NONPARAMETRIC EXTRAPOLATIVE FORECASTING: AN EVALUATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Roland Doyle Kankey, B.S., M.S., M.A.

* * * * *

The Ohio State University

1988

Dissertation Committee:

Dr. G. W. Milligan, Chairperson

Dr. J. C. Chambers

Dr. P. A. Thompson

Dr. D. A. Wolfe

Approved by

Adviser
Graduate Program in
Business Administration

This is dedicated to all those who have supported my efforts during the past five years. Especially my wife Linda and our three sons Jason, Andrew, and Adam who have managed admirably while a major portion of my attention has been concentrated on studies and research.

ACKNOWLEDGEMENTS

Many thanks to all the excellent faculty at The Ohio State University for their continual support and encouragement. In particular the members of my committee and others who encouraged hard work and research activity. In addition, my thanks to all my compatriots at the Air Force Institute of Technology, School of Systems and Logistics.

VITA

Nov 17, 1946 . . . . . Born, Batesville, Arkansas

1968 . . . . BS (Cum Laude), Wichita State Univ., Wichita, Kansas (Mathematics)

1970 . . . . MS, Oklahoma State University, Stillwater, Okla. (Mathematics)

1970 - 77 . . . . . . Management / Cost Analyst; Rome Air Development Center (Rome, New York), 51 Air Base Wing (Osan Air Base, Republic of Korea), Hq United States Air Force (Washington, D.C.)

1977 - 83 . . . . . . Instructor and Assistant Professor, School of Systems and Logistics, Air Force Institute of Technology, Wright-Patterson AFB, Ohio

1985 . . . . MA, The Ohio State University, Columbus, Ohio (Management Science)

1985 - Present . . . . Assistant Professor, School of Systems and Logistics, Air Force Institute of Technology, Wright-Patterson AFB, Ohio

PUBLICATIONS

"Life Cycle Cost Management: Analytical or Political?", Proceedings, 16th Annual Symposium, Society of Logistics Engineers, Summer 1981.

"Life Cycle Cost Management: What is it?", AFIT AOG Quarterly, Vol 3, No. 2, 1982.

"Precepts for Life Cycle Cost Management", Air Force Journal of Logistics, Vol 6, No. 1, 1982.

"Learning Curves: An Overview", National Estimator, Vol 4, No. 2, 1983.

"Time Series Analysis", Armed Forces Comptroller, Vol 28, No. 3, 1983.

"Bin Packing Problems in One Dimension: Heuristic Solutions and Confidence Intervals", with N.G. Hall, S. Ghosh, S. Narasimhan, and W.T. Rhee, College of Administrative Science, The Ohio State University, WPS 85-32, March 1985. (Accepted by Computers and Operations Research)

"Logistics Cost and Logistics Cost Analysis", National Estimator, Vol 6, No. 1, 1985; Reprinted Vol 7, No 3, 1987.

"Logistics Support Analysis in Life Cycle Cost Management", AFIT Technical Report, AU-AFIT-LSQ-1, August 18, 1986.

"Loss Functions and Forecast Accuracy Statistics: Some Relationships Between Measures Used to Compare Forecasting Techniques", with P. A. Thompson, College of Administrative Science, The Ohio State University, WPS 86-91, Sept 86.

"Dynamic Costing/Modular Costing in Logistics Systems", National Estimator, Vol 7, No. 2, 1986.

"Competition in a Noncompetitive Environment", with T. Kloppenborg, Air Force Business Research Management Center Report, BRMC-87-02, May 1987.

FIELDS OF STUDY

Major Field: Management Science
Minor Field: Logistics

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

VITA iv

LIST OF TABLES ix

LIST OF FIGURES xiii

CHAPTER

I. INTRODUCTION 1

1.1 Environment 5
1.2 Extrapolative Forecasting 7
1.3 Empirical Results 10
1.4 Nonparametric Philosophy 11

II. LITERATURE REVIEW 13

2.1 Comparisons Between Forecasting Approaches 13
2.2 Comparisons Between Extrapolative Techniques 16
2.3 Nonparametric Forecasting Techniques . 20

III. RESEARCH OBJECTIVES 24

3.1 Stationary Series 25
3.2 Series With Linear Trend 27
3.3 Limitations 27

IV. SELECTION OF ERROR MEASURES 29

4.1 How are Iterative Forecasts Evaluated? 32
4.1.1 Absolute and Squared Measures . 33
4.1.2 Raw versus Relative Measures . . 34
4.1.3 Realization Measures 36
4.2 Need for a Variety of Measures 38
4.3 Simulation Study of Error Measure Relationships. 40

4.4 Evaluation of Results 44
4.4.1 MAD vs MSE. 44
4.4.2 MAD vs MAPE 44
4.4.3 MSE vs MAPE 45
4.4.4 Theil's U vs any other measure . 46
4.5 Summary and Conclusions 61

V. PARAMETRIC AND NONPARAMETRIC TECHNIQUES STUDIED 62

5.1 Techniques for Stationary Time Series . 62
5.1.1 Parametric Techniques 63
5.1.2 Nonparametric Alternatives . . . 67
5.2 Techniques for Series With Linear Trend 70
5.2.1 Parametric Techniques 71
5.2.2 Nonparametric Alternatives . . . 73
5.3 Summary and Conclusions 75

VI. EVALUATION OF PERFORMANCE 76

6.1 Simulation Study 76
6.1.1 Performance on Stationary Series 81
6.1.1.1 Normal Errors 82
6.1.1.2 Cauchy Errors 90
6.1.1.3 Conclusions from Stationary Simulations 95
6.1.2 Performance on Linear Trend Series 98
6.1.2.1 Normal Errors 98
6.1.2.2 Cauchy Errors 99
6.1.2.3 Conclusions of Linear Trend Simulations 102
6.1.3 Simulation Conclusions 104
6.2 Empirical Study 105
6.2.1 Stationary Techniques 107
6.2.1.1 Results Using MAD 109
6.2.1.2 Results Using MSE 110
6.2.1.3 Results Using MAPE . . . 111
6.2.1.4 Results Using Theil's U . 112
6.2.1.5 Stationary Techniques Conclusion 113
6.2.2 Linear Trend Techniques 114
6.2.2.1 Results Using MAD 117
6.2.2.2 Results Using MSE 117
6.2.2.3 Results Using MAPE . . . 118
6.2.2.4 Results Using Theil's U . 120
6.2.2.5 Linear Trend Techniques Conclusion 121
6.2.3 Empirical Study Summary 122
6.3 Summary of Performance 123

VII. SUMMARY AND CONCLUSIONS 124

7.1 Summary 124
7.2 Conclusions and Recommendations for Further Study 125

APPENDICES

A: SAMPLE ERROR MEASURE EVALUATION PROGRAMS . . 132
B: SAMPLE SIMULATION PROGRAMS 139
C: DETAILED SIMULATION RESULTS 174
D: SAMPLE EMPIRICAL PROGRAMS 199

BIBLIOGRAPHY 245

LIST OF TABLES

Table		Page
1	Summary of the Newbold and Granger Study: Comparison of Box-Jenkins' and Holt-Winter's Forecasting Performance on 1Ø6 Series	17
2	Summary of the Makridakis and Hibon Study: Comparison of Ten Extrapolative Method's Forecasting Performance on 111 Series	19
3	Summary of the M-Competition: Comparison of Ten Extrapolative Method's Forecasting Performance on 1ØØ1 Series	21
4	Summary of the M-Competition: Comparison of Selected Extrapolative Method's Forecasting Performance on 111 Series	22
5	Some Accuracy Statistics Used to Compare Forecasting Techniques.	31
6	Error Measures Reported in Fourteen Recent Forecasting Articles	4Ø
7	Spearman's Correlations Between MAD and the Other Error Measures as Series Dispersion and Probability of Outliers is Varied	47
8	Spearman's Correlations Between MSE, MAPE, and Theil's U as Series Dispersion and Probability of Outliers is Varied	48
9	Best Nonparametric Results versus Best Parametric Results, Stationary Series, Normal Errors, Average Results over 1ØØ Simulated Series	88
1Ø	Best Nonparametric Results versus Best Parametric Results, Stationary Series, Cauchy Errors, Average Results over 1ØØ Simulated Series	92

11 Best Nonparametric Results versus Best Parametric Results, Stationary Series, Cauchy Errors, Median Results over 100 Simulated Series 94

12 Best Nonparametric Results versus Best Parametric Results, Trend Series, Normal Errors, Average Results over 100 Simulated Series 100

13 Best Nonparametric Results versus Best Parametric Results, Trend Series, Cauchy Errors, Average Results over 100 Simulated Series 101

14 Best Nonparametric Results versus Best Parametric Results, Trend Series, Cauchy Errors, Median Results over 100 Simulated Series 102

15 Empirical Results using Mean Absolute Deviation: Stationary Techniques' Average MAD Values over 19 Series 109

16 Empirical Results using Mean Squared Error: Stationary Techniques' Average MSE Values (Millions) over 19 Series 110

17 Empirical Results using Mean Absolute Percentage Error: Stationary Techniques' Average MAPE Values over 19 Series 111

18 Empirical Results using Theil's U: Stationary Techniques' Theil's U Values over 19 Series . 112

19 Empirical Results using Mean Absolute Deviation: Linear Trend Techniques' Average MAD Values over 92 Series 117

20 Empirical Results using Mean Squared Error: Linear Trend Techniques' Average MSE Values (Millions) over 92 Series 118

21 Empirical Results using Mean Absolute Percentage Error: Linear Trend Techniques' Average MAPE Values over 92 Series 119

22 Empirical Results using Theil's U: Linear Trend Techniques' Theil's U Values over 91 Series . 120

23 Mean MAD Ratios to Naive One MAD for Stationary Series, Normal Errors 175

24 Mean MSE Ratios to Naive One MSE for Stationary Series, Normal Errors 176

25 Mean MAPE Ratios to Naive One MAPE for Stationary Series, Normal Errors 177

26 Mean Theil's U Values for Stationary Series, Normal Errors 178

27 Mean MAD Ratios to Naive One MAD for Stationary Series, Cauchy Errors 179

28 Mean MSE Ratios to Naive One MSE for Stationary Series, Cauchy Errors 180

29 Mean MAPE Ratios to Naive One MAPE for Stationary Series, Cauchy Errors 181

30 Mean Theil's U Values for Stationary Series, Cauchy Errors 182

31 Median MAD Ratios to Median Naive One MAD for Stationary Series, Cauchy Errors 183

32 Median MSE Ratios to Median Naive One MSE for Stationary Series, Cauchy Errors 184

33 Median MAPE Ratios to Median Naive One MAPE for Stationary Series, Cauchy Errors 185

34 Median Theil's U Ratios for Stationary Series, Cauchy Errors 186

35 Mean MAD Ratios to Naive One MAD for Trend Series, Normal Errors 187

36 Mean MSE Ratios to Naive One MSE for Trend Series, Normal Errors 188

37 Mean MAPE Ratios to Naive One MAPE for Trend Series, Normal Errors 189

38 Mean Theil's U Values for Trend Series, Normal Errors 190

39 Mean MAD Ratios to Naive One MAD for Trend Series, Cauchy Errors 191

40 Mean MSE Ratios to Naive One MSE for Trend Series, Cauchy Errors 192

41 Mean MAPE Ratios to Naive One MAPE for Trend Series, Cauchy Errors 193

42 Mean Theil's U Values for Trend Series, Cauchy Errors 194

43 Median MAD Ratios to Median Naive One MAD for Trend Series, Cauchy Errors 195

44 Median MSE Ratios to Median Naive One MSE for Trend Series, Cauchy Errors 196

45 Median MAPE Ratios to Median Naive One MAPE for Trend Series, Cauchy Errors 197

46 Median Theil's U Values for Trend Series, Cauchy Errors 198

LIST OF FIGURES

Figure		Page
1	Efficiency Frontier for Time Series Forecasting Methods	22
2	Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAD and MSE	49
3	Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAD and MAPE	50
4	Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAD and Theil's U	51
5	Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MSE and MAPE	52
6	Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MSE and Theil's U	53
7	Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAPE and Theil's U	54
8	Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MSE and MAD	55

9 Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MAPE and MAD 56

10 Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MAD and Theil's U . . . 57

11 Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MSE and MAPE 58

12 Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MSE and Theil's U . . . 59

13 Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MAPE and Theil's U . . . 60

14 Development of a Walsh Average 69

15 Pictoral of One Replication 78

16 Graph of a Stationary Series with Phi = .3 . 79

17 Graph of a Stationary Series with Phi = .7 . 80

18 Sample Box Plots of Simulation MAD Results using Moving Averages on a Series with Normal Errors, and Phi = .3 83

19 Sample Box Plots of Simulation MAD Results using Running Medians on a Series with Normal Errors, and Phi = .3 84

20 Sample Box Plots of Simulation MAD Results using Moving Averages on a Series with Cauchy Errors, and Phi = .3 85

21 Sample Box Plots of Simulation MAD Results using Running Medians on a Series with Cauchy Errors, and Phi = .3 86

CHAPTER I

INTRODUCTION

The area of forecasting is often misunderstood. Overtones of mysticism (ex. Delphi method) become mixed with references to covariance matrices and multicollinearity. In general, the forecaster's job is to envision the different possible futures and decide which are more likely, given current history and certain assumptions.

Forecasts can be categorized in several ways, with the most common being to consider (1) the lead time or horizon of the forecast, and (2) the forecasting approach used. Since forecasts are made for a future time period, the question is how far into the future. The general division is into short range, intermediate range, and long range. Although Armstrong (1985) indicates these are difficult to tie down and vary by the situation, they are useful ways to divide forecasting tasks. They relate well to LaLonde's (1984) three general categories of business decisions: operational, tactical, and strategic. In contrast, Wheelwright and Makridakis (1985) use a four member taxonomy of immediate, short term, medium term, and long term forecasts. Most agree that the period of time between when a forecast is made, and the period of time for which it is made can be called the forecast horizon. The taxonomy of Wheelwright and Makridakis place forecasts with horizons of one to three months as short term forecasts, while those with horizons of two years or more are classified as long term. The distinction is clearly business specific.

A second way to categorize forecasts relates to the forecasting approach utilized. Chambers, Mullick, and Smith (1974) used the classifications of qualitative, causal, and time series/projection. Wheelwright and Makridakis (1985) classified approaches as judgmental, quantitative, and technological. Armstrong (1985) used a methodology tree with branches of subjective and objective. Objective was then further divided into naive (extrapolative) and causal. This dissertation considered the three approaches to be subjective, causal, and extrapolative. This is consistent with the categorization presented by Makridakis (1984). This dissertation deals in depth only with the extrapolative or time series approach.

There has been considerable recent activity to answer the question as to what extrapolative forecasting techniques are better. The viewpoint that the more complex techniques such as Box-Jenkins with differencing functions and autocorrelation analysis would naturally be more accurate than more automatic (simpler) approaches has been strongly questioned in the management science literature. Several empirical studies have indicated that the simpler techniques perform very well and are often superior to the more sophisticated techniques. For example, Makridakis and Hibon (1979) concluded that a forecaster would have done as well with simple methods such as deseasonalized exponential smoothing as with the more complicated Box-Jenkins analysis in forecasts from their set of 111 time series. These, and other more recent findings, suggest that the field of simple extrapolative techniques should be studied further.

Several authors including Lawrence (1983), Winkler (1983), Mahmoud (1984), Weiss and Andersen (1984), Kucukemiroglu and Ord (1985), and Gardner and McKenzie (1985) indicated that robust extrapolative techniques are needed for improved forecasting. Levenbach and Cleary (1984) indicated that robust estimators not only should be less sensitive to the true underlying distribution of the data, but also must be resistant to outliers. Robust extrapolative techniques should perform well over different types of series. Robustness is one of the reasons for the recent upsurge of interest in combinations of forecasts. Winkler (1983) indicated that the use of averaged forecasts from several techniques was robust in his study, with good results for different types of series. Sanders and Ritzman (1987) reported similar results. Some researchers, such as

Huss (1985), have found that the combination of an extrapolative forecast and subjective adjustment of that forecast provides results comparable to more sophisticated techniques. Kucukemiroglu and Ord (1985) tried to achieve robustness through use of least absolute deviation fitting, with disappointing results.

This dissertation explores a different approach toward achieving robustness for extrapolative forecasts. This approach entails blending of nonparametric statistical techniques with time series analysis. The field of nonparametric statistics has been continually expanding since the 1940's. General results of extensions of nonparametrics to date have included a marked increase in robustness. Nonparametric methods are less sensitive to violations of the usual normality assumptions and to outliers. Typically, they are based upon much less restrictive assumptions as to the underlying distribution. While parametric methods will usually give somewhat better results if the normality assumptions are true, their performance often degrades rapidly as these assumptions fail. The writer hypothesized that the extension of nonparametric concepts into the problem of extrapolative forecasting would yield many of the same benefits already achieved in other areas. Thus, this dissertation attempts to provide insight into the possible gains of extending nonparametric techniques to extrapolative forecasting.

1.1 Environment

There are many reasons why managers and analysts want to know about the future. Primarily, perfect knowledge about the future would allow perfect decision making in the present. But the lack of this foresight leads managers and analysts to do the next best thing, i.e., managers actively consider future possibilities, and make decisions and contingency plans based upon them. Sometimes the entire bundle of future possibilities is considered, and at other times only one possibility is forecasted. Virtually all decision making involves some sort of forecast. Purposes vary from cash budgeting to capital budgeting, from sales quotas to inventory control. Approaches for forecasting these futures can be divided into the three categories of subjective forecasting, causal forecasting, and extrapolative forecasting.

The subjective forecasting approach relies upon the learned opinions of the manager, the expert, and/or the analyst. Examples range from simple expert opinion to the more complex Delphi method. The subjective approach is most useful when there are severe data constraints or when the forecast horizon is long range. Mentzer and Cox (1984) indicated that most managers were very familiar with subjective methods. Makridakis, Wheelwright, and McGee (1983) cited its use for new product demand while Chambers, Mullick, and Smith (1971 and 1974) indicated a subjective approach, the Delphi method, was successfully used to estimate demand for new products, and (in 1962-63) to forecast a date for the first landing on the moon. Other examples of Delphi use include a large scale Society of Mechanical Engineers study on robotics.

The causal forecasting approach relies upon a good historical data base and properly developed causal relationships. Examples range from fairly simple equations to complex econometric models. The causal approach is most useful when the historical data are complete and when the causal relationships that have existed in the past are understood and remain unchanged. Ballou (1985) indicated that this approach is useful over forecasting horizons of intermediate and long range. Armstrong (1985) indicated that the causal approach was preferred over the extrapolative approach for long range forecasting.

