21 models, and (3) all 26 models. Examination of Table 6 shows that all approaches are considerably more accurate in estimating L for non-differenced models. Except for differencing models and the (2, 0, 0) $\phi_1 = .8$, $\phi_2 = .2$ model there is essentially no difference in accuracy of estimate between the three

Table 4

Estimates of Pre-intervention Level for $N = 40$

Model	Dependency		IDENT		AUTO-1		GENERAL	
	Lag 1	Lag 2	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
(0,0,0)	0	0	-.03	.17	-.03	.17	-.03	.17
(0,0,1)	.8	0	.01	.05	.01	.05	.01	.05
(0,0,1)	.4	0	-.01	.10	-.01	.10	-.01	.10
(0,0,1)	-.4	0	-.07	.23	-.04	.24	-.05	.23
(0,0,1)	-.8	0	-.09	.28	-.07	.32	-.08	.30
(1,0,0)	.8	0	-.12	1.02	-.12	1.02	-.20	.81
(1,0,0)	.4	0	-.06	.28	-.06	.28	-.03	.29
(1,0,0)	-.4	0	-.02	.12	-.02	.12	-.02	.12
(1,0,0)	-.8	0	-.01	.10	-.01	.10	-.01	.10
(0,1,0)	0	0	-.31	3.92	-.65	4.25	-.82	4.39
(0,1,1)	.8	0	-.10	1.16	-.25	.98	-.23	1.01
(0,1,1)	.4	0	-.22	2.83	-.73	2.72	-.56	2.72
(0,1,1)	-.4	0	-.47	5.14	-.83	5.63	-1.27	5.88
(0,1,1)	-.8	0	-1.16	6.53	-1.03	6.88	-1.71	7.34
(2,0,0)	.8	.2	-.42	3.77	-.42	3.74	-.68	3.75
(2,0,0)	-.8	.2	-.02	.11	-.02	.11	-.02	.11
(2,0,0)	.4	.4	.08	1.14	-.22	.76	-.03	1.02
(2,0,0)	-.4	.4	-.02	.17	-.04	.17	-.03	.17
(2,0,0)	.8	-.4	-.08	.24	-.08	.29	-.08	.25
(2,0,0)	-.8	-.4	-.01	.08	.00	.08	.00	.08
(2,0,0)	.8	-.2	-.12	.37	-.10	.46	-.13	.35
(2,0,0)	-.8	-.2	.00	.08	.00	.09	-.01	.08
(2,0,0)	.4	-.4	-.04	.17	-.02	.19	-.04	.17
(2,0,0)	-.4	-.4	-.01	.10	-.01	.09	-.02	.10
(2,0,0)	.2	-.4	-.03	.15	-.02	.15	-.03	.15
(2,0,0)	-1.2	-.4	.00	.07	.00	.07	.00	.07
Averages			-.13	1.09	-.18	1.12	-.23	1.15

approaches. All three approaches do poorly with differencing models, with GENERAL doing the worst.

Estimate of Post-intervention Change in Level (D)

Tables 7 and 8 present the estimates of D for series with 40 and 100 points respectively. The criterion value for D is 1.

Estimates of D were more accurate on the average for series with 100 points than for series with 40 points. All approaches were considerably less accurate with the differencing models and the $(2, 0, 0)$ $\phi_1 = .8$, $\phi_2 = .2$ model. Table 9 presents the averages of $\bar{X}$ and s for each approach across 1) the 5 differencing models, 2) the remaining 21 models, and 3) all

Table 5
Estimates of Pre-intervention Level for N = 100

Model	Dependency		IOENT		AUTO-1		GENERAL	
	Lag 1	Lag 2	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
(0,0,0)	0	0	-.01	.12	-.01	.12	-.01	.12
(0,0,1)	.8	0	.00	.03	.01	.04	.00	.03
(0,0,1)	.4	0	.00	.08	.00	.08	.00	.08
(0,0,1)	-.4	0	-.01	.16	-.02	.16	-.02	.16
(0,0,1)	-.8	0	-.02	.21	-.03	.21	-.03	.21
(1,0,0)	.8	0	-.31	.83	-.31	.83	-.30	.79
(1,0,0)	.4	0	-.02	.18	-.02	.18	-.03	.18
(1,0,0)	-.4	0	-.01	.09	-.01	.09	-.01	.09
(1,0,0)	-.8	0	-.01	.07	-.01	.07	-.01	.07
(0,1,0)	0	0	-2.24	4.05	-2.35	3.90	-2.61	3.89
(0,1,1)	.8	0	-.64	.94	-.63	.95	-.62	.95
(0,1,1)	.4	0	-1.40	2.56	-1.52	2.08	-1.61	2.38
(0,1,1)	-.4	0	-3.21	5.33	-3.30	5.35	-3.72	5.45
(0,1,1)	-.8	0	-4.45	6.68	-4.33	6.75	-4.84	7.00
(2,0,0)	.8	.2	-2.14	3.22	-2.04	3.24	-2.15	3.25
(2,0,0)	-.8	.2	-.02	.08	-.02	.08	-.02	.08
(2,0,0)	.4	.4	-.31	.63	-.17	.49	-.31	.63
(2,0,0)	-.4	.4	-.02	.12	-.02	.12	-.02	.12
(2,0,0)	.8	-.4	-.03	.19	-.04	.19	-.03	.19
(2,0,0)	-.8	-.4	.00	.06	.00	.06	.00	.06
(2,0,0)	.8	-.2	-.07	.31	-.07	.29	-.07	.31
(2,0,0)	-.8	-.2	.00	.06	.00	.06	.00	.06
(2,0,0)	.4	-.4	-.01	.13	.00	.12	-.01	.13
(2,0,0)	-.4	-.4	.00	.07	.00	.07	.00	.07
(2,0,0)	.2	-.4	.01	.11	.00	.10	.01	.11
(2,0,0)	-1.2	-.4	.01	.05	-.01	.05	.00	.05
Averages			-.58	1.01	-.57	.99	-.63	1.02

26 models. Examination of Table 9 reveals that all approaches are considerably less accurate in estimating D for models with differencing than for models without. The accuracy of estimate for non-differencing models is essentially the same for all three approaches.

Discussion

This study compared the accuracy of parameter estimation of three different approaches to analysis of interrupted time-series. This was the first time that the GENERAL approach was studied empirically and the first extensive empirical evaluation of the IDENT and

Table 6

Means and Standard Deviations of Estimates
of Pre-intervention Level Across Models

Number of Points	Model	IDENT		AUTO-1		GENERAL	
		$\bar{X}$	$\bar{S}$	$\bar{X}$	$\bar{S}$	$\bar{X}$	$\bar{S}$
N = 40	Difference Models	-.45	3.92	-.70	4.09	-.92	4.27
	Non-Difference Models	-.05	.42	-.06	.41	-.07	.40
	All Models	-.13	1.09	-.18	1.12	-.23	1.15
N = 100	Difference Models	-2.39	3.92	-2.43	3.81	-2.68	3.93
	Non-Difference Models	-.14	.32	-.13	.32	-.14	.32
	All Models	-.58	1.01	-.57	.99	-.63	1.02

AUTO-1 approaches. In general, all three approaches provided reasonable and accurate estimations of the error variance (EMIN), the pre-intervention level (L), and the post-intervention effect (D). In general, all three approaches provided equivalent results. One notable exception to these observations is a possible trend seen with the AUTO-1 estimates of EMIN for the (2, 0, 0) models. The AUTO-1 estimates are higher than the others which would tend to result in a larger standard error for estimates of L and D and therefore a reduced likelihood for finding statistical significance.