The extrapolative (time series) approach also relies upon a good historical data base, but does not require explicit identification of causal relationships, or data on causal variables. Examples range from simple naive techniques through moving averages to the more complex Box-Jenkins methodology. The extrapolative approach is most useful for short range cases, and is the focus of this study.

1.2 Extrapolative Forecasting

This approach is based upon the premise that, in the short run, historical patterns will continue. Although the causal variables are not explicitly identified, it is assumed that they are generally known and are not expected to change in their impact on the measure of concern in the forecast horizon. To the extent that this assumption cannot be met, the rationale for using the extrapolative approach is weakened. In most situations, the assumption is deemed to be adequately met and the analyst is faced with a large spectrum of possible extrapolative techniques. The analysis of the time series of data, the choice of an appropriate technique, and the development of an appropriate model within the technique are central to time series analysis and extrapolative forecasting.

Although all the techniques are variations on the theme of averaging, the mechanics can be quite complex because the time series components of trend, cycle, seasonality, and randomness must be addressed. Since the model developed is used for short range forecasting, it is commonly rerun at frequent intervals as time advances and newer data become available. The model thus seeks a proper balance between responsiveness and stability. It should respond reasonably fast to a shift in the underlying pattern, yet should not over react to the expected randomness of the data since this would suggest changes in trend, rate, etc. which have not occurred; making the management job more difficult and costly. Because these objectives are in direct opposition, the model selected must be appropriately balanced.

Two distinct phases are recommended in model development or selection. The first phase is fitting. Here the chosen model type is best fitted to the bulk of the historical data, with model parameters selected either automatically or subjectively. They are most commonly fitted automatically by a computer program using a least squares best fit criterion. In other words, the parameters are chosen which result in the smallest mean squared error for one period ahead forecasts over the historical data. This is like the historical simulation discussed by Graver (1981) where for each time t within the data base one assumes that only information up to that time is available. That information is used to generate a forecast for the next time period t+1. Comparison of the forecasts with the actuals generates a set of error terms which can be analyzed to determine the goodness of fit, or how well the model would have forecast over the entire period of the historical data used for fitting.

The assumption that a model best fit for one period ahead forecasts will also be best for longer term forecasts is questionable. Several authors have addressed this issue, with mixed results. Dalrymple and King (1981) concluded that one period ahead fitting was a reasonable approach, but that sometimes separate models for different horizons reduced forecasting errors. Carbone and Makridakis (1986) developed a technique that shifts weight from a short term model to a long term model as the horizon lengthens, thus supporting the need for separate models. This question is not addressed futher in this research. All simulation work is done using one period ahead forecasts. The empirical study in Chapter VI does consider forecasts with horizons up to six periods ahead, with the best model chosen based upon one period ahead fitting.

The second phase is forecasting. How well has the model forecasted recently? To answer this question, a set of recent hold-out data is generally used. This requires that the most recent portion of the historical data not be initially used in the fitting exercise. The best fit model is used to forecast values of the time series through the period of the hold out data. The set of error terms generated by this forecasting exercise is considered to be a better test of the model's forecasting ability. The forecaster's concern is not how well the model fit the historical data, but how well it would have forecast recently. If the forecaster were satisfied with the model's recent forecasting performance, the hold out data would then be included in the data used for fitting and the model parameters would be updated. If the model had been forecasting well, changes to the model parameters should be minor.

The extrapolative approach holds an advantage in ease of short term forecasting. This advantage is largely due to the reduced data requirements as compared to the causal approach. Since it is easily automated, thousands of short term forecasts can be updated quickly; a clear advantage over the subjective and causal approaches when the resource of time is critical.

1.3 Empirical Results

Although the extrapolative approach is less demanding in terms of time and data, it is not necessarily an inferior forecast. The superiority of quantitative forecasts over strictly subjective forecasts under repetitive forecasting conditions is generally accepted and is well supported in the literature, with few exceptions [for one exception see Lawrence (1983)]. While subjective forecasting techniques perform well on very long range (long horizon) situations, these typically are not performed repetitively. Both the causal approach and the extrapolative approach generate quantitative models. Further results which compare quantitative approaches frequently indicate that causal models perform no better than extrapolative models.

The extrapolative approach has been shown to be as good as, or better than, other forecasting approaches for short term forecasts. A physical analogy is that a large system usually has sufficient momentum to retain its speed and direction for the immediate future. Recent empirical tests have almost unanimously shown that the simpler extrapolative techniques, properly applied, compete very successfully with the more complex techniques such as Box-Jenkins. Armstrong (1985) refers to the complicated time series techniques, such as Box-Jenkins and spectral analysis, as "rain dances" and calls for a return to simple methods for short term forecasting. Wheelwright and Makridakis (1985) indicate that simpler methods do well for noisy series and may be desirable in other cases where any increase in accuracy from use of more complicated methods would require additional cost. Makridakis, Andersen, Carbone, Fildes, Hibon, Lewandowski, Newton, Parzen, and Winkler (1982) relate that simpler methods can be as accurate as more complex, sophisticated methods.

1.4 Nonparametric Philosophy

The development of nonparametrics has been significant since the 1940's. Hollander and Wolfe (1973) relate numerous nonparametric alternatives for many parametric statistical tests. Work on applying nonparametric techniques to the regression problem is continuing as evidenced by Hussain and Sprent (1983) and Hardle and Gasser (1984). The work in nonparametric regression should be compared with the robust regression approaches of Conover (1980), and the R-estimation discussed by Hogg (1979), both based on ranks. Conover, for example, ranked the raw data, did least squares best fit regression on the ranks, then interpolated for an estimate. The robust regression technique commonly referred to as M-estimation, see Draper and Smith (1981), which reduces the weight given to points determined to be outliers, is not nonparametric. Based upon review of the literature, one vital area which had not been actively considered for extension of nonparametric techniques was the extrapolative forecasting problem.

The philosophy of the nonparametrician is to make as few assumptions about the underlying population as possible. Although some of the early tests were championed for their simplicity, it has become evident through simulation studies that the true worth of nonparametric techniques is their increased power versus parametric tests when standard normality assumptions are violated, see Hollander and Wolfe (1973) and Noether (1984). Another important advantage of nonparametric statistical tests is their relative insensitivity to outliers. Robustness and insensitivity to outliers would be very desirable characteristics for a forecasting technique.

CHAPTER II

LITERATURE REVIEW

The main purpose of this chapter is to suggest that simple extrapolative techniques do not necessarily lack accuracy and to relate the apparent lack of nonparametric extrapolative forecasting techniques. Since the proper test of a forecasting technique is how well it forecasts, a simple forecasting technique that performs well is vastly superior to a complicated technique that performs no better.

2.1 Comparisons Between Forecasting Approaches

The three approaches used for forecasting (subjective, causal, and extrapolative) often have been compared for accuracy. Although any one forecasting approach may prove to be superior in a particular case, several studies have reached conclusions as to their relative accuracy in general. Armstrong (1978 and 1985) concluded that subjective forecasting is not superior to quantitative approaches, even for long range forecasts. After reviewing the psychological literature, Makridakis and Hibon (1979) concluded that quantitative methods outperformed clinical (subjective) judgment in repetitive situations. Carbone and Gorr (1985) concluded that subjective adjustment by students of extrapolative forecasts was not desired if accuracy was the criteria. These results should be tempered by the report by Armstrong (1983) that company managers and analysts can forecast annual earnings more accurately than the extrapolative forecasts. One should keep in mind, however, that company forecasters may have data that is not available for use in the extrapolative technique, and company managers have some control over annual earnings. In addition, Huss (1985) reports that a combination where extrapolative forecasts are subjectively adjusted by experts knowledgeable on the causation system is commonly used and works well in electric load forecasting. Thus, subjective forecasting retains an important role when the necessary conditions are satisfied.

How then do the two quantitative approaches compare? The causal approach would seem to have the logical advantage since more variables can be considered. Armstrong (1978) judged econometric models as not being superior to extrapolative models for short range forecasts, but superior to extrapolative models for long range forecasts. Makridakis and Hibon (1979) merely concluded that the results are mixed. Makridakis et al. (1982) concluded that the causal approach is not necessarily more accurate than the extrapolative approach.

Thus, it is clear that extrapolative forecasting is competitive with the more demanding causal forecasting in terms of accuracy when dealing with short horizons. Development of a causal, regression, or econometric model generally requires a considerable data base. This data base, of course, needs to be adjusted to common definitions throughout. It must contain not only the measure of concern (say the cost) but also various and sundry explanatory variables. Checks and adjustments need to be made during the fitting process for autocorrelation, multicollinearity, heteroscedasticity, insignificant variables, incorrect functional specifications, etc. The cost of developing and then maintaining this data base can be quite high in terms of time, effort, and money. This is in contrast to the extrapolative approach where the only data needed are the time series of the measure of concern. While the fitting requires identification of the trend, seasonal, cycle, and randomness components, this is certainly an easier task than for the causal model. Given that the simpler approach results in competitive forecasts, the rational conclusion is that it should be the first approach considered when the estimate itself is the item of principal interest.

2.2 Comparisons Between Extrapolative Techniques

Until recently it was commonly believed that the more complex extrapolative techniques, such as Box-Jenkins, were superior to all the simpler techniques. This superiority was supported by Newbold and Granger (1974) who cited "the great beauty" of the Box-Jenkins technique since the forecast function could be chosen from analysis of the data. They felt this should result in more accurate forecasts. Their results, based on the analysis of 106 time series, indicated that Box-Jenkins performed better than Holt-Winters (a good extrapolative technique) 73% of the time for horizons of one period, decreasing to 58% for horizons of 4-8 periods. Based upon mean square error, Box-Jenkins forecasts were 20% more accurate than Holt-Winters. These results were not surprising and Priestly, in a comment published with the article, concluded that since Box-Jenkins attempted to fit the model to the series it should be generally superior to the automatic techniques. For the automatic techniques the model is fixed, and only the parameters of the model are fitted. Some results of the Newbold and Granger study are summarized in Table 1.

TABLE 1

Summary of the Newbold and Granger Study: Comparison of Box-Jenkins' and Holt-Winter's Forecasting Performance on 1Ø6 Series

	Percent of the Time that the Method is Better	
Lead Times	Box-Jenkins	Holt-Winters
1	73	27
2	64	36
3	6Ø	4Ø
4	58	42
5	58	42
6	57	43
7	58	42
8	58	42

The superiority was not supported by the Makridakis and Hibon (1979) study. This study specifically addressed the accuracy of extrapolative forecasting techniques. The evaluative performance measures used were mean absolute percentage error, a relative measure of absolute error; and Theil's U, a measure of how well the models forecast change. Complete definitions of these measures are included in Chapter IV. The study dealt with 111 time series using 22 techniques or variations of techniques. They used nine nonseasonal methods on the original data, four seasonal methods on the original data, and then seasonally adjusted the data and used the same nine nonseasonal methods on the adjusted data.

Their conclusions were that the best method depended upon the particular evaluative measure used and on the amount of randomness in the data. They found that simpler techniques performed better than more complex ones, prompting reviewers to question their proper use of several techniques. Chatfield (1979), in particular, was skeptical of the results regarding the Box-Jenkins forecasts. Durbin (1979) cited some of the results as "counter intuitive". In contrast, Armstrong (1978) noted that the Holt-Winters technique had not been properly applied in the previously mentioned Newbold and Granger (1974) study, thus biasing their results towards Box-Jenkins. Results of the Makridakis and Hibon study are summarized in Table 2.

A forecasting competition was run as a follow up to the Makridakis and Hibon study, with results published in Makridakis, et. al. (1982 and 1984). This competition used a total of 1001 time series and is frequently called the M-competition. To counter some of the objections made about the first study, experts in each of the techniques performed analysis and made forecasts for either all 1001 series or a subset of 111 series. The more complex, labor intensive methods were performed only on the subset. Results were then summarized and tabulated.

Comparisons with the results of the previous study indicated that Box-Jenkins performed better than before. However, it was still not clearly better than the Winter's, deseasonalized Brown's linear, or deseasonalized Holt's linear exponential smoothing (extrapolative) techniques. In fact, it was found to be worse than simpler

TABLE 2

Summary of the Makridakis and Hibon Study:
Comparison of Ten Extrapolative Method's
Forecasting Performance on 111 Series

	Mean Absolute Percentage Error						
Method	Model Fitting	Forecasting Horizons 1	2	3	4	5	6
Naive	10.0	14.5	15.0	15.1	15.3	15.6	16.6
Single Moving Average	8.4	12.9	13.6	13.7	13.8	14.3	15.3
Single Exponential Smoothing	8.5	12.8	13.4	13.8	14.0	14.3	15.6
Adaptive Response Rate Exponential Smoothing	9.2	13.0	14.0	14.5	14.7	15.2	16.2
Double Moving Average	9.1	15.0	15.6	16.3	16.6	17.4	18.6
Brown's Linear Exponential Smoothing	8.5	12.9	14.3	14.6	14.9	15.9	17.1
Holt's Linear Exponential Smoothing	9.0	12.0	12.8	13.2	13.7	14.8	16.0
Brown's Quadratic Exponential Smoothing	8.7	12.5	14.0	14.7	15.6	17.0	18.6
Regression	11.4	19.6	20.4	21.1	21.1	21.9	22.8
Box-Jenkins	10.6	14.7	15.0	15.7	16.6	17.1	18.1

extrapolative techniques a majority of the time when a relative linear loss function was assumed. Some of the results are summarized in Tables 3 and 4. A review of the results indicates that the simple deseasonalized methods performed competitively with Box-Jenkins over the horizons reported. Over the shorter horizons, the simpler techniques were frequently superior to Box-Jenkins, while Box-Jenkins seemed to have the advantage for longer horizons. Thus, one is drawn to the conclusion that technique complexity does not assure a more accurate forecasting model.

2.3 Nonparametric Forecasting Techniques

To date, a review of the literature on forecasting indicates remarkably little on the use of nonparametric approaches. The only reference located which relates time series analysis together with nonparametric approaches is the Quantile Estimation Procedure (QEP) developed by Gorr and Hsu (1985). This adaptive filtering technique was designed to provide nonparametric estimates of multivariate regression quantiles. This procedure is not pursued in this dissertation for two reasons. First it does not fit within the category of "simple" extrapolative techniques. Wheelwright and Makridakis (1985) rated eight extrapolative techniques as to their complexity. The simplest, with a complexity rating of two was single exponential smoothing.

TABLE 3

Summary of the M-Competition:
Comparison of Ten Extrapolative Method's
Forecasting Performance on 1001 Series

	Mean Absolute Percentage Errors					
	Model	Forecast Horizon				
Method	Fitting	1	2	3	4	1-6
Naive 2	9.6	9.1	11.3	13.3	14.6	14.4
Single Exponential Smoothing	9.5	8.6	11.6	13.2	14.1	14.1
Single Moving Average	8.4	11.5	14.9	17.0	17.8	17.5
Adaptive Response Rate Exponential Smoothing	10.6	9.4	13.5	14.0	15.3	15.1
Holt's Linear Exponential Smoothing	8.8	8.7	11.0	13.3	15.2	14.8
Brown's Linear Exponential Smoothing	9.0	8.7	10.9	13.8	15.0	14.7
Brown's Quadratic Exponential Smoothing	9.3	9.8	12.7	16.6	18.8	19.1
Regression	15.6	15.5	16.9	19.1	18.3	19.1
Winter's Exponential Smoothing	9.3	8.7	10.9	13.2	14.9	14.7
Automatic AEP* Filtering	9.9	9.1	11.9	13.4	13.7	14.4

* Automatic univariate Adaptive Estimation Procedure (AEP).

TABLE 4

Summary of the M-Competition:
Comparison of Selected Extrapolative Method's
Forecasting Performance on 111 Series

	Mean Absolute Percentage Error					
	Model	Forecast Horizon				
Method	Fitting	1	2	3	4	1-6
Holt's Linear Exponential Smoothing	8.6	7.9	10.5	13.2	15.1	13.8
Brown's Linear Exponential Smoothing	8.3	8.5	10.8	13.3	14.5	13.9
Box-Jenkins	n/a	10.3	10.7	11.4	14.5	13.4

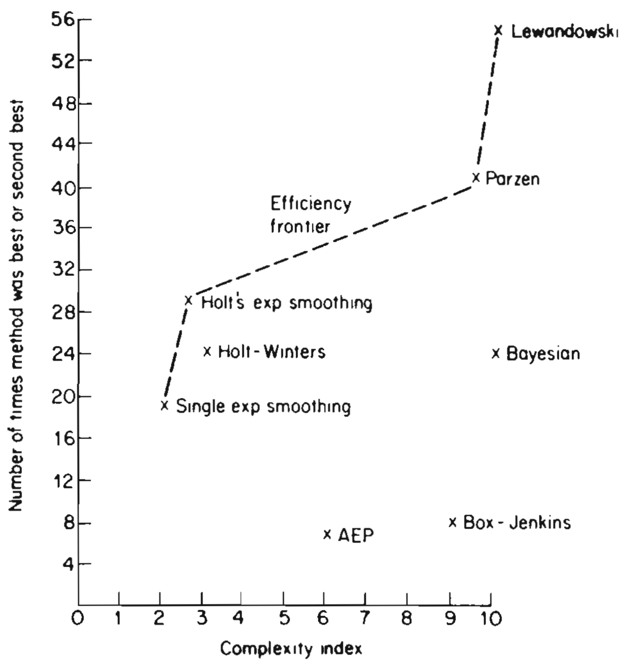

Figure 1. Efficiency Frontier for Time Series Forecasting Methods. If a forecaster is using a technique on the frontier, then higher accuracy will require higher complexity. However, if the technique used is off the frontier, improved performance may be achieved with a less complex technique. (From The Forecasting Accuracy of Major Time Series Methods by Makridakis, et. al., 1984. Reprinted by permission of John Wiley & Sons, Ltd.).

The most complex, with a complexity rating of ten was the Bayesian technique. On this scale, a similar filtering technique (AEP) was rated with a complexity of six. In contrast, the most complicated exponential smoothing technique, Holt-Winters, rated only a three. A second reason why the procedure is not pursued is the poor performance of the filtering technique in the M-competition, rating near the bottom. See Figure 1 from Wheelwright and Makridakis (1985). This figure portrays how they synthesized the results of the M-Competition. In general, they felt that one should get better accuracy for increased complexity. The efficiency frontier is consistent with the concept of "Why pay more for less?" The frontier consists of those techniques that provide the best results for the least effort (complexity). For example, Holt's exponential smoothing with a complexity of 2.5 was best or second best 28 times out of 111. In contrast, the AEP filtering technique with a complexity of 6 was best or second best only 8 times.