Another notable exception to the above observations is the estimation of L and D for all three approaches when dealing with the differencing i.e. (0, 1, 0) and (0, 1, 1) models, and with the (2, 0, 0) $\phi_1 = .8$, $\phi_2 = .2$ model. All approaches fared consistently poorly with these models. It appears that the (2, 0, 0) model is in this grouping because of its similarity to the (0, 1, 0) model. The equation for the (2, 0, 0) is

$$[3] \quad Z_t = a_t + .8Z_{t-1} + .2Z_{t-2},$$

and for the (0, 1, 0) is

$$[4] \quad Z_t = a_t + Z_{t-1}.$$

Table 7
Estimates of Intervention Effect for N = 40

Model	Dependency		IDENT		AUTO-1		GENERAL	
	Lag 1	Lag 2	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
(0,0,0)	0	0	.99	.21	.99	.22	.99	.22
(0,0,1)	.8	0	.98	.04	.95	.08	.97	.06
(0,0,1)	.4	0	.98	.12	.97	.13	.99	.12
(0,0,1)	-.4	0	1.04	.34	.99	.38	1.01	.31
(0,0,1)	-.8	0	1.07	.45	1.00	.55	1.05	.39
(1,0,0)	.8	0	1.23	1.03	1.23	1.03	1.20	1.02
(1,0,0)	.4	0	1.00	.42	1.00	.42	.96	.41
(1,0,0)	-.4	0	.98	.14	.98	.14	.99	.14
(1,0,0)	-.8	0	.98	.11	.98	.11	.98	.12
(0,1,0)	0	0	1.49	1.40	1.29	1.53	1.19	2.10
(0,1,1)	.8	0	1.02	.85	.95	.90	.92	.95
(0,1,1)	.4	0	1.35	1.28	1.10	2.24	.98	1.89
(0,1,1)	-.4	0	1.50	1.16	1.48	1.59	1.25	2.50
(0,1,1)	-.8	0	.72	.70	1.61	1.82	1.16	3.19
(2,0,0)	.8	.2	1.63	2.13	1.52	2.00	1.16	1.92
(2,0,0)	-.8	.2	.99	.15	.99	.14	1.00	.15
(2,0,0)	.4	.4	1.19	.96	1.16	1.03	1.14	1.01
(2,0,0)	-.4	.4	.96	.25	1.00	.20	.99	.21
(2,0,0)	.8	-.4	1.06	.33	1.04	.54	1.05	.33
(2,0,0)	-.8	-.4	.98	.09	.97	.08	.98	.09
(2,0,0)	.8	-.2	1.09	.51	1.07	.74	1.10	.52
(2,0,0)	-.8	-.2	.98	.10	.97	.09	.98	.10
(2,0,0)	.4	-.4	1.01	.21	.96	.31	1.01	.21
(2,0,0)	-.4	-.4	.98	.12	.97	.12	.99	.11
(2,0,0)	.2	-.4	1.00	.18	.96	.24	1.00	.18
(2,0,0)	-1.2	-.4	.98	.07	.97	.07	.98	.07
Averages			1.08	.51	1.08	.64	1.04	.70

The lack of a difference parameter might explain the failure of the AUTO-1 approach for this data. The GENERAL approach should, theoretically at least, accurately handle data from a model with a differencing parameter (Velicer & MacDonald, 1984). The fact that the GENERAL approach was approximated by a (3, 0, 0) model instead of the recommended (5, 0, 0) model may have contributed to its poor showing in these cases. Velicer and McDonald (1984) recommend tests on the residuals, followed by even higher order transformations. This might be appropriate for the troublesome cases presented here. The IDENT approach, however, explicitly involved a difference parameter. The poor results from data analyzed employing this approach, therefore, are the most troublesome. One potential explanation is that the

Table 8

Estimates of Intervention Effect for N = 100

Model	Dependency		IDENT		AUTO-1		GENERAL	
	Lag 1	Lag 2	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
(0,0,0)	0	0	1.03	.12	1.03	.12	1.03	.12
(0,0,1)	.8	0	1.00	.03	1.00	.04	1.00	.04
(0,0,1)	.4	0	1.01	.08	1.01	.08	1.01	.08
(0,0,1)	-.4	0	1.03	.17	1.04	.17	1.04	.17
(0,0,1)	-.8	0	1.04	.23	1.06	.22	1.05	.22
(1,0,0)	.8	0	1.31	.61	1.31	.61	1.24	.55
(1,0,0)	.4	0	1.05	.19	1.05	.19	1.06	.19
(1,0,0)	-.4	0	1.02	.09	1.02	.09	1.02	.09
(1,0,0)	-.8	0	1.02	.07	1.02	.07	1.02	.07
(0,1,0)	0	0	1.01	1.04	.98	1.06	.94	2.30
(0,1,1)	.8	0	1.00	.43	1.07	1.17	1.05	1.10
(0,1,1)	.4	0	1.06	.71	.75	1.24	1.00	1.43
(0,1,1)	-.4	0	.92	.92	1.02	.90	.96	3.01
(0,1,1)	-.8	0	.96	.57	1.04	.83	.97	3.77
(2,0,0)	.8	.2	.94	2.01	.93	1.09	.96	1.99
(2,0,0)	-.8	.2	1.03	.09	1.03	.09	1.03	.09
(2,0,0)	.4	.4	1.28	.41	1.27	.50	1.27	.39
(2,0,0)	-.4	.4	1.04	.12	1.04	.12	1.04	.12
(2,0,0)	.8	-.4	1.06	.21	1.06	.20	1.06	.21
(2,0,0)	-.8	-.4	1.01	.06	1.01	.06	1.01	.06
(2,0,0)	.8	-.2	1.12	.32	1.11	.28	1.12	.31
(2,0,0)	-.8	-.2	1.01	.07	1.01	.07	1.01	.07
(2,0,0)	.4	-.4	1.02	.13	1.02	.13	1.02	.13
(2,0,0)	-.4	-.4	1.01	.08	1.01	.08	1.01	.08
(2,0,0)	.2	-.4	1.02	.11	1.01	.12	1.02	.11
(2,0,0)	-1.2	-.4	1.01	.05	1.01	.05	1.01	.05
Averages			1.04	.34	1.03	.37	1.04	.65

data analysis procedures are unable to differentiate between stochastic drift, the type involved here, and deterministic drift. Judd and Kenny (1981) have suggested the inclusion of slope parameters as a method for handling data with stochastic drift. Following this suggestion, we reanalyzed our data employing both a slope and change of slope parameter for the five models which involved a differencing parameter and the (2, 0, 0) model discussed above. The results were slightly different, but not more accurate than the results previously presented.

The results of our analysis with the differencing model are sufficiently poor to suggest that an entirely different modeling approach might be necessary. One alternative is that the GENERAL approach with a higher order will handle this problem. Another, more

Table 9

Means and Standard Deviations of Estimates
of Intervention Effect Across Models

Number of Points	Model	IDENT		AUTO-1		GENERAL	
		$\bar{X}$	$\bar{S}$	$\bar{X}$	$\bar{S}$	$\bar{X}$	$\bar{S}$
N = 40	Difference Models	1.21	1.08	1.29	1.62	1.10	2.12
	Non-Difference Models	1.05	.38	1.03	.41	1.02	.37
	All Models	1.08	.51	1.08	.64	1.04	.70
N = 100	Difference Models	.99	.73	.97	1.04	.99	2.32
	Non-Difference Models	1.05	.25	1.05	.21	1.05	.25
	All Models	1.04	.34	1.03	.37	1.04	.65

pessimistic view is that once instability is introduced into a series, it will be difficult, if not impossible, to remove. At a minimum, potential analysts must be wary of results obtained from series in which a differencing process is potentially involved.

Results for 40 point series were generally as good as, and in some cases, superior to results for 100 point series. On an average, results were better for 40 point series when estimating L and better for 100 point series when estimating D. Estimates of EMIN were lower on an average for 40 point series but more widely dispersed.

These findings strongly suggest that the model identification process might be eliminated entirely in favor of the AUTO-1 or GENERAL approaches. The reason for equivocation in the preceding statement is that this study was not all inclusive. Further areas remain to be investigated before a final definitive recommendation to eliminate the model identification process may be made.

Additional areas of investigation are: (1) the effect of pre-intervention slope and post-intervention change in slope; (2) performance with types of models not included in this study such as the (1, 1, 0) and differencing and moving averages models of order 2 and higher; (3) standard errors of estimate and Type I error rates; (4) power with

intervention effects lesser and greater than 1; (5) the effect of using the (5, 0, 0) model recommended for the GENERAL approach as opposed to the (3, 0, 0) used in this study; and (6) the effect of a more accurate algorithm for finding the minimum residual error variance. The biggest obstacle to studies of this kind is the excessive amount of computer time required. The rapid improvement of computers coupled with more efficient computer programs should soon remove this obstacle.

With the restrictions discussed above in mind, the following tentative conclusions may be reached. Both the AUTO-1 and GENERAL approaches would be preferred to the IDEN approach for the following reasons: (1) These approaches are more flexible, i.e. they apply to a wider variety of potential problems; (2) The troublesome model identification step is bypassed; and (3) Simpler procedures are viewed as generally preferable to complex procedures when they result in essentially identical results. Additionally, the GENERAL approach would be preferred to the AUTO-1 for the following reasons: (1) The rationale provided for the GENERAL approach is stronger than that for the AUTO-1 approach; (2) Some limited support was provided for the assertion that the AUTO-1 approach would be inadequate for higher order models; and (3) The GENERAL approach has a greater potential for adaptation to complex designs such as cross sectional and multivariate problems.

References

- Box, G. E. P., & Jenkins, G. M. (1976). *Time-series analysis: Forecasting and control*. (Rev. Ed.) San Francisco: Holden Day.
- Box, G. E. P., and Tiao, G. C. (1965). A change in level of a nonstationary time-series. *Biometrika*, 52, 181-192.
- Deutsch, S. J. (1979). Lies, damn lies, and statistics: A rejoinder to the comment of Hay and McCleary. *Evaluation Quarterly*, 3, 315-328.
- Deutsch, S. J., & Alt, F. J. (1977). The effect of Massachusetts' gun control law on gun-related crimes in the city of Boston. *Evaluation Quarterly*, 1, 543-568.
- Glass, G. V., Bower, C., & Padia, W. L. (1974). *Computer program TSX*. University of Colorado, Boulder, CO.
- Glass, G. V., Willson, V. L., & Gottman, J. M. (1975). *Design and analysis of time-series experiments*. Boulder, Colorado: Colorado Associated University Press.
- Gottman, J. M. (1981). *Time-series analysis*. New York: Cambridge University Press.
- Hay, R. A., Jr., & McCleary, R. (1979). Box-Tiao time-series models for impact assessment: A comment on the recent work of Deutsch and Alt. *Evaluation Quarterly*, 3, 277-314.
- Jones, R. R., Weinrott, M., & Vaught, R. S. (1978). Effects of serial dependency on the agreement between visual and statistical inference. *Journal of Applied Behavior Analysis*, 11, 277-283.
- Judd, C. M., & Kenny, D. A. (1981). *Estimating the effects of social interventions*. New York: Cambridge University Press.

John W. Harrop and Wayne F. Velicer

- McCleary, R., & Hay, R. A., Jr. (1980). *Applied time series analysis for the social sciences*. Beverly Hills, CA: SAGE.
- Padia, W. L. (1973). Effect of autocorrelation on probability statements about the mean. Masters thesis, Laboratory of Educational Research, University of Colorado.
- Padia, W. L. (1975). The consequences of model misidentification in the interrupted time-series experiment (Doctoral dissertation, University of Colorado). *Dissertation Abstracts International*, 1976, 36, 4875a. (University Microfilms No. 76-3938.)
- Simonton, D. K. (1977). Cross-sectional time-series experiments: Some suggested statistical analyses. *Psychological Bulletin*, 84, 489-502.
- Velicer, W. F., & Harrop, J. W. (1983). The reliability and accuracy of the time-series model identification. *Evaluation Review*, 7, 551-560.
- Velicer, W. F., & McDonald, R. P. (1984). Time-series analysis without model identification. *Multivariate Behavioral Research*, 19, 33-47.
- Zinkgraf, S. A., & Willson, V. L. (1981). The use of the Box-Jenkins approach in causal modeling: An investigation of the cost of the misidentification of selected stationary models. In Anderson, O. D., & Perryman, M. R. (Eds.), *Time Series Analysis*. New York: North-Holland.

Appendix B: Simulation Program

N.B.: This program is designed to run with programs TSX, GENTS, and ITSE which have been modified to run as subroutines. These programs are available elsewhere and are copyrighted, so they are not included here.

```
C*****
C*****      TIME-SERIES GENERATION PROGRAM      *****
C*****      LAST REVISED 10/29/86 BY JOHN HARROP *****
C*****
C****THIS PROGRAM HAS 5 COMPONENTS:      **
C**** 1)SIMWITH:MAIN PROGRAM WHICH ACCEPTS INPUTS AND GENERATES SERIES**
C****      SERIES GENERATION WAS ADAPTED FROM PADIA'S DISSERTATION.      **
C**** 2)CORREL:MODIFIED TO RUN AS A SUBROUTINE & ALLOW LESS PRINTOUT. **
C**** 3)TSX:   MODIFIED TO RUN AS A SUBROUTINE & ALLOW LESS PRINTOUT. **
C**** 4)GENTS: MODIFIED TO RUN AS A SUBROUTINE & ALLOW LESS PRINTOUT. **
C**** 5)ITSE:  MODIFIED TO RUN AS A SUBROUTINE & ALLOW LESS PRINTOUT. **
C*****
C**** THIS PROGRAM REQUIRES USE OF IMSL ROUTINES GGNML,BECORI,      ***
C****      LINV3F,VMULFF, & VMULFM.      ***
C****      IMSL SINGLE PRECISION IS USED.      ***
C*****
C-----INPUT INSTRUCTIONS---INPUT INSTRUCTIONS---INPUT INSTRUCTIONS-----
C*****
C CARD 1: COLS 1-80-HED1(20A4),PROBLEM HEADING
C CARD 2: COLS 1-80-HED2(20A4),PROBLEM HEADING CONTINUED
C*****
C CARD 3: COLS 1- 5 - NIT (15), THE NUMBER OF ITERATIONS
C          COLS 6- 10 - NASS (15),NUMBER OF ASSUMED MODELS
C          COLS 11-15 - NREAL (15),NUMBER OF REAL CASES
C          COLS 17-20 SELECT ANALYSES TO BE DONE IN ADDITION TO GENERATE.
C          COL 17      - DOCORR (15)=1, DO CORREL.
C          COL 18      - DOTSX  (15)=1, DO TSX.
C          COL 19      - DOGENT (15)=1, DO GENTS100.
C          COL 20      - DOITSE (15)=1, DO ITSE.
C          COLS 21-25 - OUTFRM (15) = 0 - MINIMUM PRINT OPTION
C                   - OUTFRM (15) = 1 - PRINT RANDOM NUMBERS
C                   - OUTFRM (15) = 2 - PLOT ALSO
C                   - OUTFRM (15) = 3 -PUNCH ALSO(TO UNIT 7 IN 10F8.3)
C          COLS 26-30 - ZWRITE (15) = 1 -WRITE SERIES TO UNIT 8 IN 10F8.3
C                   THIS REQUIRES A 'GO.FT08F001 DD' STATEMENT.
```

```

C          AS CURRENTLY IMPLEMENTED, IT IS PREFERRED TO
C          USE A LRECL OF 80.
C      COLS 31-35 - BWRITE (15) = 1 WRITE TSX SOLUTIONS TO UNIT 9,
C          GENTS SOLUTIONS TO UNIT 10, AND ITSE SOLUTIONS
C          TO UNIT 11.
C          THIS REQUIRES 'GO.FT0??F001 DD' STATEMENTS.
C          AS CURRENTLY IMPLEMENTED, IT IS PREFERRED TO USE
C          A LRECL OF 80.SEE END OF INPUT SPECS FOR FORMAT.
C      COLS 36-40 - ICODE (15) = 0,2 PARAMETERS ESTIMATED
C          ICODE (15) =1,4 PARAMETERS ESTIMATED
C          ICODE (15) =2,M PARAMETERS ESTIMATED
C      COLS 41-45 - ISPEC (15)= 0,NO 'ASSUMED' VALUES TO TSX
C          ISPEC (15)= 1,ALL CARD 5 VALUES SENT TO TSX
C          ISPEC (15)= 2, ONLY P,D,&Q VALUES ON CARD 5 SENT
C      COLS 46-50 - IDV (15),INC IN HUNDREDTHS FOR PHI&THETA ITERATIO
C      COLS 51-55 - MAXFUN (15)= 0(DEFAULT), MAX NUMBER OF FUNCTION
C          EVALUATIONS TO BE PERFORMED BY GENTS. THE
C          DEFAULT IS 100.
C      COLS 56-60 - NORD (15)= THE ORDER OF THE AUTOREGRESSIVE MODEL
C          TO BE USED BY ITSE.
C*****
C CARD 4, COLS 1- 10 - SEED (D10.5),START VALUE OF RANDOM NUMBER,0 TO 1
C      COLS 11-20 - XMEAN (F10.5), LEVEL OF THE SERIES
C      COLS 21-30 - SIGMA (F10.5),THE S. D. OF RANDOM COMPONENT
C      COLS 31-40 - DELTA (F10.5),THE INTERVENTION EFFECT
C      COLS 41-50 - SLOPE (F10.5),PRE-INTERVENTION SLOPE(TANGENT)
C      COLS 51-60 - DSLOPE (F10.5),POST-INTERVYN CHANGE SLOPE(TANGENT)
C      COLS 61-65 - N1 (15),NUMBER OF PRE-1 POINTS
C      COLS 66-70 - N2 (15),NUMBER OF POST-1 POINTS
C*****
C CARD 5 VALUES THIS CARD INPUT TO GENERATION PROGRAM.
C CARD 5 COLS 1- 5 - P (15),          TRUE P
C          6-10 - D (15),          TRUE D
C          11-15 - Q (15),          TRUE Q
C          16-20 - PHI(1) (F5.2), TRUE PHI(1)
C          21-25 - PHI(2) (F5.2), TRUE PHI(2)
C          26-30 - PHI(3) (F5.2), TRUE PHI(3)
C          31-35 - THETA(1) (F5.2),TRUE THETA(1)
C          36-40 - THETA(2) (F5.2),TRUE THETA(2)
C          41-45 - THETA(3) (F5.2),TRUE THETA(3)
C*****
C CARD 6:THIS CARD READ AND DATA SENT TO TSX IF ISPEC = 1 OR 2.
C CARD 6:COLS 1- 5 - PSTAR(15): ASSUMED P
C          6-10 - DSTAR(15): ASSUMED D
C          11-15 - QSTAR(15): ASSUMED Q
C          16-20 - PHSTAR(1)-(F5.2): ASSUMED PHI1 (ONLY IF ISPEC = 1)
C          21-25 - PHSTAR(2)-(F5.2): ASSUMED PHI2 (ONLY IF ISPEC = 1)
C          26-30 - PHSTAR(3)-(F5.2): ASSUMED PHI3 (ONLY IF ISPEC = 1)
C          31-35 - THSTAR(1)-(F5.2): ASSUMED TH1 (ONLY IF ISPEC = 1)
C          36-40 - THSTAR(2)-(F5.2): ASSUMED TH2 (ONLY IF ISPEC = 1)
C          41-45 - THSTAR(3)-(F5.2): ASSUMED TH3 (ONLY IF ISPEC = 1)
C*****