Other nonparametric techniques such as those by Hollander and Wolfe (1973) and Conover (1980) have been used for estimation purposes. These will be discussed later as alternatives to the standard simple extrapolative forecasting techniques when the time series has a linear trend.

CHAPTER III

RESEARCH OBJECTIVES

The objective of this research is to explore the area of simple nonparametric extrapolative forecasting, and to determine the relative performance of these techniques in comparison with standard extrapolative techniques. To accomplish this, nonparametric extrapolative methods are proposed and their performance is compared to standard extrapolative methods on a number of different types of simulated series. Only stationary series and series with linear trend are considered in the simulation. Seasonality is not addressed to limit the research to a more manageable size. Also, Makridakis et al. (1982) found that standard deseasonalizing approaches are adequate to make simple extrapolative techniques competitive with the more complex techniques that are designed to explicitly deal with seasonality. When techniques are compared based upon their ability to forecast seasonal data, the effect of the deseasonalize and reseasonalize steps would need to be included. In this study, none of the techniques are designed to handle seasonality. Following the simulation, the performance of the nonparametric techniques is compared to the extrapolative technique's on deseasonalized series used in the M-competition.

3.1 Stationary Series

According to Chatfield (1980), a stationary time series has no systematic change in its mean, or in its variance, over time. In time series analysis, the mean is called the level of the series. No systematic change in the level requires that the series not have a trend, and that strict periodicities not be present. Stationarity does not mean that the observed time values are independent. In fact, the autocorrelation of values "tau" time periods apart is a significant tool for more complex extrapolative methods such as Box-Jenkins. Thus there are any number of models for a stationary time series. A series is said to be weakly stationary if the expected value of the series for any time period is the level; the variance is constant; and if the autocorrelation between any two points in the series is a function of their separation in time. For example, the simplest type of stationary series is that with a constant level, constant variance, and no autocorrelation between error terms. Chatfield (1980) called this a purely random process or white noise. Stationary series also can be generated by moving average processes, by autoregressive processes, or by mixed processes which are a combination of the two.

This research could not perform simulation studies on every possible type of stationary series, but instead used an autoregressive model of order one with the form:

$$X_t = \Psi + (\Phi) X_{t-1} + e_t . \qquad (3.1)$$

Very different time series models are then generated by varying the values of psi and phi.

One problem with actual time series is that the level will sometimes change. If the measure of concern being forecasted is a demand rate for a particular item in inventory, the technique used to generate this forecast needs to react to this increased or decreased demand. Unfortunately, an outlier looks like a change in level until more data points are received. The technique thus needs to balance responsiveness, the ability to quickly react to a change in level, with stability, the ability to weather the expected variation in the data. The simulation study does not consider changes in level.

A simulation experiment comparing forecasting results for simple parametric and nonparametric techniques is discussed in Chapter V. Error term distributions used were normal and Cauchy. The results are tabulated and performance is evaluated based upon the measures discussed in Chapter IV.

The techniques are then applied to a deseasonalized, no monotonic trend subset of the series used in the M-competition. Results are tabulated and discussed.

3.2 Series With Linear Trend

An ideal series with linear trend is simply an inclined straight line with the stationary series as discussed in the previous section superimposed around it. This portion of the experiment considers two standard techniques used with this type series. They are the double moving average and Brown's (1963) linear exponential smoothing. One nonparametric alternative is similar to the approach to nonparametric regression related by Hollander and Wolfe (1973). A second nonparametric approach considered is the double moving median. This approach is similar to the double moving average, but uses medians. A third nonparametric technique is the double smoothed median, analogous to Brown's linear exponential smoothing.

These techniques are applied in a simulation study on series with linear trend, with results tabulated and compared in a manner similar to the stationary series case.

Following the simulation study, the techniques are applied to a deseasonalized, monotonic trend subset of the series used in the M-competition. Results are again tabulated and discussed.

3.3 Limitations

As was noted, the complexities of cycle and seasonality are not included in this study. Likewise, only selected standard techniques are compared to selected nonparametric alternatives. In addition, since there are an infinity of possible stationary series, only a small subset of these are considered in the simulation study. The empirical study deals only with the 111 series subset of the M-competition data, as selected by Makridakis et al. (1982).

While fitting during the simulation and empirical portions allows identification of separate models by error measure, one period ahead fitting is used throughout the study. All simulation results deal with horizons of length one. The empirical study deals with horizons of length one through six, ignoring series observations beyond that point.

All fitting utilizes a grid search technique, thus actual best fit parameters are not determined. Simulation results in Chapter VI are reported by the grid values.

CHAPTER IV

SELECTION OF ERROR MEASURES

An early step in any analysis must be the determination of how the results will be evaluated. Failure to make this determination before a study has been performed can result in biased reports since the evaluation method that supports the desired conclusions could be selected afterwards. To preclude the possibility in this study, a thorough review was made of how forecasting techniques are evaluated so that proper measures could be used in the simulation and empirical tests. This incorporates a specific study of the forecasting literature for qualitative and quantitative evaluation criteria, resulting in identification of several distinct types of quantitative error measures. Since this research deals with simulated and actual series for which qualitative evaluation is either impossible or impractical, further work is done to select the particular quantitative measures for use by both canvassing error measures used in recent forecasting studies and through a simulation study which evaluates the relationship between four selected error measures. The simulation then considers their relationships under conditions of different dispersion and different probabilities of outliers.

While the techniques used in this research are evaluated based strictly upon statistical measures, a number of nonstatistical criteria are used to evaluate forecasting techniques and forecasting models. Before a forecast can fulfill its desired role, the forecast must be accepted and decisions made based upon it. It is thus important for the decision maker to believe the forecast and to understand the technique used. A poor match between the decision maker and the technique will decrease the likelihood of acceptance of the forecast. One nonstatistical criterion therefore must be a reasonable match between the decisionmaker's present expectations/capabilities and the technique used for the forecast. Several authors including Wheelwright and Clark (1976) indicate that models for a nonsophisticated company should be based upon simple logic and techniques. Other nonstatistical criteria include the cost of the forecast, and the reasonableness of the forecast. Techniques whose costs (in terms of computer usage, labor hours, data collection, etc.) outweigh the potential benefits should not be used. One should note, however, that the importance of these various costs may change with the advances in computer technology and the resulting drop in the cost of computations.

Regarding statistical measures, Mahmoud (1984) surveyed over 100 forecasting articles. He listed several statistical measures used to evaluate iterative forecasting techniques, i.e. where forecasts are reaccomplished on a regular basis. Other measures also had been proposed. A number of these are listed in Table 5.

Table 5
Some Accuracy Statistics Used to
Compare Forecasting Techniques

Mean Error *
Average Ranking +
Mean Squared Error *
Theil's U Statistic *
Sum of Squared Errors #
Root Mean Squared Error *
Mean Absolute Deviation *
Coefficient of Variation *
Adjusted Theil's U Statistic *
McLaughlin's Batting Average -
Coefficient of Determination *
Mean Squared Percentage Error **
Mean Absolute Percentage Error *
Median Absolute Percentage Error +
Adjusted Mean Absolute Percentage Error *

* Mahmoud (1984)
\- McLaughlin (1972)
\+ Makridakis, et. al. (1982)
\# Granger and Ramanthan (1984)
** Wright (1986)

It should be clear that use of all the error measures listed in Table 5 is not possible. Further research is performed to get a better understanding of the principal error measurement approaches.

4.1 How are iterative forecasts evaluated?

Iterative forecasting techniques are usually compared based upon the errors that would have occurred if the techniques had been used with the historical data. Generally, the historical data are divided into two sets as discussed in Section 1.2. The earliest set typically is used to fit the selected model. The "best fit" model then is used to forecast over the most recent set of data. The techniques are compared based upon how well they would have forecasted. A technique that would have generated smaller errors on average would usually be considered the better technique. This leads to a problem in that an unbiased technique will have an average error of about zero. An additional complication is that sometimes the proper impact of an error is better indicated by taking the absolute value of the error term, sometimes by squaring the error term, sometimes by considering one of the preceding as a percentage of the series value, and sometimes by considering the amount of change properly forecasted. The choice depends upon the forecaster's (or more ideally, the manager's) loss function. While a complete discussion of loss functions is outside the scope of this effort, it is generally correct to say that sometimes several small errors are preferable to one large one. A somewhat more extensive discussion of loss functions is in Kankey and Thompson (1986). These realities led to consideration of several forms of error measures.

4.1.1 Absolute and Squared Measures

Let the time series of data be expressed as a set of X_t, with t denoting the index of the time period. Let the forecast for time period t be denoted by F_t. The error term for that time period is defined as:

$$e_t = X_t - F_t \ . \tag{4.1}$$

The absolute error is then designated as:

$$|e_t| = |X_t - F_t| \ . \tag{4.2}$$

The average absolute error, more commonly called the Mean Absolute Deviation (MAD), is then noted as:

$$MAD = [\ \Sigma\ |e_t|] \ / \ n \ , \tag{4.3}$$

where indicates the summation of the elements in the brackets, and n denotes the number of error terms used in the calculation. Here n may or may not be the total number of data points available. The definition holds in either case. If holdout data are being used, then the measure of fit (MAD in this case) over the holdout data period would be of concern. Thus, n would be the number of holdout data points. The MAD can of course be calculated for the fitting period, or for the whole set of historical data, prompting different possibilities for the number n. This interpretation for n is consistent across the different error measures considered. Use of MAD for a model is consistent with the existence of a linear loss function.

Given the same definitions, the squared error term for time period t can be denoted by:

$$(e_t)^2 = (X_t - F_t)^2 \ . \qquad (4.4)$$

The average squared error, more commonly called the Mean Squared Error (MSE), is then computed as:

$$MSE = [\ \Sigma\ (e_t)^2] \ /\ n \ . \qquad (4.5)$$

Use of the MSE is consistent with a squared loss function.

The choice of whether an absolute or squared error measure is more appropriate is driven by the loss function perceived by the management of the firm. Since the results of this research should be generalizable to either situation, both absolute and squared error measures must be considered.

4.1.2 Raw versus Relative Measures

The raw measures of error, such as those discussed in the previous section, are expressed in terms of the time series itself. If one series is in kilowatts, then the MAD is in kilowatts while the MSE is in kilowatts squared. A second series might be in millions of dollars, with a MAD likewise in millions of dollars and an MSE in millions of dollars squared. A natural question that should be addressed is whether averaging raw error measures over different series, with different levels and different dispersions, is reasonable. Guerts (1983) states they are not useful. Newbold (1983) indicates that technique performance rankings based upon average MSE can be affected by scaling of some of the series. Makridakis (1983) acknowledges that there are problems with averaging, but supports it as necessary.

In contrast, relative measures are manipulated so that the units are lost. They are in unitless or relative terms. They are scale invariate. Thus they avoid some of the criticism levied on raw error terms.

Using the same terminology as in the previous section, the percentage error is a relative measure of the error at time period t. It can be expressed as:

$$PE_t = [(X_t - F_t)/X_t] * 100 . \qquad (4.6)$$

As can be seen, this is merely the error at time t divided by the actual value at time t, then converted to a percentage. The most common relative measure developed from the percentage error is the Mean Absolute Percentage Error (MAPE). The MAPE can be written as:

$$MAPE = [\Sigma |PE_t|] / n . \qquad (4.7)$$

The MAPE is a relative absolute error measure and is consistent with a relative absolute loss function.

Relative squared error measures are not commonly used. One measure is the Mean Squared Percentage Error discussed and used by Wright (1986) and by Kankey and Thompson (1986). It is defined as:

$$MSPE = [\Sigma (PE_t)^2] / n . \qquad (4.8)$$

The use of this error measure is consistent with a relative squared error loss function.

It should be noted that the relative measures share a common weakness. While powerful and meaningful in series where X_t values remain well above zero, they tend to become distorted as series values near zero occur. Suppose, for example, that at one point in time, the series value was zero, and that the forecast was positive. That point would cause the MAPE to be undefined. Values of X_t close to zero can produce very large MAPE or MSPE values.

4.1.3 Realization Measures

Other measures have been proposed which attempt to relate how well a forecasting model estimates change. The concept is that one should be concerned with how much of the actual change was anticipated by the forecaster or the forecast model. Both Theil (1966) and McLaughlin (1972) proposed such measures. Theil called his measure the Inequality Coefficient, while McLaughlin called his the Standardized Realization Percent (SR%). Both consider the ratio of predicted change to actual change. McLaughlin's (1972) paper uses absolute values where Theil's (1966) work uses squared differences, but either can be modified. The concepts are equivalent. McLaughlin changes the measure into a percentage and reverses the scale by subtraction from a constant. Perfect forecasting would result in a Theil's U statistic of Ø and a SR% of 4ØØ. The Naive One forecasting technique assumes there will be no change in the future and forecasts future periods at the latest reported value. Use of the Naive One forecast technique would result in a Theil's U of 2 and a SR% of 300. Probably due to the scaling choice with values above 300 interpreted as good, and those below 300 as bad, SR% has become commonly known as McLaughlin's Batting Average. McLaughlin (1972) explained that he found it more intuitive for the better technique to score higher, thus the reversal of scales. Since for all other error measures considered in this study (MAD, MSE, MAPE, MSPE) better techniques have lower values, this argument seems inappropriate here. Because there is a one-to-one relationship between the two types of measures, only the Theil's type measure will be considered.

Theil's U statistic can be explained as the comparison of errors from the most naive forecasting method, called Naive One (discussed briefly above and in more detail in Section 5.1.1), with those from the technique being evaluated. Theil (1966) squared these errors, while Makridakis, Wheelwright, and McGee (1983) essentially squared the percentage errors. Slight modifications to the formulas allow use of absolute errors or absolute percentage errors. Regardless of the formulation of the statistic, the interpretation is consistent. If a forecaster is to devote resources to a forecast technique, the technique should show superior performance at forecasting change, versus the no effort Naive One technique. The statistic values are bounded below by zero (perfect forecasting), yet have no upper bound.

Theil expressed his Inequality Coefficient as:

$$U^2 = \Sigma(P_t - A_t)^2 / \Sigma(A_t)^2 \quad (4.9)$$

where

$$P_t = F_t - X_{t-1} \quad \text{(Predicted Change)}, \quad (4.10)$$

$$A_t = X_t - X_{t-1} \quad \text{(Actual Change)}. \quad (4.11)$$

Theil's Inequality Coefficient, now called Theil's U, is then calculated as the positive square root of the value above. It should be clear that the use of absolute values or ratios in the above equation would be uncomplicated.

These measures also are relative in that they are unitless. They share a problem similar to the other relative measures. Here an instance where no actual change occurs in the series can prompt an undefined measure for that point in time. Theil's U is a measure useful when comparing forecasting techniques over several forecasts, but it may, in fact, not even exist for any one forecast.

4.2 Need for a Variety of Measures

Mahmoud (1984) and those references given in Section 4.1 indicate a wide variety of quantitative error measures for iterative forecasts. Given the division of error measure approaches into absolute versus squared, and raw versus relative, it seems necessary to at least consider the use of one measure of each type. The realization measures that deal with how well techniques forecast change add another dimension to the problem, with the potential for a Theil's type index for each basic measure. A quick count indicates the possibility of eight error measures to cover the spectrum. Before proceeding further in selection of error measures to be used, a survey of recent forecasting literature is presented for those measures most commonly used.

A number of recent forecasting articles are reviewed. The review results are presented in Table 6. Among the articles listed in the table, it is clear that MAPE and MSE are most commonly used. Thus, the choice of a relative absolute error measure and a raw squared error measure is clear. The only raw absolute error measure reported is the MAD, and although only a small proportion of the sample used the MAD, the measure is retained in this study. Based upon the limited use that MSPE received, and upon the conclusions by Kankey and Thompson (1986) that MAPE and MSPE relate closely in most situations, MSPE is not considered further. Although Theil's U is not extensively used in the literature, it is a representative of the realization measures and one form is retained in the study on that basis. The particular form of Theil's U retained is that used by Makridakis, Wheelwright, and McGee (1983), based on squared relative changes.

Table 6

Error Measures Reported in
Fourteen Recent Forecasting Articles

Date	Study	Error Measures			
1974	Newbold & Granger	-	MSE	-	-
1979	Makridakis & Hibon	-	-	MAPE	Theil's U
1980	Gardner & Dannenbring	MAD*	MSE	MAPE*	-
1982	Makridakis, et al	-	MSE	MAPE+	-
1983	Winkler & Makridakis	-	MSE*	MAPE	-
1984	Granger & Ramanthan	-	SSE	-	-
	Weiss & Andersen	MAD	MSE	MAPE	-
1985	Carbone & Gorr	-	-	MAPE	-
	Gardner & McKenzie	-	-	MAPE+	-
	Gorr & Hsu	-	-	MAPE	-
1986	Bustos & Yohai	-	MSE	-	-
	Carbone & Makridakis	-	-	MAPE+	-
	Fomby	MAD	RMSE	-	-
	Wright	MAD	MSE	MAPE	MSPE

KEY: MAD - Mean Absolute Deviation
MAPE - Mean Absolute Percentage Error
MSE - Mean Squared Error
MSPE - Mean Squared Percentage Error
RMSE - Root Mean Squared Error
SSE - Sum of Squared Error
+ Also reported Median Absolute Percentage Error.
* Computed this statistic, but did not report it.

4.3 Simulation Study of Error Measure Relationships

Given the four selected error measures, are they sufficiently different to merit retention of all of them in the evaluation of the nonparametric forecasting techniques? If two of these are highly related in all situations, then perhaps the study can be reduced by dropping one of them. No references explicitly compared the behavior of these error measures under varying conditions. Thus, a simulation study is performed to address this question.

The intent is to evaluate the relationship between each pair of error measures. To do this, simulation results are displayed on simple scatter diagrams and the Spearman's correlations between measures are evaluated.

A time series of data and a set of forecasts are generated. Each element of the time series is distributed about a level (mean) of ten. A forecasting model is then used with this series to generate the described error terms and measures. One might think that it would have been sufficient to generate a set of error terms and then perform the necessary operations to calculate the various measures. But MAPE and Theil's U both require series values for computation. Because of this need, the series of X_t is generated and one of the simplest forecasting techniques, appropriate for this type of series, is used to generate the forecasts. While some would suggest that different forecast techniques might generate error measures that are related differently, this determination is deferred for later work. For now, it is acknowledged that different models will drive errors with different autocorrelation properties. It is felt that the dimension under consideration here, that is across error measures, would not be affected by the model used. Forecasts for this simulation are generated by a single exponential smoothing model (single exponential smoothing is explained in Section 5.1.1).