```

```

C-----TSX-----TSX-----TSX-----TSX-----
C-----OUTPUT FORMAT---OUTPUT FORMAT---OUTPUT FORMAT---OUTPUT FORMAT-----
C*****
C FORMAT OF UNIT 9 OUTPUT-----THIS IS FOR EACH NASS X NREAL RUN.-----
C RECORDS 1 TO 6- REPRODUCE CARDS 1 TO 6 OF THE INPUT TO THIS PROGRAM.
C RECORD 7- ISUMKE(15),THIS IS THE NUMBER OF TIMES THAT TSX FOUND MORE
C          THAN ONE MINIMUM ERROR VARIANCE SOLUTION.
C RECORD 8-12(F5.3,1X) THE NUMBER OF TIMES THAT THE ESTIMATES OF LEVEL,
C          LEVELCHG,SLOPE, & SLOPECHG WERE SIGNIFICANT
C          AT THE P=.1,.05, & .01 LEVELS RESPECTIVELY,
C          DIVIDED BY NIT(THE NUMBER OF ITERATIONS).
C          THE ORDER IS L.10,L.05,L.01,DL.10,DL.05 ETC.
C RECORD 9 - (F10.5,2(F10.5,F12.8,F8.2))
C MEAN OF: EMIN & (ESTIMATE,STAND. ERR.,& T-STAT OF LEVEL & LEVELCHG)
C RECORD 10- (F10.5,2(F10.5,F12.8,F8.2))
C SIGMA OF: EMIN & (ESTIMATE,STAND. ERR.,& T-STAT OF LEVEL & LEVELCHG)
C RECORD 11- (F10.5,2(F10.5,F12.8,F8.2))
C MEAN OF: Z0 & ( ESTIMATE,STAND. ERR.,& T-STAT OF SLOPE & SLOPE CHG)
C RECORD 12- (F10.5,2(F10.5,F12.8,F8.2))
C SIGMA OF: Z0 & ( ESTIMATE,STAND. ERR.,& T-STAT OF SLOPE & SLOPE CHG)
C RECORD 13-(5F12.6) MEAN OF: SUMSQ,SS1,SS2,SS3,SS4
C RECORD 14-(5F12.6) SIGMA OF: SUMSQ,SS1,SS2,SS3,SS4
C THIS IS FOLLOWED BY WRITING THE RESULTS FROM EACH ITERATION, NIT TIMES.
C RECORD '15'-(11,1X,12,1X,12,1X,14,6F10.5)
C          KE(THE # OF EMIN SOLUTIONS FOUND BY TSX),NRLCNT,NASCNT,
C          NITCNT,PHI1,PHI2,PHI3,THETA1,THETA2,& THETA3.
C RECORD '16'-(F10.5,2(F10.5,F12.8,F8.2))
C          EMIN & (ESTIMATE,STAND. ERR.,& T-STAT OF LEVEL & LEVEL CHG)
C RECORD '17'-(F10.5,2(F10.5,F12.8,F8.2))
C          Z0 & ( ESTIMATE,STAND. ERR.,& T-STAT OF SLOPE & SLOPE CHG)
C RECORD '18'-(5F12.6) SUMSQ,SS1,SS2,SS3,SS4
C-----END OF UNIT 9 OUTPUT FORMAT-----
C-----
C*****
C-----
REAL HED1(20),HED2(20),Z(300),PHI(3),THETA(3),THSTAR(3),PHSTAR(3),
1XMEAN,SIGMA,Z0,SLOPE,DSLOPE,
2PARMS(1000,25),XPARMS(1000,19),XM(19),S(19),R(190),
3GBETA(4),BGENTS(1000,4),XBGENT(1000,4),GXM(4),GS(4),GR(10),
4IBETA(8),BITSE(1000,8),XBITSE(1000,8),IXM(8),IS(8),IR(36)
INTEGER L10,L05,L01,DL10,DL05,DL01,S10,S05,S01,DS10,DS05,DS01
INTEGER N,M,IX,IER
INTEGER D,DSTAR,P,Q,PSTAR,QSTAR,ITHETA(3),IPHI(3), N1,N2,
1ICODE,ISPEC,OUTFRM,ISAVE1,ISAVE2,ISAVE3,ZWRITE,BWRITE,
2IPARM(100,4),
3PROGS,DOCORR,DOTSX,DOGENT,DOITSE,
4MAXFUN,ISTART,GIER,
5NORD,IER
DOUBLE PRECISION SEED,SAVE4
DATA OLDP/0./,NEWP/40/,OLDQ/0./,NEWQ/40/
CALL ERRSET(207,256,0,1,0,208)
1ISPEC=0

```

```

      IERROR=0
C*****
C READ CARDS 1 & 2
C*****
      READ(5,100) HED1
      100 FORMAT(20A4)
      READ(5,100) HED2
C*****
C READ CARD 3
C*****
      READ(5,102) NIT,NASS,NREAL,DOCORR,DOTSX,DOGENT,DOITSE,
      +OUTFRM,ZWRITE,BWRITE,ICODE,ISPEC,IDV,MAXFUN,NORD
      102 FORMAT(3I5,1X,4I1,9I5)
      WRITE(6,669) NIT,NASS,NREAL,DOCORR,DOTSX,DOGENT,DOITSE,
      +OUTFRM,ZWRITE,BWRITE,ICODE,ISPEC,IDV,MAXFUN,NORD
      669 FORMAT('1  NIT  NASS  NREAL  PROGS  OUTFRM  ZWRITE  BWRITE  ICODE
      +  ISPEC  IDVMAXFUNNORD'/1X,3(1X,I5,1X),2X,4I1,1X,1X,
      +2(1X,I5,1X),2X,I5,1X,4(I5,1X),I3)
      ISAVE1=OUTFRM
      ISAVE2=ICODE
      ISAVE3=ISPEC
C*****
C READ CARD 4
C*****
      READ(5,11) SEED,XMEAN,SIGMA,DELTA,SLOPE,DSLOPE,N1,N2
      11 FORMAT(D10.5,5F10.5,2I5)
      SAVE4=SEED
      WRITE(6,12) SEED,XMEAN,SIGMA,DELTA,SLOPE,DSLOPE,N1,N2
      12 FORMAT(  71H          SEED      XMEAN      SIGMA      DELTA      SLOPE
      1  DSLOPE  N1  N2/1X,6F10.5,2I5)
C*****
C OUTERMOST LOOP; ITERATE TO END 'NREAL' TIMES
C*****
      DO 10 JJ=1,NREAL
      NRLCNT=JJ
C*****
C READ CARD 5
C*****
      2 READ(5,101) P,D,Q,PHI,THETA
      101 FORMAT(3I5,6F5.2)
C*****
C SECOND LOOP; ITERATE 'NASS' TIMES FOR EACH 'REAL' MODEL
C*****
      DO 10 KK=1,NASS
      NASCNT=KK
      IF(NASCNT.GE.2.AND.1 ISPEC.EQ.1) ISPEC=2
C*****
C CHECK WHETHER ACTUAL OR 'ASSUMED' VALUES PASSED TO TSX
C*****
      IF (ISPEC.NE.0) GO TO 69
C*****
C SET 'ASSUMED' VALUES EQUAL ACTUAL AND SKIP READING OF ASSUMED

```



```

        PSTAR = P
        DSTAR = D
        QSTAR = Q
    DO 70 K=1,3
70  IPHI(K)=100.*PHI(K)
    DO 71 L=1,3
71  ITHETA(L)=100.*THETA(L)
    GO TO 699
C*****
C READ CARD '6'
C READ 'ASSUMED' VALUES TO BE PASSED TO TSX IF ISPEC = 1 OR 2
C*****
    69 READ(5,101) PSTAR,DSTAR,QSTAR,PHSTAR,THSTAR
        IF(ISPEC.EQ.2) GO TO 698
C MULTIPLY BY 100 TO ENTER TSX
    DO 5 J=1,3
        ITHETA(J)=100.*THSTAR(J)
    5  IPHI(J)=100.*PHSTAR(J)
    698 IF(ISPEC.EQ.2) IISPEC=1
        IF(ISPEC.EQ.2) ISPEC=0
C*****
C INITIALIZE A BUNCH OF VARIABLES TO ZERO.
C*****
    699 L10=0
        L05=0
        L01=0
        DL10=0
        DL05=0
        DL01=0
        S10=0
        S05=0
        S01=0
        DS10=0
        DS05=0
        DS01=0
        SVAR=0.
        SSTE=0.
        ISUMKE=0
C*****
C IF REQUESTED,WRITE 5 RECORD HEADER FOR TSX RESULTS UNIT 9 OUTPUT FILE,
C GENTS UNIT 10, AND ITSE UNIT 11.
C NOTE THAT ISAVE3 IS INPUT VALUE OF ISPEC.
C*****
    IF(BWRITE.NE.1) GO TO 501
    IF (DOTSX .NE. 1 .AND. DOGENT .NE. 1 .AND. DOITSE .NE. 1)
    1 WRITE (6,500)
    500 FORMAT('0*****NOTE:YOU REQUESTED THAT THE RESULTS OF TSX,GENT
+ S,OR */' ' *****ITSE BE WRITTEN TO DISK ,BUT DID NOT REQUEST
+*****'/' *****THAT ANY ANALYSES BE PERFORMED.
+ *****')
    PROGS = DOTSX + DOGENT + DOITSE
    IF (PROGS .EQ. 0) GO TO 501

```

```

C-----
C IF REQUESTED, WRITE TSX DISK HEADER-----
  IF (DOTSX .NE. 1) GO TO 511
C WRITE 'INPUT CARDS 1 & 2' -----
  WRITE(9,100) HED1
  WRITE(9,100) HED2
C WRITE 'INPUT CARD 3' -----
  WRITE(9,102)NIT,NASS,NREAL,DOCORR,DOT SX,DOGENT,DOITSE,
  +OUTFRM,ZWRITE,BWRITE,ICODE,ISPEC,IDV,MAXFUN,NORD
C WRITE 'INPUT CARD 4' -----
  WRITE(9,11) SEED,XMEAN,SIGMA,DELTA,SLOPE,DSLOPE,N1,N2
C WRITE 'INPUT CARD 5' -----
  WRITE(9,101) P,D,Q,PHI,THETA
C WRITE 'INPUT CARD 6' -----
  IF(ISAVE3.EQ.0) WRITE(9,502)
  502 FORMAT(' THIS CARD NOT USED IN THIS RUN.')
```

```

  IF(ISAVE3.NE.0) WRITE(9,101) PSTAR,DSTAR,QSTAR,PHSTAR,THSTAR
C-----
C IF REQUESTED, WRITE GENTS DISK HEADER-----
  511 IF (DOGENT .NE. 1) GO TO 512
C WRITE 'INPUT CARD 1' -----
  WRITE(10,100) HED1
  WRITE(10,100) HED2
C WRITE 'INPUT CARD 2' -----
  WRITE(10,102)NIT,NASS,NREAL,DOCORR,DOT SX,DOGENT,DOITSE,
  +OUTFRM,ZWRITE,BWRITE,ICODE,ISPEC,IDV,MAXFUN,NORD
C WRITE 'INPUT CARD 3' -----
  WRITE(10,11) SEED,XMEAN,SIGMA,DELTA,SLOPE,DSLOPE,N1,N2
C WRITE 'INPUT CARD 4' -----
  WRITE(10,101) P,D,Q,PHI,THETA
C WRITE 'INPUT CARD 5' -----
  IF(ISAVE3.EQ.0) WRITE(10,502)
  IF(ISAVE3.NE.0) WRITE(10,101) PSTAR,DSTAR,QSTAR,PHSTAR,THSTAR
C-----
C IF REQUESTED, WRITE ITSE DISK HEADER-----
  512 IF (DOITSE .NE. 1) GO TO 501
C WRITE 'INPUT CARDS 1 & 2' -----
  WRITE(11,100) HED1
  WRITE(11,100) HED2
C WRITE 'INPUT CARD 3' -----
  WRITE(11,102)NIT,NASS,NREAL,DOCORR,DOT SX,DOGENT,DOITSE,
  +OUTFRM,ZWRITE,BWRITE,ICODE,ISPEC,IDV,MAXFUN,NORD
C WRITE 'INPUT CARD 4' -----
  WRITE(11,11) SEED,XMEAN,SIGMA,DELTA,SLOPE,DSLOPE,N1,N2
C WRITE 'INPUT CARD 5' -----
  WRITE(11,101) P,D,Q,PHI,THETA
C WRITE 'INPUT CARD 6' -----
  IF(ISAVE3.EQ.0) WRITE(11,502)
  IF(ISAVE3.NE.0) WRITE(11,101) PSTAR,DSTAR,QSTAR,PHSTAR,THSTAR
C*****
C THIRD LOOP;ITERATE 'NIT' TIMES FOR EACH 'REAL' AND EACH 'ASSUMED'
  501 DO 1 I=1,NIT
```

```

      NITCNT=1
C WRITE VARIOUS AND SUNDRY
      IF (OUTFRM.EQ.0.AND.NITCNT.GE.2) GO TO 700
      WRITE(6,98) HED1,HED2
    98  FORMAT(1H0,20A4/1H ,20A4)
      WRITE(6,123) P,D,Q
    123  FORMAT(/18H TRUE MODEL ARIMA(,12,2H ,,12,2H ,,12,2H ))
      WRITE(6,125) PHI
    125  FORMAT(5X,11H PHI VALUES,2X,3F6.2)
      WRITE(6,124) THETA
    124  FORMAT(5X,13H THETA VALUES,3F6.2)
C*****
C THE NEXT CARD MOVED FROM IMMEDIATELY PRECEEDING CALL TSX
      ISPEC=ISAVE3
C*****
C WRITE ASSUMED MODEL(PASSED TO TSX) IF NOT THE SAME AS THE TRUE MODEL
      IF(ISPEC.EQ.2 .OR. ISPEC.EQ.1) WRITE(6,106) PSTAR,DSTAR,QSTAR
    106  FORMAT(21H ASSUMED MODEL ARIMA(,12,2H ,,12,2H ,,12,2H ))
      IF(ISPEC.EQ.1) WRITE(6,125) PHSTAR
      IF(ISPEC.EQ.1) WRITE(6,124) THSTAR
C*****
      WRITE(6,121) SEED
    121  FORMAT(/8H SEED = ,F16.3)
      WRITE(6,127) SIGMA
    127  FORMAT(42H STANDARD DEVIATION OF RANDOM COMPONENT = ,
    1F8.4)
      WRITE(6,115) XMEAN
    115  FORMAT(26H PREINTERVENTION LEVEL = ,F8.4)
      WRITE(6,116) DELTA
    116  FORMAT(26H POST-I CHANGE IN LEVEL = ,F8.4)
      WRITE(6,117) SLOPE
    117  FORMAT(26H PREINTERVENTION SLOPE = ,F8.4)
      WRITE(6,118) DSLOPE
    118  FORMAT(26H POST-I CHANGE IN SLOPE = ,F8.4)
      WRITE(6,131) NREAL,NRLCNT
    131  FORMAT(/25H NUMBER OF TRUE MODELS = ,15,11H - COUNT = ,15)
      WRITE(6,132) NASS,NASCNT
    132  FORMAT(/28H NUMBER OF ASSUMED MODELS = ,15,11H - COUNT = ,15)
      WRITE(6,134) N1,N2
    134  FORMAT(/6H N1 = ,13,4X,6H N2 = ,13)
    700  WRITE(6,133) NIT,NITCNT
    133  FORMAT(/24H NUMBER OF ITERATIONS = ,15,11H - COUNT = ,15,/)
C GENERATE TIME-SERIES WITH TRUE PARAMETERS
      CALL GENER(NPROB,P,D,Q,N1,N2,SEED,XMEAN,DELTA,PHI,THETA,Z,
    1OUTFRM,SIGMA,Z0,SLOPE,DSLOPE)
C IF WRITE SERIES SELECTED(ZWRITE=1), WRITE NOW TO UNIT 8.
      NTOT=N1+N2
      IF(ZWRITE.EQ.1) WRITE(8,97) (Z(K), K=1,NTOT)
    97  FORMAT(10F8.3)
C*****
C CALL CORREL, TSX, GENTS, AND/OR ITSE AS REQUESTED. *****
C*****