The building block of the simulation is a set of twenty observations. For each set of twenty observations, MAD, MSE, MAPE, and Theil's U values are calculated. Twenty-six sets of these error measures are calculated per replication. The measures for the first twenty observations of each replication are discarded for two reasons. The first reason is that the forecasts are started by assuming that the first exponential average is the first observed value, and the effects of this choice could conceivably affect the results if the earliest observations are used. The second reason is that both the MAPE and Theil's U calculations lack a prior X value at the first time period. Thus, twenty-six sets of error measures provide twenty-five usable sets of error measures. This is replicated twenty times, resulting in 500 sets of error measures from 10,000 observations. All 500 sets of these error measures are used in the scatter diagrams. Since the relationships are nonlinear, Spearman correlation coefficients are calculated between error measures for each of the 20 replications. Data for box plots as discussed by Tukey (1977) are generated and can be found in Tables 7 and 8.

The series X_t was generated using the following formula:

$$X_t = 10.0 + k*RANNOR(SEED) \qquad (4.12)$$

where RANNOR is the Statistical Analysis System (SAS) standard normal function generator (reference SAS Users Guide: Basics, Version 5 Edition, 1985, pg. 267). This simulation study was performed using the VMS SAS Release 5.03. To test the effect of dispersion on the relationships between the error measures, the values of .6, 1, 1.8, and 2.2 are used for k. The effect of outliers upon the relationships is evaluated by using an approach similar to Bustos and Yohai (1986). As the series values are generated, a uniform random variable is used to control from which of two distributions the error term is drawn. Relatively small probabilities of outliers are used (.005,.01,.02), with the second distribution having dramatically larger dispersion (k = 10 vs. k = 1). Since a very small value from the normal distribution might still not be an outlier in this situation, these results all stipulate that:

$$P(\text{outlier}) \leq Y \qquad (4.13)$$

where $\leq$ is the standard symbol for "less than or equal to", and Y takes on the values .005, .010, and .020. Sample programs are in Appendix A.

4.4 Evaluation of Results

4.4.1 MAD vs. MSE

The data in Table 7, Figure 2, and Figure 8 indicate that these two raw error measures relate well in all dispersion and outlier cases. The most notable, and expected effect is the clear nonlinear relationship as the dispersion or the probability of outliers is increased. MSE should be expected to increase much more rapidly than MAD as errors get larger.

4.4.2 MAD vs. MAPE

The nominal case (sigma = 1, no outliers) results in the highly linear scatter plot illustrated in the lower portion of Figure 3. For this case, the coefficient of variation is .10 which keeps the denominator in the MAPE calculations well away from zero. As the dispersion of the error term is increased to a C.V. of .22, this linear relationship began to include points with higher MAPEs, illustrated by the larger dispersion in the upper portion of Figure 3. It is clear that as the likelihood of X_t values near zero increased, the MAPE calculation:

$$|e_t|/X_t \qquad (4.14)$$

results in some large values on occasion which have a significant impact on the measure. As might be expected, outliers largely destroy this linear relationship, as shown by the scatter plot in Figure 9. The data in Table 7 also reflect the gradual effect of increased dispersion (probably better interpreted as increased Coefficient of Variation). Table 7 data also indicate that outliers have little effect on the Spearman's correlation between MAD and MAPE.

4.4.3 MSE vs. MAPE

Based upon Table 8 and Figure 6, these measures relate well in the nominal case and degrade slowly as dispersion is increased. The relationship is markedly affected by outliers. It is clear from Figure 11 that, in the case of outliers, a high MSE or a low MSE do not assure a high or low MAPE. Several cases with a very high MSE had a low MAPE, while several cases with a very high MAPE had a low MSE. Regardless of these cases, however, the data in Table 8 indicate that the Spearman's Correlation between the two measures tends to increase with increased probability of outliers. Carbone and Armstrong (1982) report that MSE has had a larger usage in the forecasting community.

4.4.4 Theil's U vs. any other measure

It is somewhat surprising to find that the relationship between Theil's U and MAD, MSE, or MAPE tends to be somewhat negative. This is reflected in Tables 7 and 8, and Figures 4, 6, 7, 10, 12, and 13. This lack of systematic relationship is due to the different objective of the standardized realization measures. The realization measures, such as Theil's U, are designed to reflect the ability of a model to forecast change. The others deal with forecasting errors, rather than the amount of change properly forecasted. Based upon the data in Table 8, the relationships between Theil's U and MSE or MAPE tend to be somewhat negative. The results in Table 7 indicate that MAD is consistent with MSE. As the amount of dispersion increases or as the probability of outliers increases, the forecasting model does relatively better than a strict Naive One model. Thus, Theil's U tends to decrease (that is, reflect improvement) while the other error measures tend to increase, therefore the somewhat negative correlation with the other measures. This relationship is consistent when dispersion is increased or when the probability of outliers is increased.

TABLE 7

Spearman's Correlations Between MAD and the Other Error Measures as Series Dispersion and Probabilities of Outliers are Varied

		Characteristic Being Varied							
		Dispersion (Sigma values)					Outliers		
Comparison	Quantile	.6	1.0	1.4	1.8	2.2	.005	.010	.020
MAD	Maximum	.973	.973	.973	.973	.973	.972	.972	.982
vs.	75%	.953	.953	.953	.953	.953	.954	.957	.967
MSE	Mean	.920	.920	.920	.920	.920	.927	.940	.955
	Median	.937	.937	.937	.937	.937	.933	.942	.959
	25%	.881	.881	.881	.881	.881	.905	.925	.949
	Minimum	.838	.838	.838	.838	.838	.838	.875	.906
MAD	Maximum	.994	.990	.979	.958	.940	.990	.992	.992
vs.	75%	.991	.983	.961	.939	.898	.979	.980	.978
MAPE	Mean	.984	.968	.939	.906	.862	.966	.971	.965
	Median	.987	.972	.945	.920	.875	.970	.973	.965
	25%	.982	.954	.923	.886	.846	.957	.962	.953
	Minimum	.960	.934	.864	.788	.713	.934	.945	.937
MAD	Maximum	.454	.436	.442	.419	.384	.493	.384	.257
vs.	75%	.129	.122	.127	.108	.122	.145	.069	-.003
Theil's	Mean	-.090	-.098	-.108	-.124	-.132	-.103	-.136	-.161
U	Median	-.107	-.110	-.080	-.083	-.097	-.081	-.170	-.207
	25%	-.305	-.313	-.354	-.385	-.381	-.367	-.352	-.322
	Minimum	-.448	-.468	-.495	-.529	-.518	-.542	-.517	-.625

TABLE 8

Spearman's Correlations Between MSE, MAPE, and Theil's U as Series Dispersion and Probabilities of Outliers are Varied

		Characteristic Being Varied							
		Dispersion					Outliers		
Comparison	Quantile	.6	1.0	1.4	1.8	2.2	.005	.010	.020
MSE	Maximum	.978	.975	.969	.962	.952	.968	.982	.982
vs.	75%	.949	.942	.945	.931	.908	.941	.949	.958
MAPE	Mean	.918	.911	.900	.886	.861	.909	.925	.937
	Median	.936	.927	.905	.907	.873	.919	.927	.933
	25%	.879	.877	.874	.856	.828	.884	.913	.918
	Minimum	.820	.778	.762	.732	.685	.778	.858	.868
MSE	Maximum	.575	.523	.487	.444	.387	.492	.388	.241
vs.	75%	.127	.088	.023	.029	.003	.145	.086	.023
Theil's	Mean	-.086	-.104	-.126	-.159	-.185	-.138	-.130	-.153
U	Median	-.100	-.096	-.073	-.070	-.163	-.146	-.113	-.212
	25%	-.242	-.290	-.310	-.406	-.473	-.415	-.375	-.312
	Minimum	-.608	-.628	-.658	-.680	-.660	-.602	-.512	-.520
MAPE	Maximum	.448	.437	.433	.420	.369	.492	.355	.224
vs.	75%	.133	.139	.101	.036	-.076	.145	.019	-.122
Theil's	Mean	-.097	-.117	-.166	-.223	-.283	-.138	-.178	-.242
U	Median	-.109	-.178	-.220	-.227	-.282	-.146	-.203	-.218
	25%	-.350	-.360	-.416	-.511	-.544	-.415	-.414	-.402
	Minimum	-.409	-.435	-.567	-.637	-.658	-.602	-.584	-.697

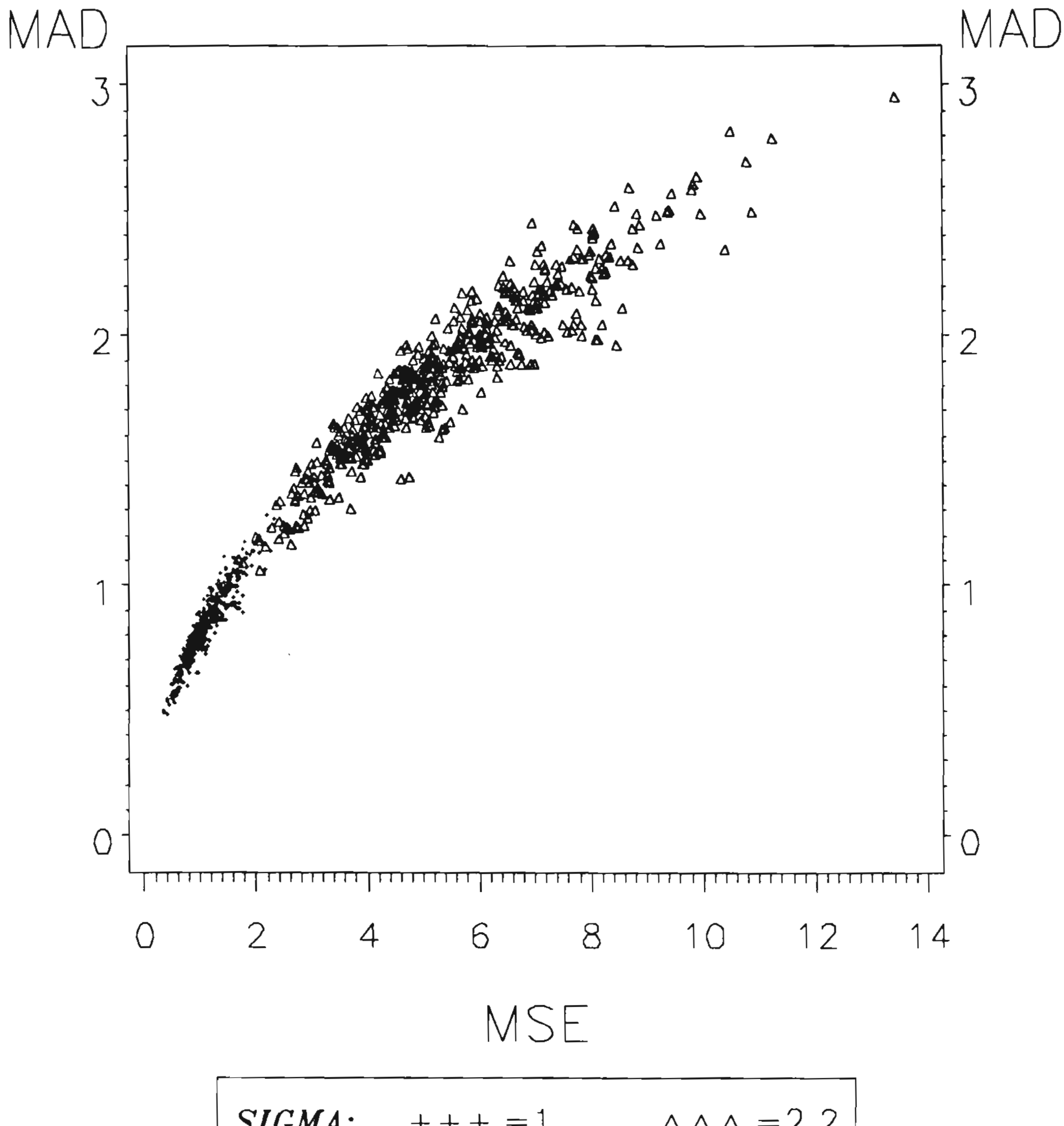

Figure 2. Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAD and MSE. Note the clearly nonlinear relationship between the effect of the smaller deviation (lower left) and the larger deviation (upper right).

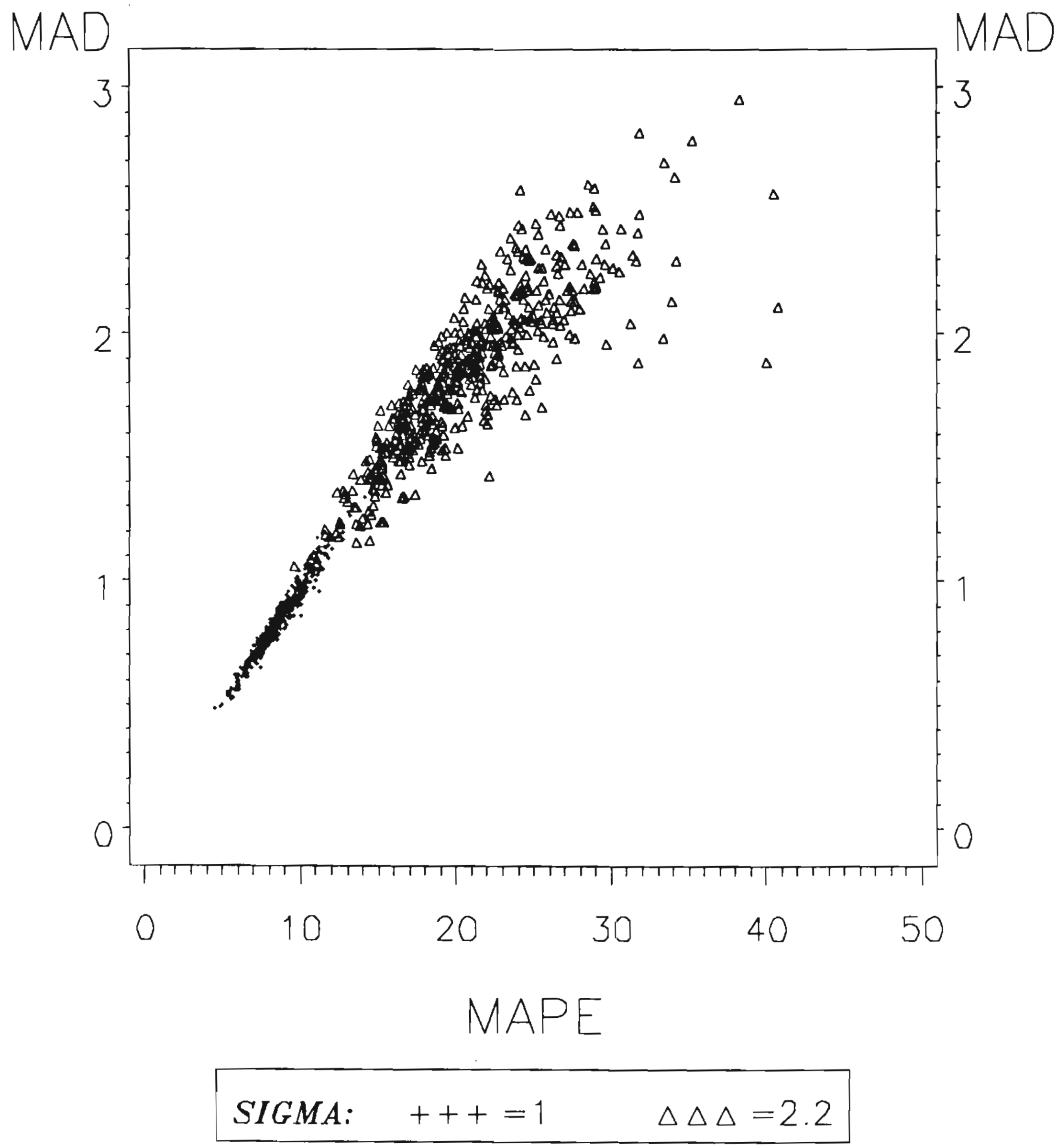

Figure 3. Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAD and MAPE. Note that the relationship remains fairly linear, but becomes less symmetrical with the higher dispersion, with MAPE values tending to increase faster than MAD values.

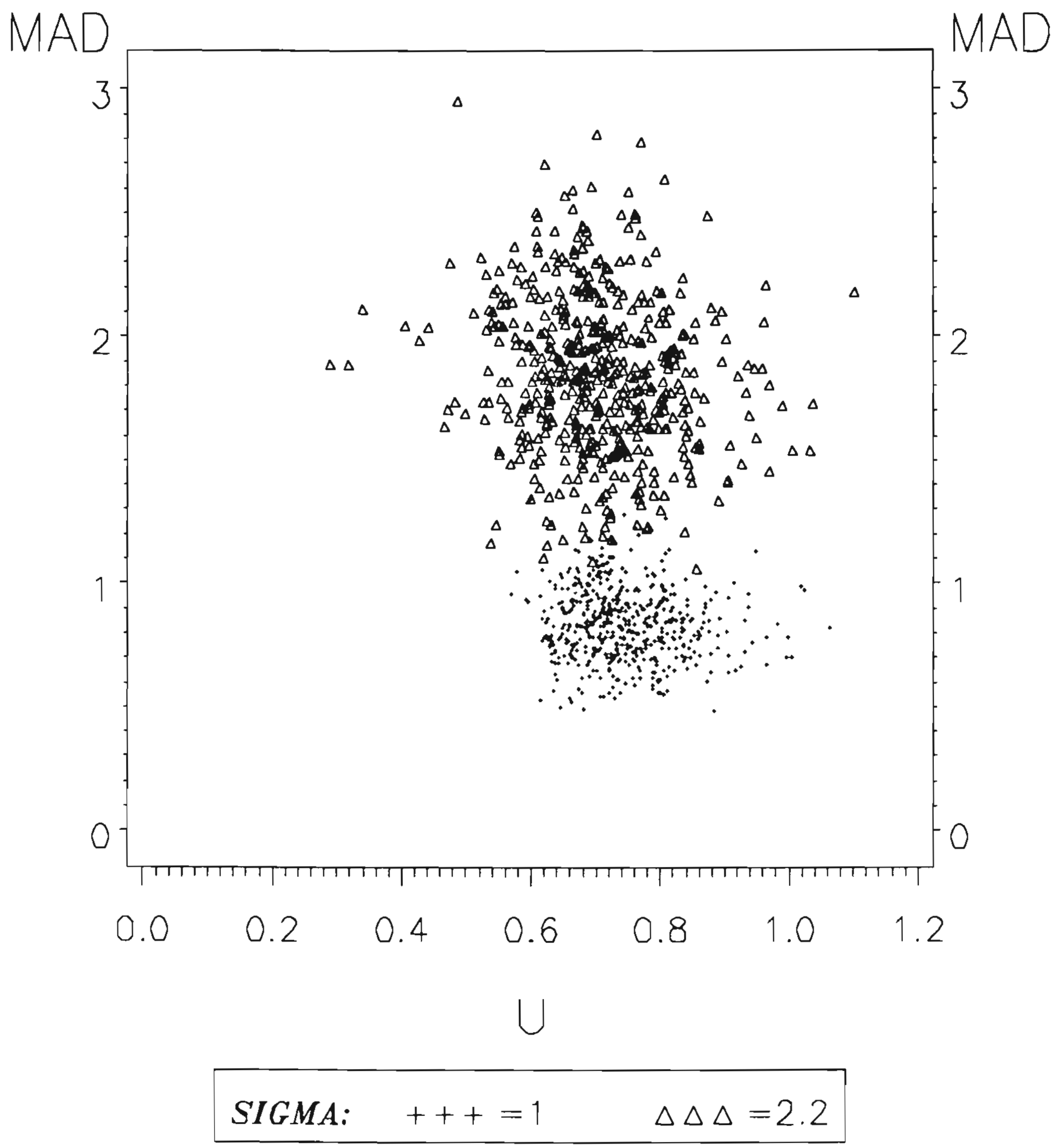

Figure 4. Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAD and Theil's U. Note that while the increased dispersion clearly increases the MAD values, the Theil's U values tend to be somewhat smaller. This is consistent with the small negative relationship indicated in Table 7.

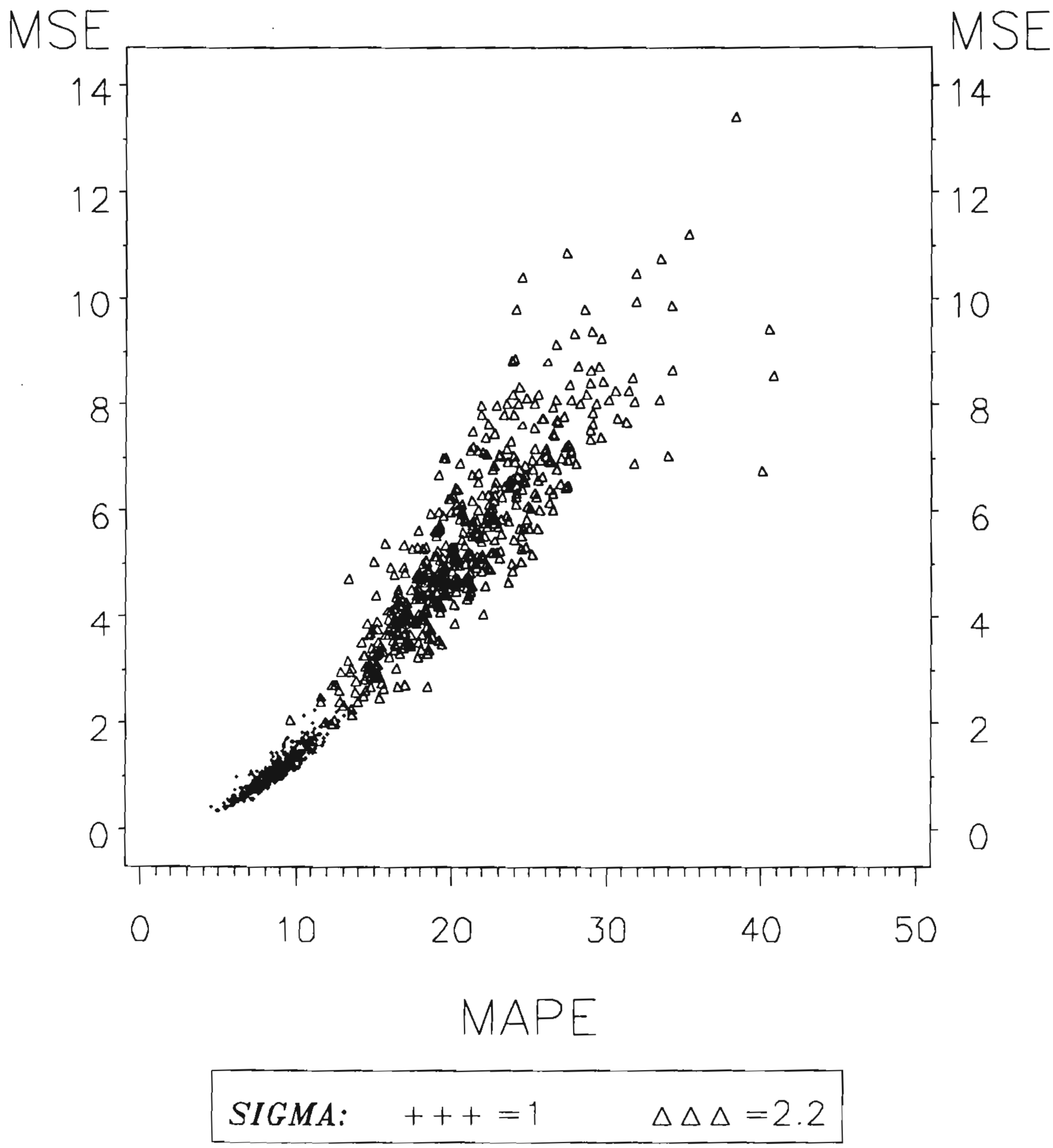

Figure 5. Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MSE and MAPE. Again note the clearly nonlinear relationship.

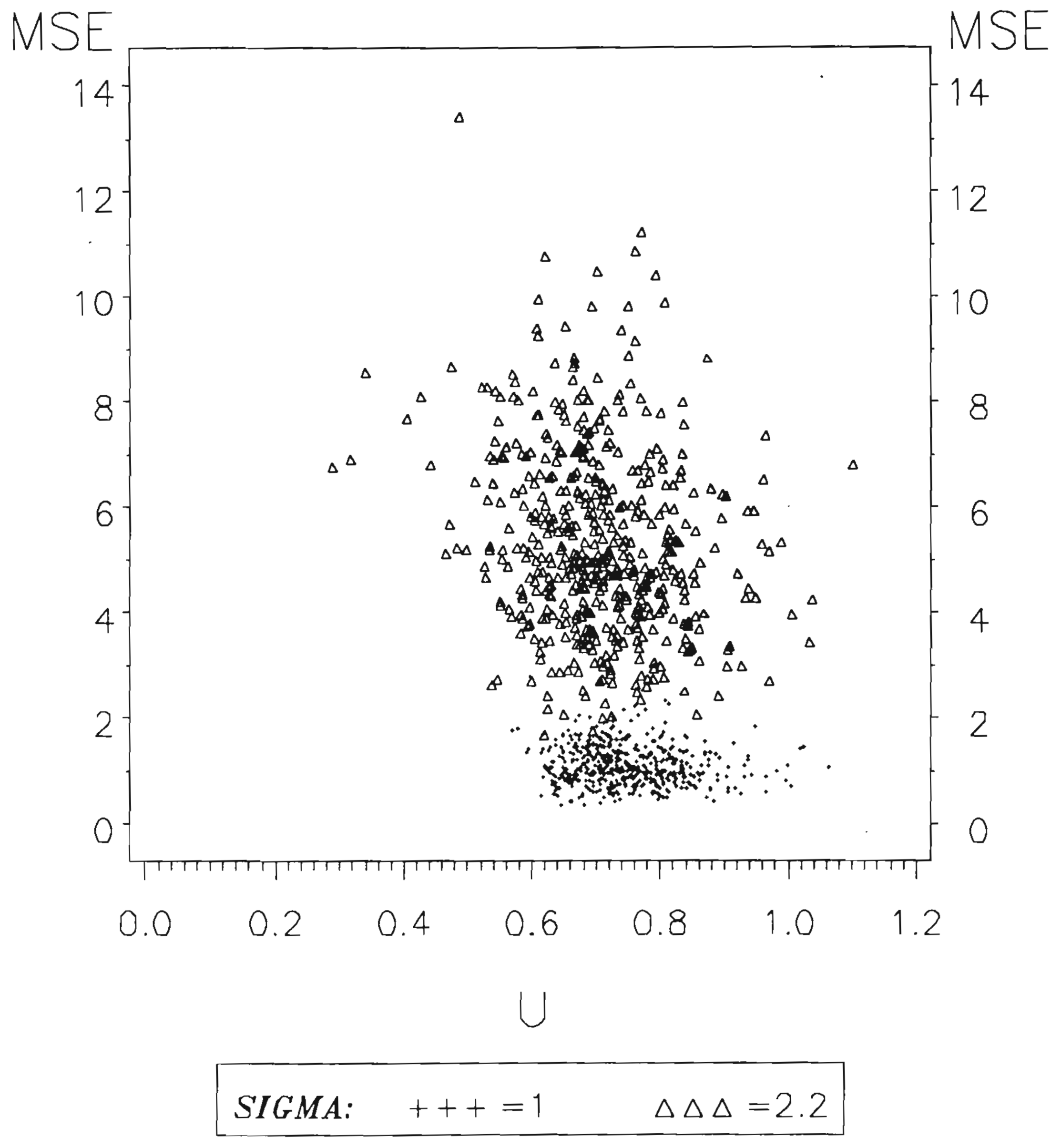

Figure 6. Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MSE and Theil's U. Clearly indicates that increased dispersion had little effect on Theil's U values. Reflects the small negative relationship indicated in Table 8.

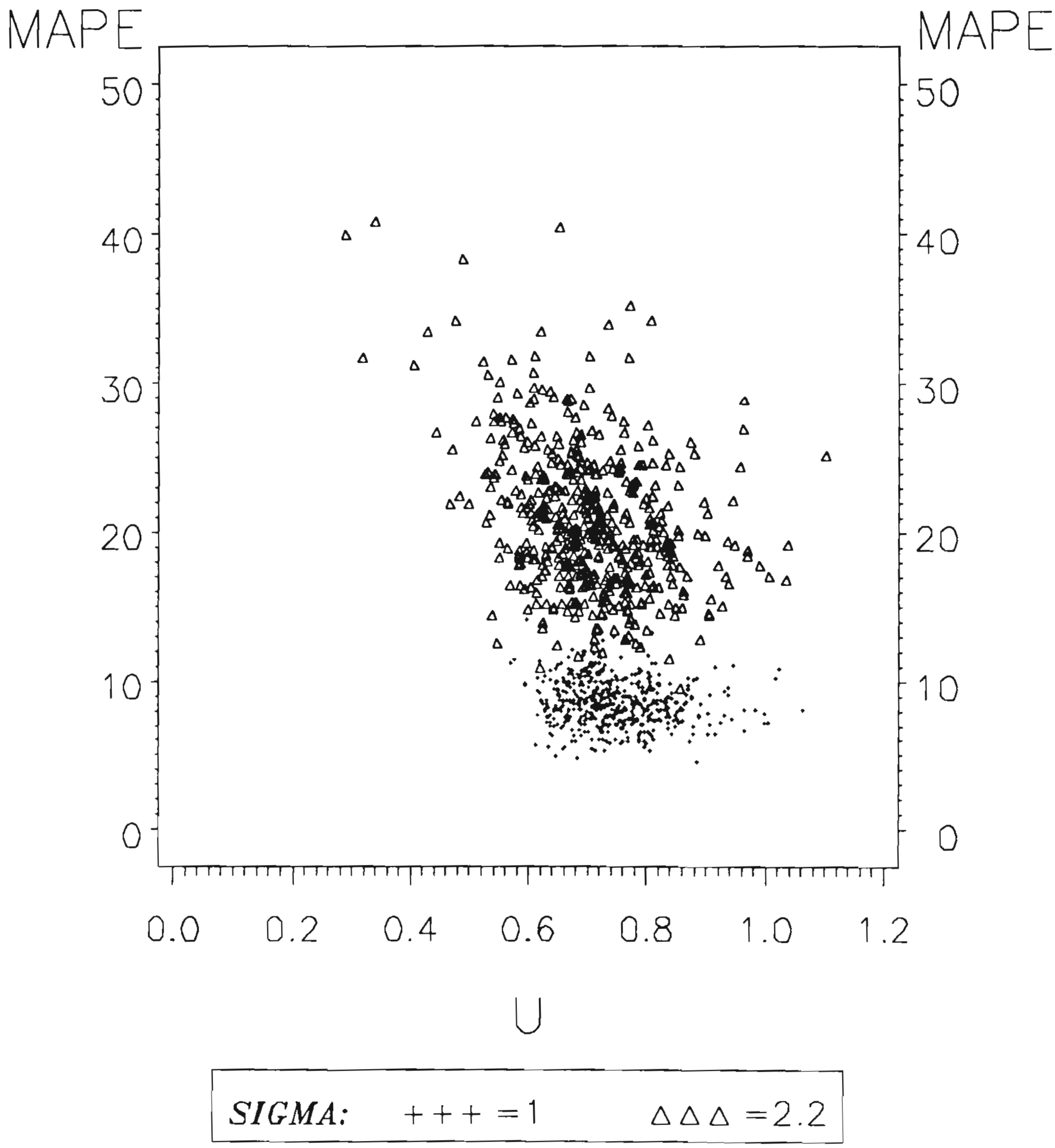

Figure 7. Scatterplot Illustrating the Effect of an Increased Disturbance Standard Deviation of 2.2 versus 1.0 on the Relationship Between MAPE and Theil's U. Note that although both are unitless measures, there is little relationship between them.

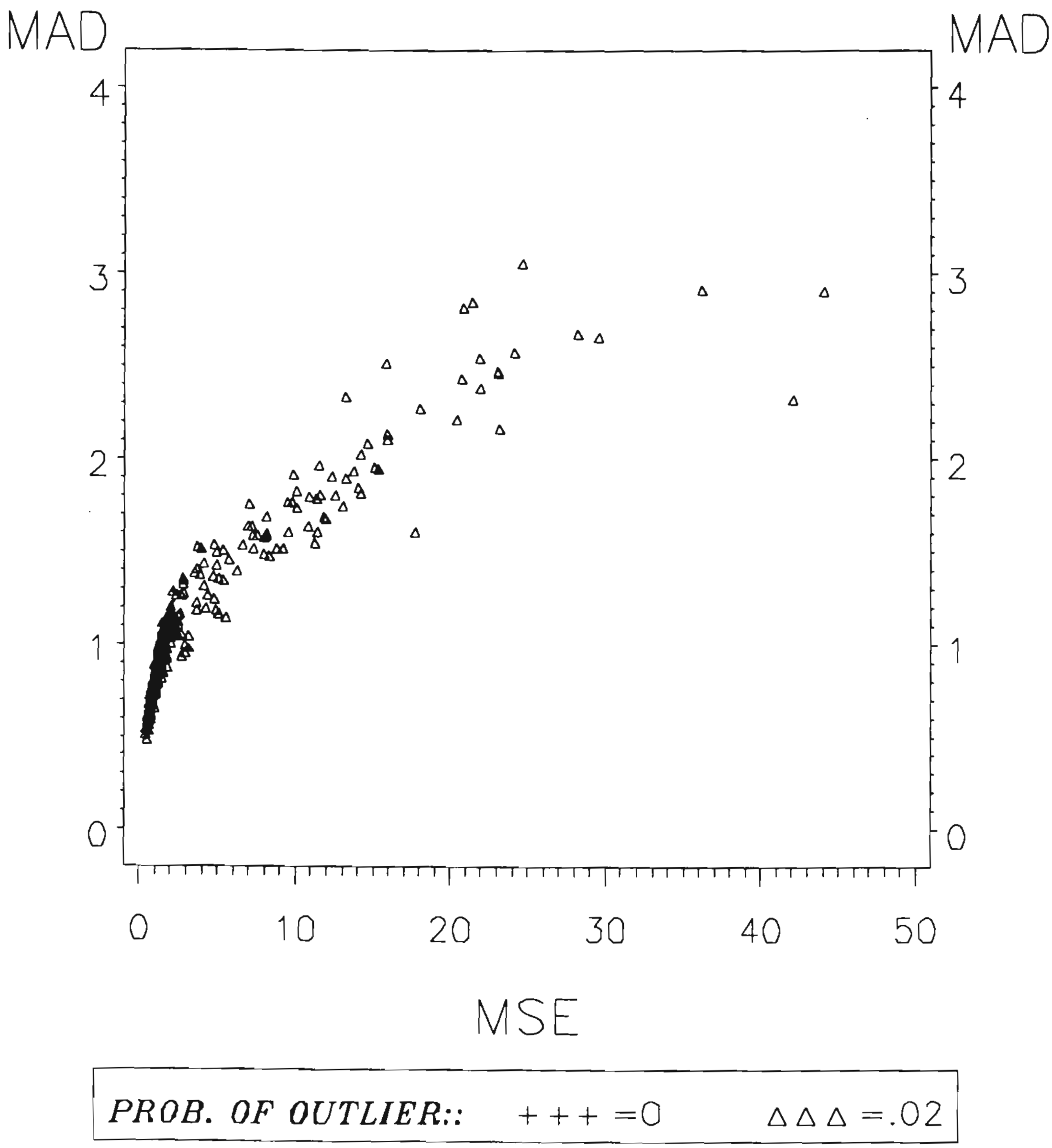

Figure 8. Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MSE and MAD. Reflects a nonlinear relationship with MSE increasing much faster than MAD with the higher level of outliers.