```

```

PROGS = DOCORR + DOTSX + DOGENT + DOITSE
IF (PROGS .EQ. 0) GO TO 1
IF (DOCORR .EQ. 0) GO TO 513

```

```

C-----
C CALL CORREL
C-----

```

```

      CALL CORREL(N1,N2,Z)
513 IF (DOTSX .EQ. 0) GO TO 514

```

```

C-----
C CALL TSX WITH ASSUMED VALUES
C-----

```

```

      OUTFRM=ISAVE1
      ICODE=ISAVE2
      CALL TSX(PSTAR,DSTAR,QSTAR,N1,N2,ICODE,ISPEC,
1OLDP,NEWP,IPHI,OLDQ,NEWQ,ITHETA,Z,OUTFRM,
2NITCNT,HED1,HED2,PHSTAR,THSTAR,
3IDV,VAR,B1,SE1,T1,B2,SE2,T2,B3,SE3,T3,B4,SE4,T4,KE,
4SUMSQ,SS1,SS2,SS3,SS4,
5 PHI1,PHI2,PHI3,THETA1,THETA2,THETA3)

```

```

C-----
C KEEP TRACK OF HOW MANY TIMES TSX FOUND MORE THAN ONE EMIN.
C-----

```

```

      IF(KE.NE.1) ISUMKE=ISUMKE+1

```

```

C-----
C ACCUMULATE SUCCESSES AT .10, .05, AND .01 LEVELS
C-----

```

```

      AT1=ABS(T1)
      AT2=ABS(T2)
      AT3=ABS(T3)
      AT4=ABS(T4)
      IF(AT1.GE.1.658) L10= L10+1
      IF(AT1.GE.1.980) L05= L05+1
      IF(AT1.GE.2.617) L01= L01+1
      IF(AT2.GE.1.658) DL10=DL10+1
      IF(AT2.GE.1.980) DL05=DL05+1
      IF(AT2.GE.2.617) DL01=DL01+1
      IF(AT3.GE.1.658) S10= S10+1
      IF(AT3.GE.1.980) S05= S05+1
      IF(AT3.GE.2.617) S01= S01+1
      IF(AT4.GE.1.658) DS10=DS10+1
      IF(AT4.GE.1.980) DS05=DS05+1
      IF(AT4.GE.2.617) DS01=DS01+1

```

```

C-----

```

```

      PARMS(1,1) = VAR
      PARMS(1,2) = B1
      PARMS(1,3) = T1
      PARMS(1,4) = SE1
      PARMS(1,5) = B2
      PARMS(1,6) = T2
      PARMS(1,7) = SE2
      PARMS(1,8) = B3
      PARMS(1,9) = T3

```

```

PARMS(1,10) = SE3
PARMS(1,11) = B4
PARMS(1,12) = T4
PARMS(1,13) = SE4
PARMS(1,14) = Z0
PARMS(1,15) = SUMSQ
PARMS(1,16) = SS1
PARMS(1,17) = SS2
PARMS(1,18) = SS3
PARMS(1,19) = SS4
PARMS(1,20) = PHI1
PARMS(1,21) = PHI2
PARMS(1,22) = PHI3
PARMS(1,23) = THETA1
PARMS(1,24) = THETA2
PARMS(1,25) = THETA3
IPARM(1,1) = KE
IPARM(1,2) = NRLCNT
IPARM(1,3) = NASCNT
IPARM(1,4) = NITCNT

```

```

C-----
C CALL GENTS
C-----

```

```

514 IF (DOGENT .EQ. 0) GO TO 516
    ICODE = ISAVE2
    OUTFRM = ISAVE1
    ISTART = 0
    CALL GENTS(NTOT,N1,ICODE,Z,MAXFUN,ISTART,GBETA,OUTFRM,BWRITE,
1                                NRLCNT,NASCNT,NITCNT)
    DO 515 K=1,4,1
        BGENTS(1,K) = GBETA(K)
515 CONTINUE

```

```

C-----
C CALL ITSE: IBETA(1)="LEVEL",IBETA(2)="POST-L",IBETA(3)="DLEVEL",
C            IBETA(4)=SLOPE,IBETA(5)=POST-S,IBETA(6)=DSLOPE,
C            IBETA(7)=EMIN,IBETA(8)=SUM OF SQUARES
C-----

```

```

516 IF (DOITSE .EQ. 0) GO TO 1
    IF (IERROR .EQ. 1) GO TO 1
    CALL ITSE(N1,N2,NORD,BWRITE,OUTFRM,Z,IERROR,IBETA,
1                                NRLCNT,NASCNT,NITCNT)
    IF (IERROR .EQ. 1) GO TO 1
    DO 517 K = 1,8,1
        BITSE(1,K) = IBETA(K)
517 CONTINUE

```

```

C-----
1 CONTINUE

```

```

C*****

```

```

    PROGS = DOTSX + DOGENT + DOITSE
    IF (PROGS .EQ. 0) GO TO 10

```

```

C*****

```

```

C INITIALIZE FOR AND CALL BECORI(COMPUTES MEAN & SIGMA OF TSX/GENTS/ITSE

```

```

C PARAMETER ESTIMATES OVER NIT ITERATIONS)
C*****
C COMPUTE MEAN & SIGMA OF TSX PARAMETER ESTIMATES.
    IF (DOTSX .EQ. 0) GO TO 520
    DO 510 I=1,NIT,1
    DO 509 J=1,19,1
    XPARMS(I,J)=PARMS(I,J)
509 CONTINUE
510 CONTINUE
    N=NIT
    M=19
    IX=1000
    CALL BECORI(XPARMS,N,M,IX,XM,S,R,IER)
C*****
C COMPUTE TYPE I ERROR RATES FOR TSX SOLUTIONS.
C*****
    AL10= L10
    AL05= L05
    AL01= L01
    ADL10=DL10
    ADL05=DL05
    ADL01=DL01
    AS10= S10
    AS05= S05
    AS01= S01
    ADS10=DS10
    ADS05=DS05
    ADS01=DS01
    ANIT=NIT
    PL10=AL10/ANIT
    PL05=AL05/ANIT
    PL01=AL01/ANIT
    PDL10=ADL10/ANIT
    PDL05=ADL05/ANIT
    PDL01=ADL01/ANIT
    PS10=AS10/ANIT
    PS05=AS05/ANIT
    PS01=AS01/ANIT
    PDS10=ADS10/ANIT
    PDS05=ADS05/ANIT
    PDS01=ADS01/ANIT
C*****
C COMPUTE MEAN & SIGMA FOR GENTS' BETAS.
C*****
520 IF (DOGENT .EQ. 0) GO TO 525
    DO 521 I=1,NIT,1
    DO 522 J=1,4,1
    XBAGENT(I,J) = BGENTS(I,J)
522 CONTINUE
521 CONTINUE
    N = NIT
    M = 4

```

```

IX = 1000
CALL BECORI(XBGENT,N,M,IX,GXM,GS,GR,GIER)
C*****
C COMPUTE MEAN & SIGMA FOR ITSE PARAMETERS.
C*****
525 IF (DOITSE .EQ. 0) GO TO 530
    IF (IERROR .EQ. 1) GO TO 530
    DO 526 I=1,NIT,1
    DO 526 J=1,8,1
    XBITSE(I,J) = BITSE(I,J)
527 CONTINUE
526 CONTINUE
    N = NIT
    M = 8
    IX = 1000
    CALL BECORI(XBITSE,N,M,IX,IXM,IS,IR,IIER)
C*****
C PRINT OUT RESULTS AND WRITE 'SUMMARY' TO UNIT 12
C*****
530 WRITE(6,141) HED1,HED2
141 FORMAT('1',20A4,/, ' ',20A4)
    WRITE(12,100) HED1
    WRITE(12,100) HED2
    WRITE (6,140) NITCNT
    WRITE (12,140) NITCNT
140 FORMAT (' THE FOLLOWING ARE THE RESULTS AFTER ',15,
1' ITERATIONS THRU TSX/GENTS/ITSE:')
    WRITE(6,103) P,D,Q
    WRITE(12,103) P,D,Q
103 FORMAT(18H TRUE MODEL ARIMA(,12,2H ,,12,2H ,,12,2H ))
    WRITE(6,105) PHI
    WRITE(12,105) PHI
105 FORMAT(5X,11H PHI VALUES,2X,3F6.2)
    WRITE(6,104) THETA
    WRITE(12,104) THETA
104 FORMAT(5X,13H THETA VALUES,3F6.2)
C*****
C WRITE ASSUMED MODEL(PASSED TO TSX) IF NOT THE SAME AS THE TRUE MODEL
    IF(1SPEC.EQ.2 .OR. 1SPEC.EQ.1) WRITE(6,106) PSTAR,DSTAR,QSTAR
    IF(1SPEC.EQ.2 .OR. 1SPEC.EQ.1) WRITE(12,106) PSTAR,DSTAR,QSTAR
    IF(1SPEC.EQ.1) WRITE(6,105) PHSTAR
    IF(1SPEC.EQ.1) WRITE(12,105) PHSTAR
    IF(1SPEC.EQ.1) WRITE(6,104) THSTAR
    IF(1SPEC.EQ.1) WRITE(12,104) THSTAR
C*****
    WRITE(6,127) SIGMA
    WRITE(12,127) SIGMA
    WRITE(6,115) XMEAN
    WRITE(12,115) XMEAN
    WRITE(6,116) DELTA
    WRITE(12,116) DELTA
    WRITE(6,117) SLOPE