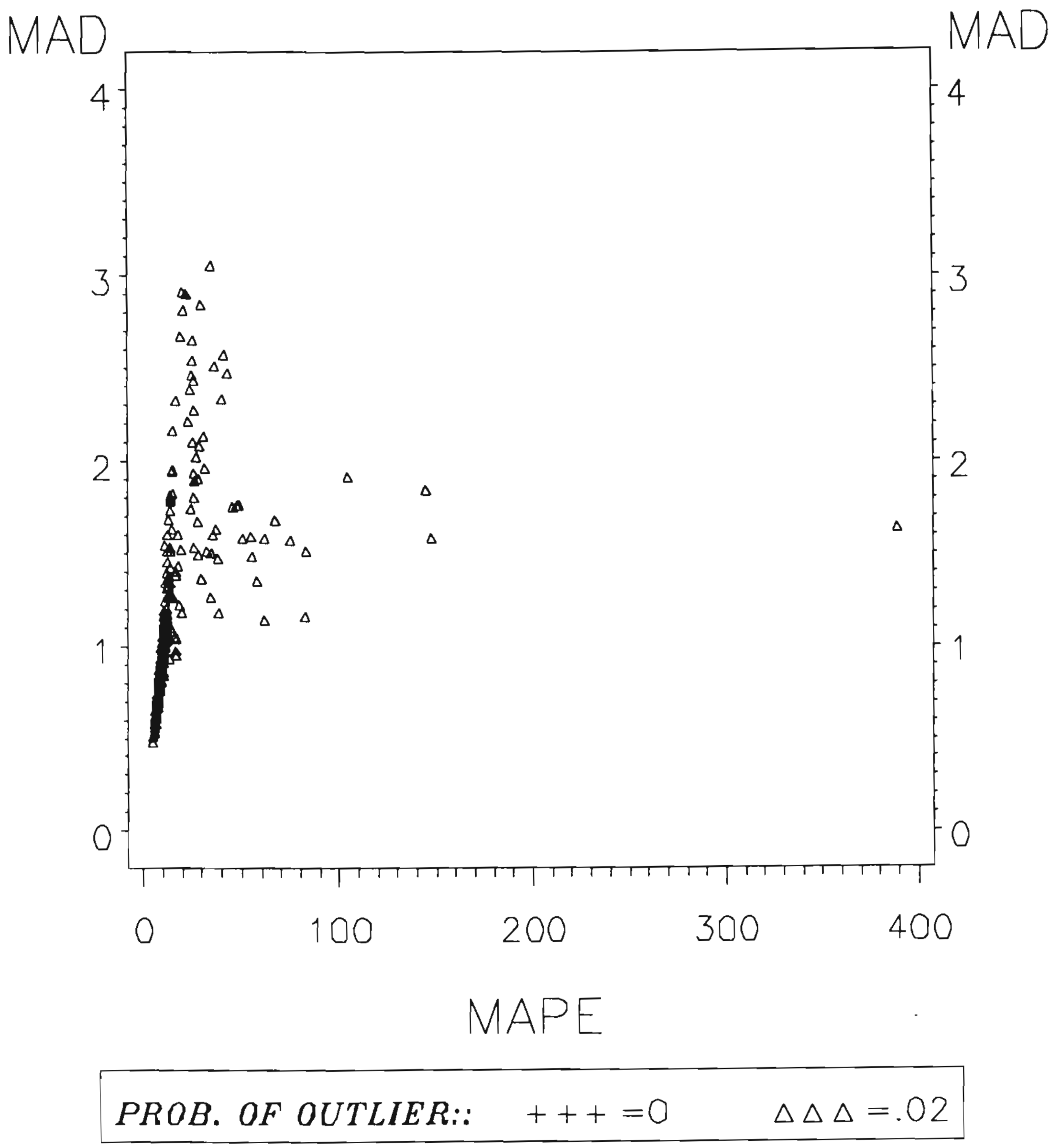

Figure 9. Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MAPE and MAD. Note that on occasion outliers can drive a very large MAPE value, although in most cases the relationship is fairly linear. High MAPE values are driven by observations where the actual series value is small.

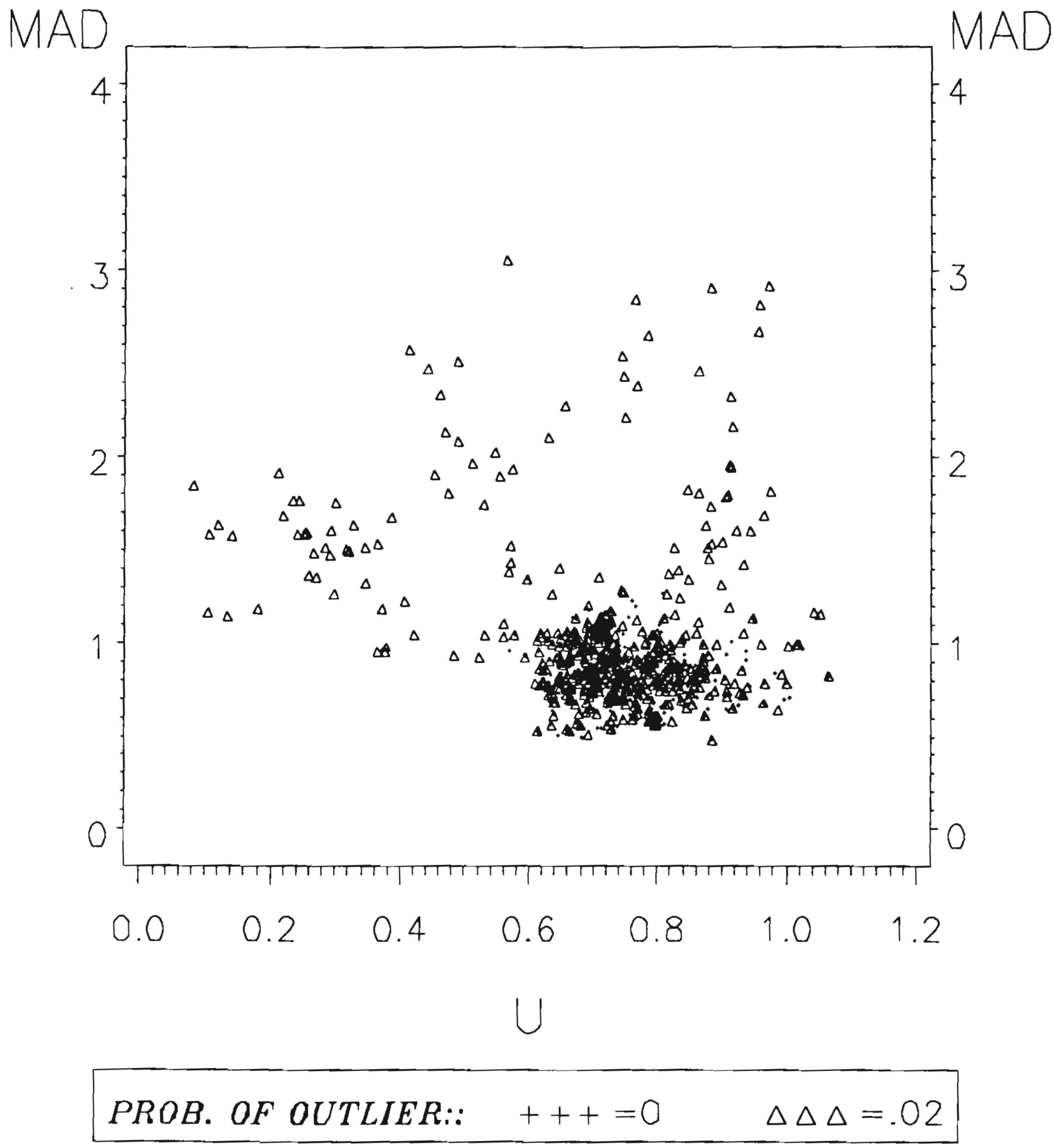

Figure 1Ø. Scatterplot Illustrating the Effect of a .Ø2 Probability of Outliers on the Relationship Between MAD and Theil's U. Note that while the outliers tend to increase MAD, they tend to decrease Theil's U.

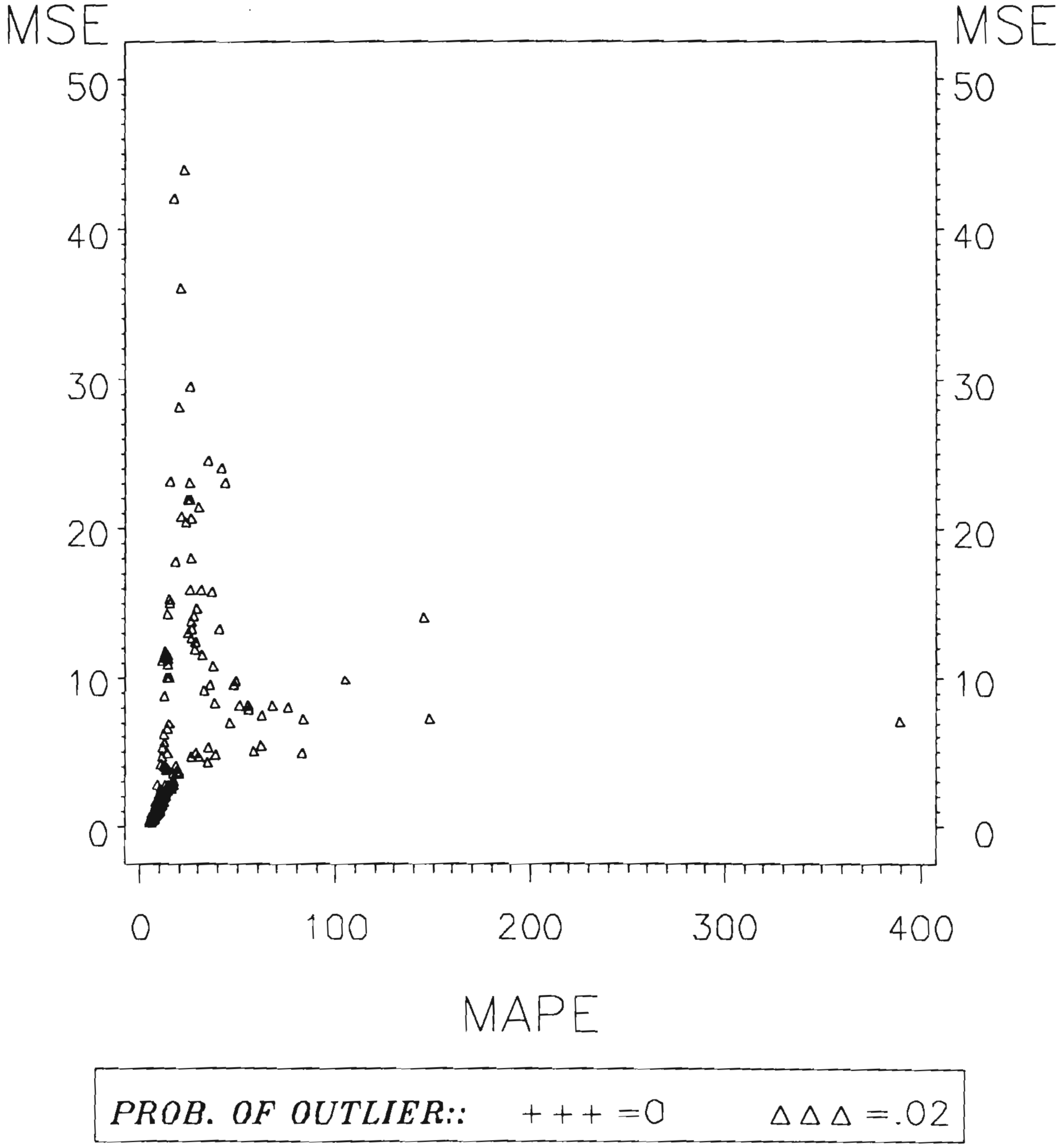

Figure 11. Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MSE and MAPE. Note that you can have a large MSE with a small MAPE, or a large MAPE with a small MSE.

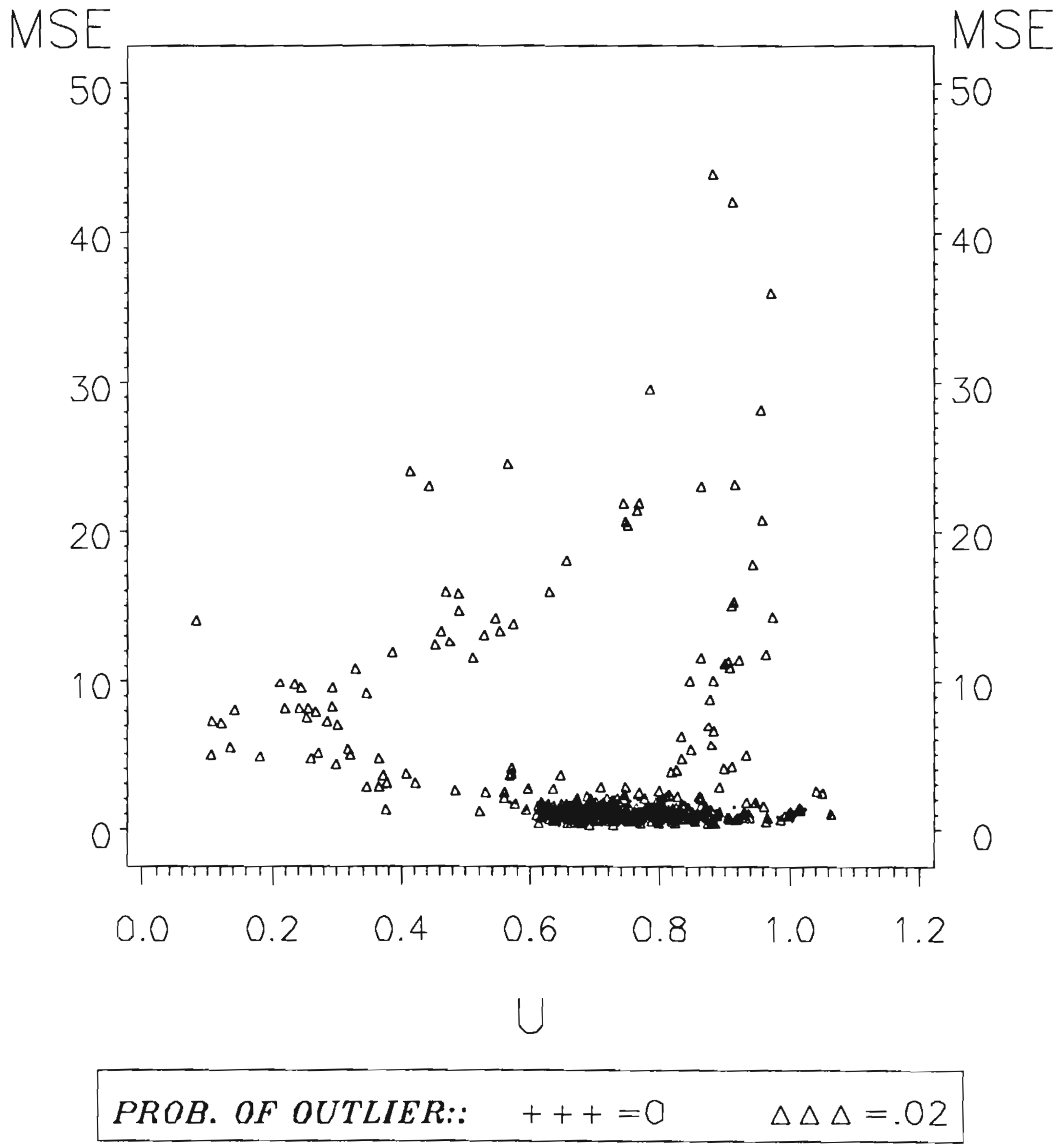

Figure 12. Scatterplot Illustrating the Effect of a .Ø2 Probability of Outliers on the Relationship Between MSE and Theil's U. MSE values tend to increase while Theil's U values tend to decrease, but there seems to be no pattern to this relationship.

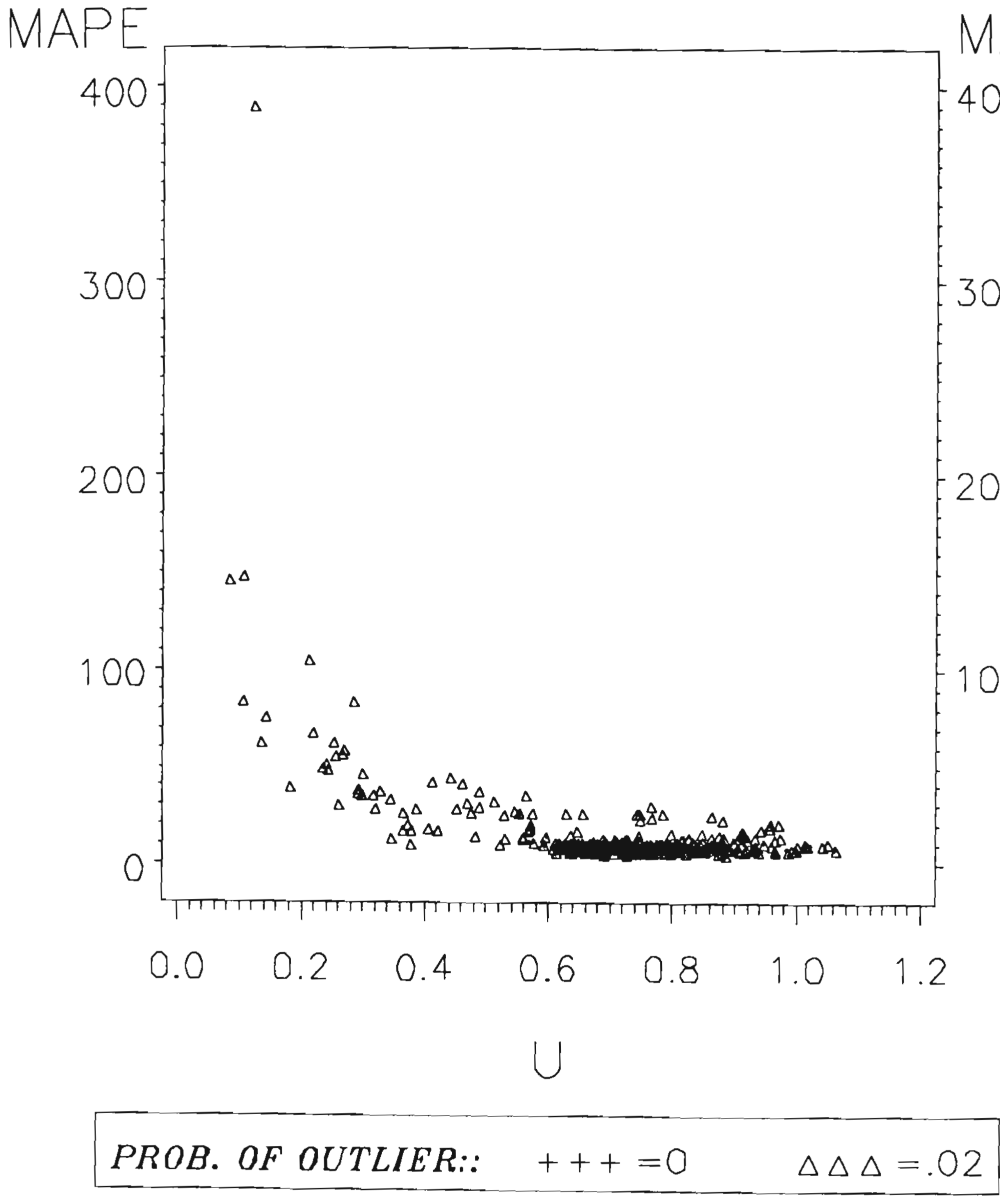

Figure 13. Scatterplot Illustrating the Effect of a .02 Probability of Outliers on the Relationship Between MAPE and Theil's U. With outliers, the Theil's U values clearly drop as the MAPE values tend to increase.