```

```

WRITE(12,117) SLOPE
WRITE(6,118) DSLOPE
WRITE(12,118) DSLOPE

```

```

C-----
C  PRINT TSX SUMMARY RESULTS
C-----
      IF (DOTSX .EQ. 0) GO TO 531
      WRITE(6,107)
      WRITE(12,107)
107  FORMAT(/33H NOMINAL ALPHA: .10 .05 .01)
      WRITE(6,108) PL10,PL05,PL01,PDL10,PDL05,PDL01,
+                PS10,PS05,PS01,PDS10,PDS05,PDS01
      WRITE(12,108) PL10,PL05,PL01,PDL10,PDL05,PDL01,
+                PS10,PS05,PS01,PDS10,PDS05,PDS01
108  FORMAT(' LEVEL: ',7X,3(F5.3,2X)/
+        ' LEVELCHG: ',4X,3(F5.3,2X)/
+        ' SLOPE: ',7X,3(F5.3,2X)/
+        ' SLOPECHG: ',4X,3(F5.3,2X))
      WRITE(6,110) XM(1),XM(14),S(1),S(14)
      WRITE(12,110) XM(1),XM(14),S(1),S(14)
      WRITE(6,114) XM(15),XM(16),XM(17),XM(18),XM(19),
+ S(15),S(16),S(17),S(18),S(19)
      WRITE(12,114) XM(15),XM(16),XM(17),XM(18),XM(19),
+ S(15),S(16),S(17),S(18),S(19)
114  FORMAT( 28X,14H SUM OF SQUARES,/ 9X,12H RESIDUAL ,
+ 12H LEVEL ,12H LEVELCHG ,12H SLOPE ,12H SLOPECHG ,
+ /1X,6H MEAN: ,5F12.6,8H :MEAN,
+ / 7H SIGMA: ,5F12.6,9H :SIGMA)
110  FORMAT(' ----- TSX SUMMARY RESULTS -----',/,
+ 7X,28HEMIN(MINIMUM ERROR VARIANCE),
+ 7X,22HVALUE OF SERIES AT T=0/1X,6H MEAN: ,F10.3,
+ 36X,F10.3,8H :MEAN,/ 7H SIGMA: ,F10.3,36X,F10.3,9H :SIGMA)
      IF(1SUMKE.NE.0) WRITE(6,503) 1SUMKE
      IF(1SUMKE.NE.0) WRITE(12,503) 1SUMKE
503  FORMAT(/,' WARNING: ON ',15,' OCCASIONS TSX FOUND MORE THAN 1 EMIN
+ SOLUTION. ')
      WRITE(6,111)
      WRITE(12,111)
111  FORMAT(' ',27X,'PARAMETER ESTIMATES',/,
+ 9X,65H LEVEL STANEROR T-STAT LEVELCHG STANEROR T
+ -STAT)
      WRITE(6,112) XM(2),XM(4),XM(3),XM(5),XM(7),XM(6),
+ S(2),S(4),S(3),S(5),S(7),S(6)
      WRITE(12,112) XM(2),XM(4),XM(3),XM(5),XM(7),XM(6),
+ S(2),S(4),S(3),S(5),S(7),S(6)
112  FORMAT(1X,6H MEAN: ,F10.5,1X,F10.6,1X,F9.3,6X,F10.5,1X,F10.6,
+ F9.3,5H:MEAN/
+ 7H SIGMA: ,F10.5,1X,F10.6,1X,F9.3,6X,F10.5,1X,F10.6,F9.3,
+ 6H:SIGMA,/)
      WRITE(6,113)
      WRITE(12,113)
113  FORMAT( 9X,65H SLOPE STANEROR T-STAT SLOPECHG ST

```



```

+ANEROR  T-STAT)
WRITE(6,112) XM(8),XM(10),XM(9),XM(11),XM(13),XM(12),
+ S(8),S(10),S(9),S(11),S(13),S(12)
WRITE(12,112) XM(8),XM(10),XM(9),XM(11),XM(13),XM(12),
+ S(8),S(10),S(9),S(11),S(13),S(12)
C-----
C PRINT GENTS SUMMARY RESULTS. -----
C-----
531 IF (DOGENT .EQ. 0) GO TO 540
WRITE(6,533) (I,GXM(I),GS(I),I=1,4)
WRITE(12,533) (I,GXM(I),GS(I),I=1,4)
533 FORMAT('-----GENTS SUMMARY RESULTS-----',
1 /,' BETA # MEAN SIGMA',
2 /,' ',6X,13,5X,F10.5,3X,F10.5))
C-----
C PRINT ITSE SUMMARY RESULTS-----
C-----
540 IF (DOITSE .EQ. 0) GO TO 545
IF (IERROR .EQ. 1) WRITE(6,544)
IF (IERROR .EQ. 1) WRITE(12,544)
544 FORMAT(' SEE ERROR MESSAGE, ITSE SUMMARY NOT PRINTED')
IF (IERROR .EQ. 1) WRITE(11,544)
IF (IERROR .EQ. 1) GO TO 545
WRITE (6,542) (IXM(I),IS(I),I=1,8)
WRITE (12,542) (IXM(I),IS(I),I=1,8)
542 FORMAT('----- ITSE SUMMARY RESULTS -----',
1 /,' VARIABLE MEAN SIGMA ',
1 5X,' VARIABLE MEAN SIGMA',
2 /,' ',3X,'LEVEL',4X,F10.5,5X,F10.5,
2 8X,'POSTL',4X,F10.5,5X,F10.5,
3 /,' ',3X,'DLEVEL',3X,F10.5,5X,F10.5,
4 /,' ',3X,'SLOPE',4X,F10.5,5X,F10.5,
4 8X,'POSTS',4X,F10.5,5X,F10.5,
5 /,' ',3X,'DSLOPE',3X,F10.5,5X,F10.5,
6 /,' ',3X,'EMIN',5X,F10.5,5X,F10.5,
6 8X,'SUMSQ',4X,F10.5,5X,F10.5)
C-----
C*****
C WRITE ORIGINAL INPUT VALUES.NOTE ISAVE1,ISAVE2,ISAVE3,& SAVE4 ARE
C OUTFRM,ICODE,ISPEC & SEED ORIGINAL VALUES RESPECTIVELY.
C*****
545 WRITE(6,670) NIT,NASS,NREAL,DOCORR,DOTSX,DOGENT,DOITSE,
+ISAVE1,ZWRITE,BWRITE,ISAVE2,ISAVE3,IDV,MAXFUN,NORD
WRITE(12,670) NIT,NASS,NREAL,DOCORR,DOTSX,DOGENT,DOITSE,
+ISAVE1,ZWRITE,BWRITE,ISAVE2,ISAVE3,IDV,MAXFUN,NORD
670 FORMAT(' NIT NASS NREAL PROGS OUTFRM ZWRITE BWRITE ICODE
+ ISPEC IDVMAXFUNNORD'/1X,3(1X,15,1X),2X,411,1X,1X,
+2(1X,15,1X),2X,15,1X,4(15,1X),13)
WRITE(6,671) SAVE4,XMEAN,SIGMA,DELTA,SLOPE,DSLOPE,N1,N2
WRITE(12,671) SAVE4,XMEAN,SIGMA,DELTA,SLOPE,DSLOPE,N1,N2
671 FORMAT( 71H SEED XMEAN SIGMA DELTA SLOPE
1 DSLOPE N1 N2/1X,6F10.5,215)

```

```

C*****
C IF REQUESTED, WRITE SUMMARY RESULTS OF ANALYSES TO DISK OUTFILES.
C*****
      PROGS = DOT SX + DOGENT + DOITSE
      IF (PROGS .EQ. 0) GO TO 10
      IF (BWRITE .NE. 1) GO TO 10
      IF (DOT SX .EQ. 0) GO TO 550

C-----
C  WRITE TSX SUMMARIES TO UNIT 9  -----
C-----
C ***** WRITE RECORD 6. *****
      WRITE(9,507) ISUMKE
507  FORMAT(15)
C ***** WRITE RECORDS 7 THRU 13. *****
      WRITE(9,504)
      +      PL10,PL05,PL01,PDL10,PDL05,PDL01,
      +      PS10,PS05,PS01,PDS10,PDS05,PDS01,
      +      XM(1),XM(2),XM(4),XM(3),XM(5),XM(7),XM(6),
      +      S(1),S(2),S(4),S(3),S(5),S(7),S(6),
      +      XM(14),XM(8),XM(10),XM(9),XM(11),XM(13),XM(12),
      +      S(14),S(8),S(10),S(9),S(11),S(13),S(12),
      +      XM(15),XM(16),XM(17),XM(18),XM(19),
      +      S(15),S(16),S(17),S(18),S(19)
504  FORMAT(12(F5.3,1X)/
      +      F10.5,2(F10.5,F12.8,F8.2)/
      +      F10.5,2(F10.5,F12.8,F8.2)/
      +      F10.5,2(F10.5,F12.8,F8.2)/
      +      F10.5,2(F10.5,F12.8,F8.2)/
      +      5F12.6/5F12.6)
C ***** WRITE RECORDS '14' THRU '17'. *****
      DO 506 I=1,NIT,1
      WRITE(9,505) IPARM(1,1),IPARM(1,2),IPARM(1,3),IPARM(1,4),
      +PARMS(1,20),PARMS(1,21),PARMS(1,22),
      +PARMS(1,23),PARMS(1,24),PARMS(1,25),
      +PARMS(1,1),PARMS(1,2),PARMS(1,4),PARMS(1,3),
      +PARMS(1,5),PARMS(1,7),PARMS(1,6),
      +PARMS(1,14),PARMS(1,8),PARMS(1,10),PARMS(1,9),
      +PARMS(1,11),PARMS(1,13),PARMS(1,12),
      +PARMS(1,15),PARMS(1,16),PARMS(1,17),PARMS(1,18),PARMS(1,19)
505  FORMAT(11,1X,12,1X,12,1X,14,6F10.5/
      +      F10.5,2(F10.5,F12.8,F8.2)/
      +      F10.5,2(F10.5,F12.8,F8.2)/
      +      5F12.6)
506  CONTINUE

C-----
C  WRITE GENTS SUMMARIES TO UNIT 10  -----
C-----
550 IF (DOGENT .EQ. 0) GO TO 555
      WRITE(10,551) (1,GXM(1),GS(1),I=1,4)
551  FORMAT('-----GENTS SUMMARY RESULTS-----',
      1  /'      BETA #      MEAN      SIGMA',
      2  /(7X,13,5X,F10.5,3X,F10.5))

```

```

C-----
C  WRITE ITSE SUMMARIES TO UNIT 11  -----
C-----
555 IF (DOITSE .EQ. 0) GO TO 10
    WRITE (11,556) (IXM(I),IS(I),I=1,6)
556 FORMAT('----- ITSE SUMMARY RESULTS -----',
1 /,' ' VARIABLE           MEAN           SIGMA',
2 /,' ',5X,'LEVEL',4X,F10.5,5X,F10.5,
3 /,' ',5X,'DLEVEL',3X,F10.5,5X,F10.5,
4 /,' ',5X,'SLOPE',4X,F10.5,5X,F10.5,
5 /,' ',5X,'DSLOPE',3X,F10.5,5X,F10.5,
6 /,' ',5X,'EMIN',5X,F10.5,5X,F10.5,
7 /,' ',5X,'SUMSQ',4X,F10.5,5X,F10.5)
C-----
C*****
C THIS IS THE END OF THE 'SECOND LOOP' I.E. DO NASS (NUMBER OF ASSUMED
C   MODELS) TIMES FOR EACH NREAL (NUMBER OF REAL MODELS).
C*****
10  CONTINUE
C*****
    STOP
    END
C*****
    SUBROUTINE GENER(NPROB,IP,DIFF,IQ,N1,N2,SEED,XMEAN,DELTA,
1PHI,THETA,Z,OUTFRM,SIGMA,Z0,SLOPE,DSLOPE)
    DOUBLE PRECISION SEED
    REAL Z(300),PHI(3),THETA(3),RAND(310),C(5,7),PHID(7),ZP10(310),
1SIGMA,Z0,SLOPE,DSLOPE,SINC,DSINC
    INTEGER DIFF,OUTFRM
    DATAC/0.,-1.,-2.,-3.,-4.,2*0.,1.,3.,6.,3*0.,-1.,-4.,4*0.,1.,15*0./
    SINC=0.
    DSINC=0.
    DO 1 I=1,7
    IF(I.LE.IP) PHID(I)=PHI(I)
    IF(I.GT.IP) PHID(I)=0.
1  CONTINUE
4  NTOT=N1+N2
    IPD=IP+DIFF
    NZ=NTOT+10
    N101=N1+11
    CALL GGNML(SEED,NZ,RAND)
    DO 6 JJ=1,NZ
C MULTIPLY RANDOM NUMBER N(0,1) BY SIGMA FOR N(0,SIGMA**2)
    RAND(JJ)=RAND(JJ)*SIGMA
    ERR=RAND(JJ)
C ADD THE 'LEVEL' TO EACH OBSERVATION
    ZT=ERR+XMEAN
C COMPUTE INCREMENT DUE TO SLOPE (STARTS AT 11TH POINT)
    IF (JJ .GT. 10) SINC=(JJ-10)*SLOPE
C COMPUTE INCREMENT DUE TO CHANGE IN SLOPE (STARTS AT N1+11TH POINT)
    IF (JJ .GE. N101) DSINC=(JJ-N1-10)*DSLOPE
C COMPUTE MOVING AVERAGES IF ANY

```

```

        IF(IQ.EQ.0) GO TO 13
        DO 7 I=1,IQ
7        IF((JJ-1).GE.1) ZT=ZT-THETA(I)*RAND(JJ-1)
C COMPUTE AUTOREGRESSIVES & DIFFERENCES IF ANY
        IF(IPD.EQ.0) GO TO 12
13       DO 8 I=1,IPD
        IF((JJ-1).LT.1) GO TO 8
        SUM=    PHID(I)-C(DIFF+1,I)
        IM1=I-1
        IF(IM1.LT.1) GO TO 14
        DO 9 J=1,IM1
9        SUM=SUM+PHID(J)*C(DIFF+1,I-J)
14       CONS=0.
        IF((JJ-1).GE.N101) CONS=DELTA+DSINC
        ZT=ZT+SUM*(ZP10(JJ-1)-XMEAN-CONS-SINC)
8        CONTINUE
12       CONTINUE
        I=JJ
C ADD THE INCREMENT DUE TO SLOPE *****
        ZT=ZT+SINC
C ADD THE INCREMENTS DUE TO DELTA & POST-I SLOPE IF T=POST-I.*****
        IF(JJ.GE.N101) ZT=ZT+DELTA+DSINC
        ZP10(JJ)=ZT
6        CONTINUE
        Z0=ZP10(10)
        DO 10 K=11,NZ
        I=K-10
10       Z(I)=ZP10(K)
        WRITE(6,126) Z0
126      FORMAT(26H VALUE OF SERIES AT T=0 IS,F6.2)
        IF(OUTFRM.EQ.0) GO TO 800
        IF(OUTFRM.GE.1) WRITE(6,98)
98       FORMAT(/10X,41HTHESE ARE THE RANDOM NUMBERS: NID(0,1**2))
        IF(OUTFRM.GE.1) WRITE(6,100) (RAND(I),I=1,NZ)
        IF(OUTFRM.GE.1) WRITE(6,99)
99       FORMAT(/10X,33HTHIS IS THE GENERATED TIME-SERIES)
        IF(OUTFRM.GE.1) WRITE(6,100) (Z(I),I=1,NTOT)
        IF(OUTFRM.GE.2) WRITE (6,109)
109      FORMAT(1H1)
        IF(OUTFRM.GE.2) CALL PLOTT (Z,NTOT,1.,1.)
        IF(OUTFRM.EQ.3) WRITE(7,100) (Z(I), I=1,NTOT)
100      FORMAT(10F8.3)
800      RETURN
        END

```