4.5 Summary and Conclusions

The information in the previous sections indicate how the error measures were selected for use in this study. Four measures were selected based upon a thorough literature review of types of quantitative error measures. A simulation study indicates that the three measures of MAD, MSE, and MAPE are reasonably well related so long as we have a well behaved series with a small coefficient of variation. Increased dispersion or outliers do not radically affect the relationships. These relationships are not deemed strong enough to exclude MAD, MSE, MAPE, or Theil's U from the study. It is acknowledged that the use of these error measures will complicate the research to some extent, but their use seems prudent at least through the simulation study in Chapter VI. Further discussion of the validity of raw error measures in the empirical test is included in Section 6.2. The use of these four measures should encourage those researchers and practitioners who have selected a favored measure to evaluate the results. It should increase the generality of the conclusions.

CHAPTER V

PARAMETRIC AND NONPARAMETRIC TECHNIQUES STUDIED

5.1 Techniques for Stationary Time Series

Techniques for stationary time series do not deal with trend or seasonality. Only the level and randomness need to be considered. While all time series techniques could be applied in this case, those with parameters designed to cope with trend and seasonality would be overspecified (have insignificant parameters) and would violate Box and Jenkins (1976) ground rules which call for parsimony in time series models. Simple parametric techniques designed for stationary series include Naive One, single exponential smoothing, single moving average, and adaptive response rate exponential smoothing. While adaptive response rate exponential smoothing has some nice theoretical features, it is dropped from further consideration due to its increased complexity and reputation for poorer performance than single exponential smoothing (Hibon, 1984). The literature review did not indicate any nonparametric techniques for extrapolative forecasting. As a consequence, the parametric techniques are explained in the next section, while nonparametric alternatives are introduced in Section 5.1.2.

5.1.1 Parametric Techniques

The techniques of Naive One, single moving average, and single exponential smoothing are simple and yet perform reasonably well. The Naive One technique presumes that tomorrow will be like today, or that:

$$F_{t+1} = X_t \quad . \tag{5.1}$$

McLaughlin (1975) considered this no-change model to be a standard for measuring a forecasting technique. If the forecaster or technique does no better than a no effort Naive One technique, then the extra effort is for nought. Besides its value for comparative purposes, Naive One is used in the development and explanation of both McLaughlin's Standardized Realization Percent (or batting average) and Theil's U statistic. McLaughlin (1975) notes that this technique has surprisingly good properties around turning points, i.e. points in time when the underlying trend or level change. Naive One can be considered the most responsive, and least stable technique for stationary series. While this technique is not seriously considered as a competitor, it is used as a standard for comparison in many tables.

The single moving average technique uses the average of the last m time series values as the forecast, so that:

$$MA1_t = (X_t + X_{t-1} + \ldots + X_{t-m+1}) / m \ , \tag{5.2}$$

represents the average and:

$$F_{t+1} = MA1_t \tag{5.3}$$

represents the forecast for time period t+1. The integer m can vary from 1 upward, so that the single one period moving average is identical to the Naive One technique. Smaller m values result in a more responsive, less stable model. Larger m values result in a less responsive, more stable model. While Hibon (1984) dismisses this technique as generally no better than Naive One or exponential smoothing, it should be noted that the results of the Makridakis and Hibon (1979) study indicated that single moving average yielded a MAPE of 12.9 and an adjusted Theil's U of 1.61, while single exponential smoothing yielded values of 12.8 and 1.66 on their 111 series. Note also that the M-competition ran all automatic techniques on all 1001 series. For those series with trend, the fitting procedures may have allowed a more responsive model for exponential smoothing than for moving averages, thus perhaps unjustly penalizing the moving average's performance. Appendix 2 of the M-competition results (Makridakis, et. al., 1984) does not indicate what restrictions were used on either the number of periods in the moving average or on the exponential average smoothing parameter. Any restrictions could have affected the results. For all 1001 series, moving average forecasts exceeded those of exponential smoothing by 34% using MAPE and 216% using MSE, a rather sharp contrast to the previous results. The empirical portion of this study attempts to screen the M-competition series into more appropriate categories prior to fitting and forecasting. The single moving average is expected to be more competitive when used on series that do not display monotonic trend.

The single exponential smoothing technique also uses the average of the historical data as the forecast. But in this case, all the historical data are used and weighted exponentially with more weight on recent time periods' data and less weight on old data. The equation for the exponential average can be expressed as:

$$EA1_t = \underline{c}X_t + (1-\underline{c})EA1_{t-1} \quad , \tag{5.4}$$

where $EA1_t$ refers to the single exponential average at time t and $\underline{c}$ refers to the smoothing constant or parameter of the model. The smoothing constant can vary between 0 and 1 but Brown (1963), one of the originators of exponential smoothing, suggests that the use of constants above .3 usually result from trying to use the technique for inappropriate series. The smaller the smoothing constant, the less weight is given to the most recent data points, and the more stable the model. The recursive equation 5.4 leads to the question of how to start the forecasts. Possible ways to begin include letting:

$$EA1_1 = X_1 \quad , \tag{5.5}$$

or

$$EA1_1 = \Sigma\ (X_t)/n \quad , \tag{5.6}$$

where the summation is over the set of data used for fitting. Due to the rapid decrease in weight applied to the first value, its effect on forecast accuracy is minimal so long as the series of data used for fitting is of sufficient length. For that reason the simulation study used the first series value as $EA1_1$.

Although a very simple technique, single exponential smoothing did well in the Makridakis and Hibon (1979) study. On seasonally adjusted data, its MAPE was within 7.5% of the best reported method, while Box-Jenkins was high by 22.5%. In the M-competition (Makridakis, et. al., 1982) this technique had the lowest average MAPE for one period ahead forecasts, and the lowest average MAPE over horizons 1-18, among the standard techniques tested over all 1001 series and over the 111 series (Table 2(a) and Table 2(b), respectively, of the Makridakis article). Since the results in Chapter IV indicated that MAPE and MSE are not necessarily highly related, it should not be surprising that in terms of MSE, simple exponential smoothing did not perform as well. Of course, there remains the question of whether the average of raw measures of error across series of different levels and dispersion is meaningful. Guerts (1983) indicated that while MSE is an excellent accuracy measure for a single time series, its average over several series can be distorted by the magnitudes of the series.

5.1.2 Nonparametric Alternatives

While no nonparametric extrapolative techniques were found in the literature review, some nonparametric techniques were found that could be applied to or modified for the extrapolative forecasting problem. One promising nonparametric technique is the running median as discussed by Tukey (1977). This technique is used as a smoothing device in data analysis, not in extrapolative forecasting, but its application appears straight forward. Due to the general strength of exponential smoothing techniques, a second technique, a variation on running medians similar to an exponential average, is proposed and evaluated. A third nonparametric technique with promise for this stationary case is Walsh Averages.

Similar techniques not included in the present study include trimmed means and medial averages. Levenbach and Cleary (1984) suggested using the trimmed mean as a robust measure of location. The trimmed mean is developed by deleting a specified proportion of the ordered observations from the top and from the bottom. That is, a percentage of the largest and smallest values are deleted, then the mean is calculated from the data remaining. Wheelwright and Makridakis (1985) used a medial average when smoothing historical data for calculation of seasonals. These are similar to the trimmed mean, but only the single largest and single smallest values are deleted. Neither the trimmed mean, nor the medial average are strictly speaking nonparametric techniques; although they do address the problem of outliers.

Running medians are directly analogous to moving averages. While moving averages consist of the arithmetic average (mean) of m consecutive time series values, running medians as discussed by Tukey (1977) consist of the median of m consecutive time series values. Thus,

$$RM1_t = MED(X_t , X_{t-1} , \ldots ,X_{t-m+1}) , \quad (5.7)$$

and

$$F_{t+1} = RM1_t \quad . \quad (5.8)$$

The running median is expected to be very robust to outliers at the cost of being somewhat slower to respond to changes in level.

A second nonparametric technique is a modification to the running median which incorporates some of the features of exponential smoothing. The idea is to reduce the expected variability of forecasts over time from running medians by inclusion of an exponential weighting to these medians. The smoothed median can be expressed as:

$$SM1_t = \underline{c}RM1_t + (1-\underline{c})SM1_{t-1} , \quad (5.9)$$

where;

$SM1_t$ is the single smoothed median at time t,

$RM1_t$ is the running median for time t,

$\underline{c}$ is the smoothing constant, with a value ranging from zero to one.

This is again a recursive equation and is started by setting $SM1_1 = RM1_1$. Since the items smoothed are medians rather than series values, it is suspected that relatively few time periods should be used in each median and that a relatively large smoothing constant (compared to exponential smoothing) would be better.

The third nonparametric technique for stationary series is the median of the Walsh Averages as discussed by Hollander and Wolfe (1973). Given a set of m consecutive values, X_t, X_{t-1}, ..., X_{t-m+1} , from the series, the m(m-1)/2 Walsh Averages can be expressed as:

$$WA_t = [(X_i + X_j)/2 \ , \ i < j], \qquad (5.10)$$

where i and j are chosen from [t,t-1,...,t-m+1]. This is perhaps easier to see from a small example. Suppose the median of the Walsh Averages of 1, 2, and 5 was desired. A table can be arranged as in Figure 2.

values		1	2	5
values	1	2	3	6
	2	-	4	7
	5	-	-	10

Figure 14. Development of a Walsh Average.

where the values from the series are in the upper and the left column margins. The 3 by 3 body consists of the pairwise sums of upper row and left column elements. Only the upper diagonal matrix is considered when calculating the Walsh Averages. While the average of each matrix element value could be performed at any time, it is more efficient to perform the division just once, after the median has been determined. The median of the six elements in this upper diagonal matrix is the average of the third and fourth ranked elements or 5. The forecast is then 5/2 or 2.5. If the data had been symmetric the median (2 in this case) would have equaled the median Walsh Average. Because the data was "skewed right", the median Walsh Average was larger than the median. The median Walsh Average is more responsive than the median, yet still relatively insensitive to outliers.

5.2 Techniques for Series with Linear Trend

Simple techniques for linear trend must deal with level and trend. All those considered use a linear forecast. A level, a_t, is estimated for time period t. A slope or expected change in level per time period, b_t, also is estimated. Given these two estimates, forecasts for future time periods are easily computed. Again, many time series techniques can be applied to nonstationary series. While nonlinear trends are certainly possible, as are changes in the underlying series behavior, the techniques used here are designed for series with linear trend. The parametric techniques used were double moving average and

Brown's Linear Exponential Smoothing. As in the stationary case, no nonparametric extrapolative forecasting techniques were found. Rather, a robust regression technique and modifications to the previous section's nonparametric techniques are considered.

5.2.1 Parametric Techniques

Two parametric techniques for linear trend time series were considered. These were the double moving average and Brown's linear exponential smoothing. The single moving average was discussed in Section 5.1.1 where the basic equation was expressed as:

$$MA1_t = (X_t + X_{t-1} + \ldots + X_{t-m+1})/m \,. \quad (5.11)$$

The double moving average, $MA2_t$, can be expressed as:

$$MA2_t = (MA1_t + MA1_{t-1} + \ldots + MA1_{t-m+1})/m \,, \quad (5.12)$$

where m is consistently used as the number of periods in the average. It should be clear that in the case of a perfect linear trend MA1 will lag the proper time series value by (m-1)/2 time periods, while MA2 will lag the proper value by another (m-1)/2 periods. The proper level at time t, a_t, can then be expressed as:

$$a_t = MA1_t + (MA1_t - MA2_t) \quad (5.13)$$

$$\text{or} \quad = 2MA1_t - MA2_t \,. \quad (5.14)$$

To forecast a linear trend, a slope for time period t also is needed. Since the difference between $MA1_t$ and $MA2_t$ is (m-1)/2 time periods of change, the slope can be estimated as:

$$b_t = [2/(m-1)]\ [MA1_t - MA2_t] \ . \qquad (5.15)$$

The double moving average forecast for any future time period t+k can then be expressed as:

$$F_{t+k} = a_t + b_t k \ . \qquad (5.16)$$

For the purposes of the simulation study, k is fixed at one, while the empirical study considers k = 1,2,...,6.

Double moving averages did not perform well in the Makridakis and Hibon (1979) study, yet it was only three tenths of a percentage point worse than Box-Jenkins for one period ahead forecasts (MAPE of 15.0 versus 14.7). The technique apparently was not used in the M-competition. While this technique is not expected to be the best performer in the present study, it is included because it is quite similar to the proposed double running medians technique. Since this study does not run the linear trend techniques on clearly inappropriate series, this technique is expected to perform better than in the previous studies.

The second and last parametric technique used for the nonstationary series is Brown's linear exponential smoothing. The single exponential average is expressed in Equation 5.4 of Section 5.1.1 as:

$$EA1_t = \underline{c}X_t + (1-\underline{c})EA1_{t-1} \ . \qquad (5.17)$$

The double exponential average at time t is:

$$EA2_t = \underline{c}EA1_t + (1-\underline{c})EA2_{t-1} \ , \qquad (5.18)$$

where the smoothing constant $\underline{c}$ remains a fixed value for both averages. Forecasting using Brown's Linear

Exponential Smoothing technique again requires an estimated level at time t and an estimated slope. Here:

$$a_t = 2EA1_t - EA2_t \,, \tag{5.19}$$

$$b_t = [\underline{c}/(1-\underline{c})][EA1_t - EA2_t]. \tag{5.20}$$

Brown (1963) explained the development of these formulas. As before, the forecast for future time period t+k is:

$$F_{t+k} = a_t + b_t k, \tag{5.21}$$

while k is fixed at one for the simulation study.

Brown's linear exponential smoothing did rather well in the Makridakis and Hibon (1979) study and in the M-competition (Makridakis, et. al., 1982). In terms of MAPE, Brown's one parameter method tied Holt's two parameter method for one period forecasts and was slightly better than Holt's for two period ahead forecasts over all 1001 series. Gardner (1983) indicates that Brown's technique should be preferred over Holt's on theoretical grounds.

5.2.2 Nonparametric Alternatives.

The first nonparametric technique considered is a modification of robust regression as discussed by Hollander and Wolfe (1973). In the regression problem, their technique determines the proper slope as the median of the set of slopes generated from each pair of points. Their intercept term is determined by running lines with this slope through all the data points, generating a set of possible intercept values. The median of this set of possible values is the intercept for the regression equation:

$$Y = a + bX . \qquad (5.22)$$

Modifying this technique for the extrapolative forecasting problem required similar computations for each time period advance, using the last m time periods of data for each computation. The modified technique uses the same approach for the development of a forecasting slope, then runs a forecasting line through the point (median of t, median of X).

A second proposed approach is the double running median. Given a perfect linear trend the formulas that gave forecasts for double moving averages also will work for medians:

$$RM1_t = \text{median}(X_t, X_{t-1}, \ldots , X_{t-m+1}) , \qquad (5.23)$$

$$RM2_t = \text{median}(RM1_t, RM1_{t-1}, \ldots, RM1_{t-m+1}), \qquad (5.24)$$

$$a_t = 2RM1_t - RM2_t , \qquad (5.25)$$

$$b_t = [2/(m-1)][RM1_t - RM2_t] . \qquad (5.26)$$

A third possible approach is a double smoothed median. Since in a perfect linear case the smoothed median is always (m-1)/2 time periods behind X_t, the inclusion of this multiplier results in the following equations:

$$SM1_t = \underline{c}RM_t + (1-\underline{c})SM_{t-1} , \tag{5.27}$$

$$SM2_t = \underline{c}SM1_t + (1-\underline{c})SM2_{t-1} , \tag{5.28}$$

$$a_t = 2SM1_t - SM2_t , \tag{5.29}$$

$$b_t = [\underline{c}/(1-\underline{c})][SM1_t - SM2_t] , \tag{5.30}$$

$$F_{t+1} = a_t + [(m+1)/2]\ b_t . \tag{5.31}$$

Conover's (1980) robust regression technique, regression on ranks and interpolation, was considered but it is not included in the present study because forecasting and interpolation are largely opposite concepts. Modifications similar to the trimmed means or medial averages techniques also are not pursued for this study since, strictly speaking, they are not nonparametric.

5.3 Summary and Conclusions

Simple parametric techniques have been identified that have performed well in recent forecasting competitions. Nonparametric techniques that also can be applied to similar time series extrapolation problems have been identified. Chapter VI compares the performance of these techniques under a variety of simulated conditions and on monotonic trend and no monotonic trend subsets of the M-competition data.

CHAPTER VI

EVALUATION OF PERFORMANCE

Performance of the parametric and nonparametric techniques discussed in Chapter V are compared on simulated and actual time series. The simulation study deals with both stationary and linear trend series with results presented in Section 6.1. Sample program listings are included in Appendix B, while detailed supporting tables are included in Appendix C. An empirical study is then performed using a subset of 111 deseasonalized time series from the M-competition with discussion and results presented in Section 6.2. Sample program listings for the empirical study are included in Appendix D.

6.1 Simulation Study

The purpose of the simulation is to try to identify strengths and weaknesses, advantages and disadvantages among and between the nonparametric and the parametric techniques as the type of series and error distribution vary. The series chosen for the simulation are all autoregressive of order one. In the stationary case, a series of this type can be expressed as:

$$X_t = \Phi * X_{t-1} + U + e_t , \quad (6.1)$$

where X_t is the value of the series at time t,

Φ (Phi) is the autoregressive coefficient,

U is a constant, and

e_t is an error term from a chosen distribution.

In the linear trend case the equation for the series value can be expressed as:

$$X_t = \Phi * X_{t-1} + V + b * t + e_t , \quad (6.2)$$

where V is a constant, and

b is a slope.

For each case in this simulation, the combination of Φ, U, V, and b is adjusted to maintain an expected series value of 20. When normal errors are used, the expected coefficient of variation is 15%. While the coefficient of variation is not defined when using Cauchy errors, the Cauchy scale parameter is adjusted so that the 25th and 75th quartiles are equal for the series of normal and Cauchy errors.

According to Makridakis, Wheelwright, and McGee (1983), the value of the autoregressive coefficient is restricted to be between -1 and +1. Series with coefficients ranging from -.9 to +.9 were generated, plotted, and evaluated for use. Series with negative coefficients ossilate about their mean. A large value is followed by a small value, with the effect damping out at a speed determined by the size of the coefficient. This type of behavior is considered unlikely to be encountered in the business world. As a result, three positive coefficient values are chosen for study. These are $\Phi = .3$, $\Phi = .5$, and $\Phi = .7$. The series generated with these coefficients display very similar behavior as evidenced by Figures 15 and 16, but yet are sufficiently different to shift the fitting requirements from the most stable parameters to the most responsive parameters. They are felt to be sufficiently different to study the various techniques.

Two hundred replications are made for each combination of autoregressive coefficient and error term distribution, half for the stationary case and half for the linear case. Sixty observations are drawn for each replication with each technique restarted for each replication. The twenty observations after each start are disregarded for two reasons. One is to assure independence of the replications, the second is to assure that the method used to initiate some of the techniques does not taint the results. Pictorially, one replication can be illustrated as:

```
|        Forecasts          |        Forty Observations        |
|<-- Calculated, But --->|<---- Used for Error Measure -->|
|    No Error Measures      |           Calculations           |
|___________________________|__________________________________|

1                          21                                 60

                         Observations
```

Figure 15. Pictoral of One Replication

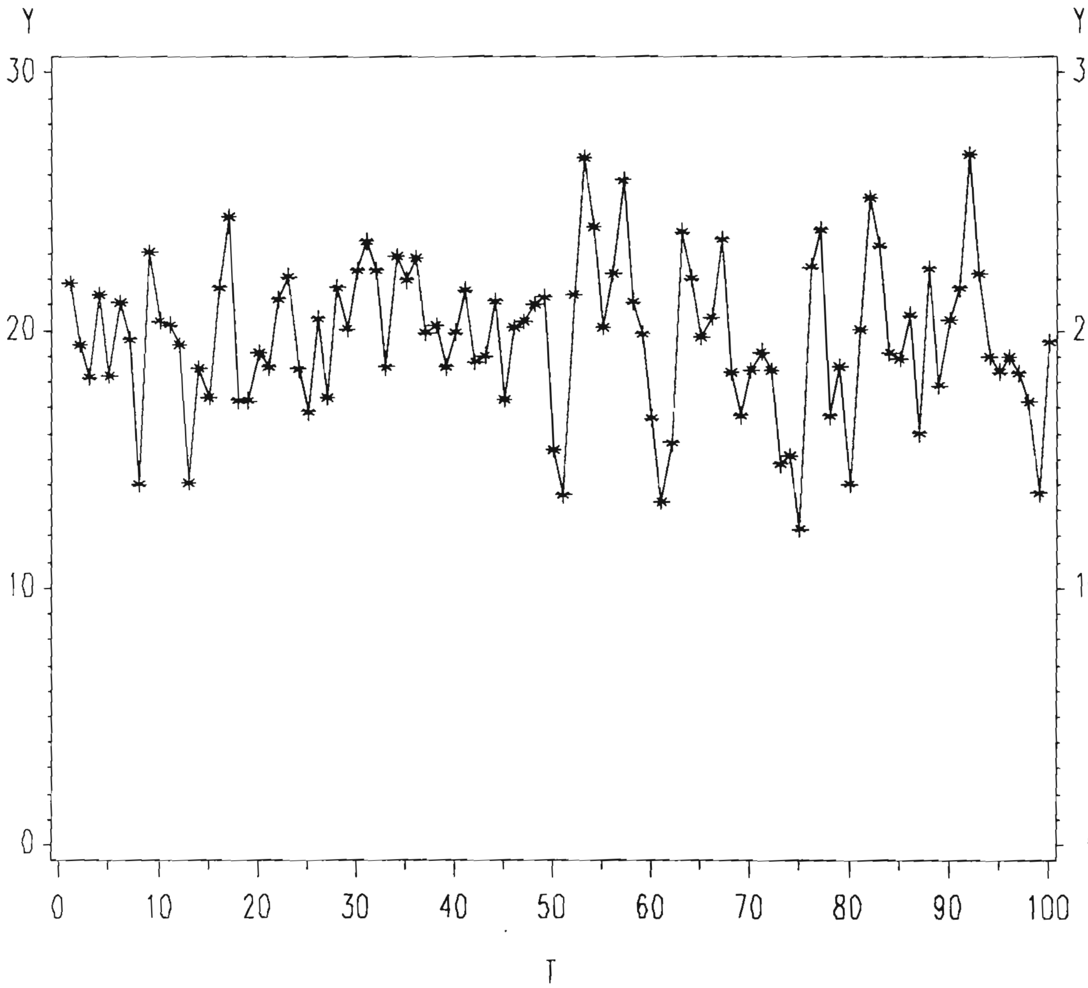

Figure 16. Graph of a Stationary Series with Phi = .3

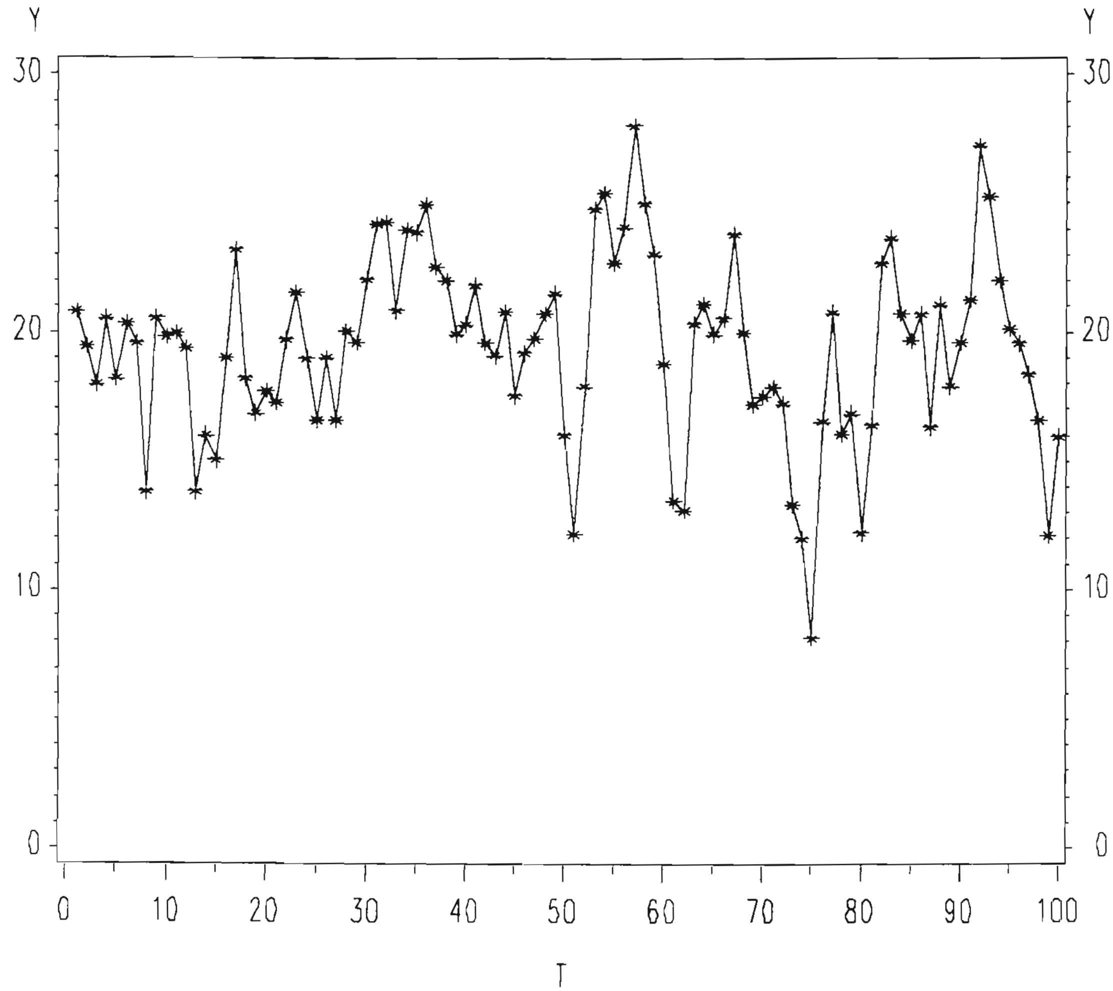

Figure 17. Graph of a Stationary Series with Phi = .7. Note the similarity to Figure 16, with the primary difference being the larger maximum and smaller minimum values.

All forecasts are for the next time period. For each technique used in the simulation and for each replication, the error measures chosen in Chapter IV are calculated. Over the 100 replications the maximum, upper quartile, mean, median, lower quartile, and minimum error measure values are calculated. Thus, each error measure represents results for 4000 observations.

One should avoid the flaw of numeracy cited by Pack (1983). Thus, the detailed results from the simulation study are expressed as ratios in Tables 23 - 46 of Appendix C. The ratios represent the actual error measure value for the cell, to the error measure that resulted from use of the Naive One technique on the series. Tables 9 - 14 within this chapter deal with comparisons of the best parametric and best nonparametric techniques and are felt to be integral to the chapter. Including the detailed tables within the chapter was felt to distract from the presentation. The detailed tables thus are included in Appendix C.

6.1.1 Performance on Stationary Series

The extrapolative techniques of single moving average, single exponential smoothing, running median, smoothed median, and Walsh average are run for the six stationary cases (Φ = .3, .5, .7; normal and Cauchy errors). A grid of parameter values is established so that performance can be related to parameters used in the models. For example, the smoothed median uses smoothing constants of .2, .3, .4, .5, and .6 on medians of length 3, 4, 5, and 6. Performance results are then calculated for each of the 24 combinations of parameters. Box plots in Figures 18 and 19 illustrate the MAD performance of different length moving averages and running medians on a stationary series with normal errors. Figures 20 and 21 illustrate the performance of moving averages and running medians on a stationary series with Cauchy errors. Note that in the case of Cauchy errors, the mean error measure is typically near the upper quartile value. For this reason, both mean and median error measure results are considered when distributions have Cauchy errors.

6.1.1.1 Normal Errors

If results of previous nonparametric work can be generalized into the time series arena, the parametric techniques should on average perform better than comparable nonparametric techniques on well behaved stationary series with normal error terms. This expectation is realized for each of the error measures considered. The parametric technique of exponential smoothing is always superior to all other techniques when the errors are normally distributed. It should be mentioned that although

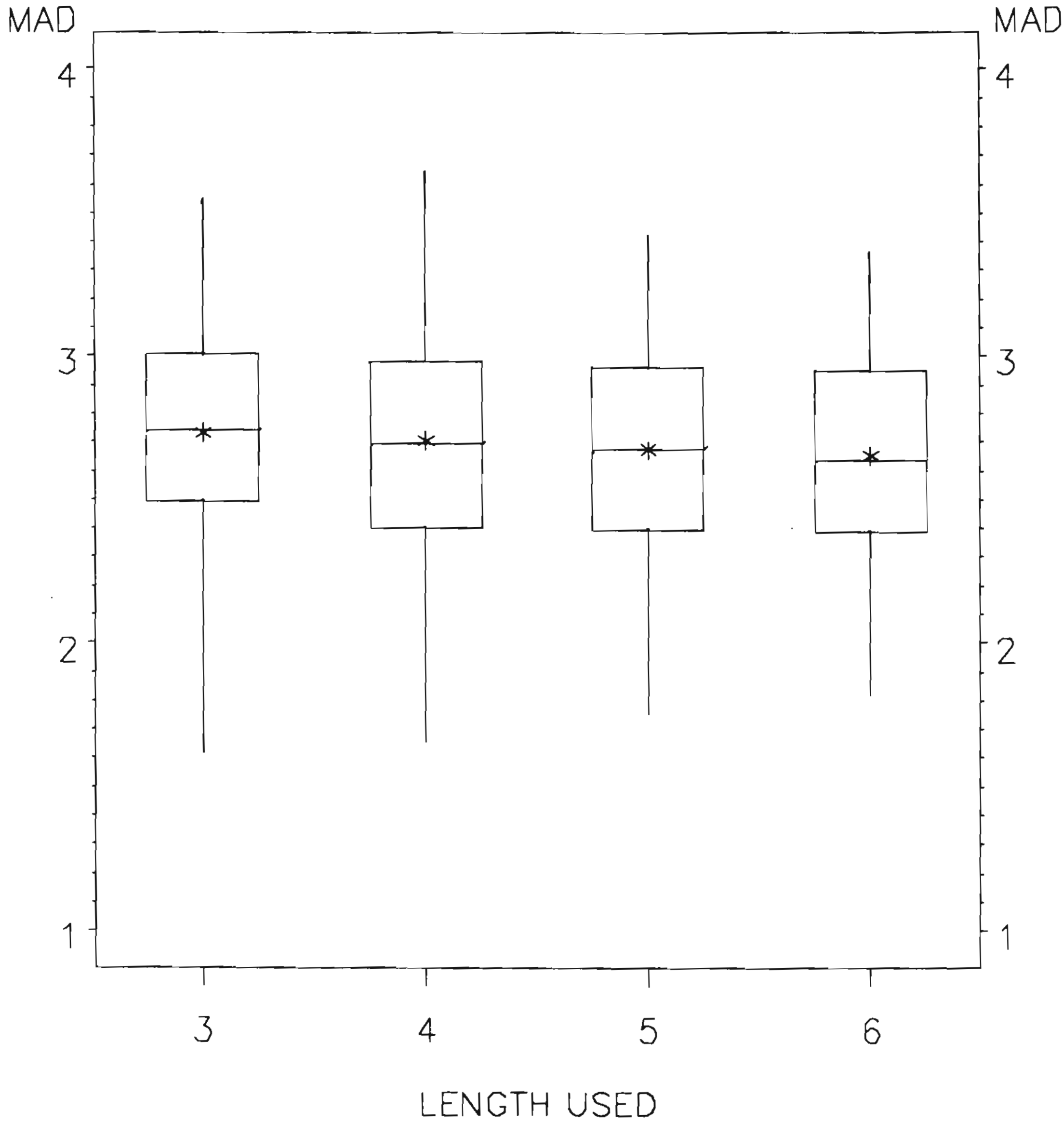

Figure 18. Sample Box Plots of Simulation MAD Results using Moving Averages on a Series with Normal Errors, and Phi = .3

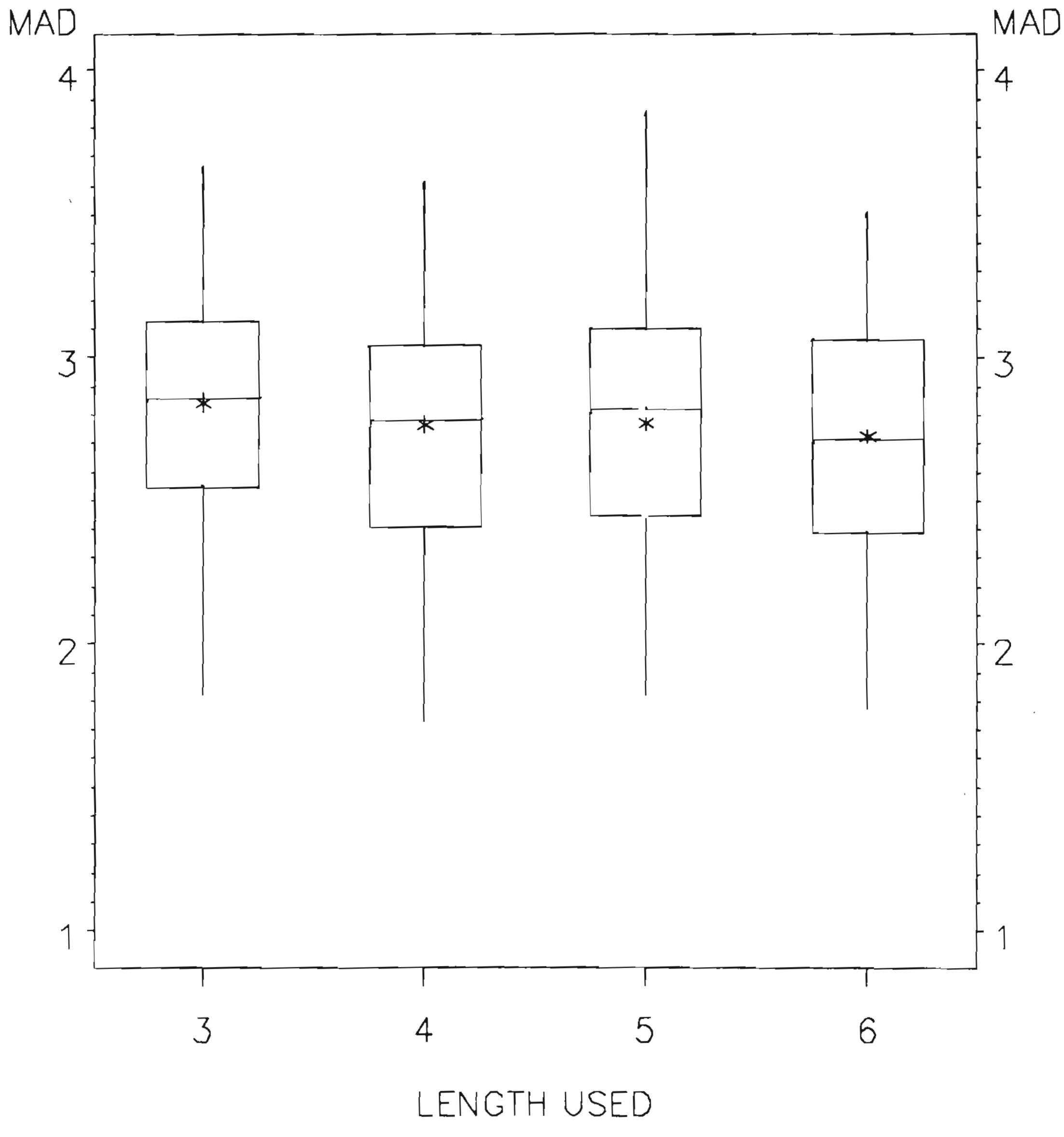

Figure 19. Sample Box Plots of Simulation MAD Results using Running Medians on a Series with Normal Errors, and Phi = .3

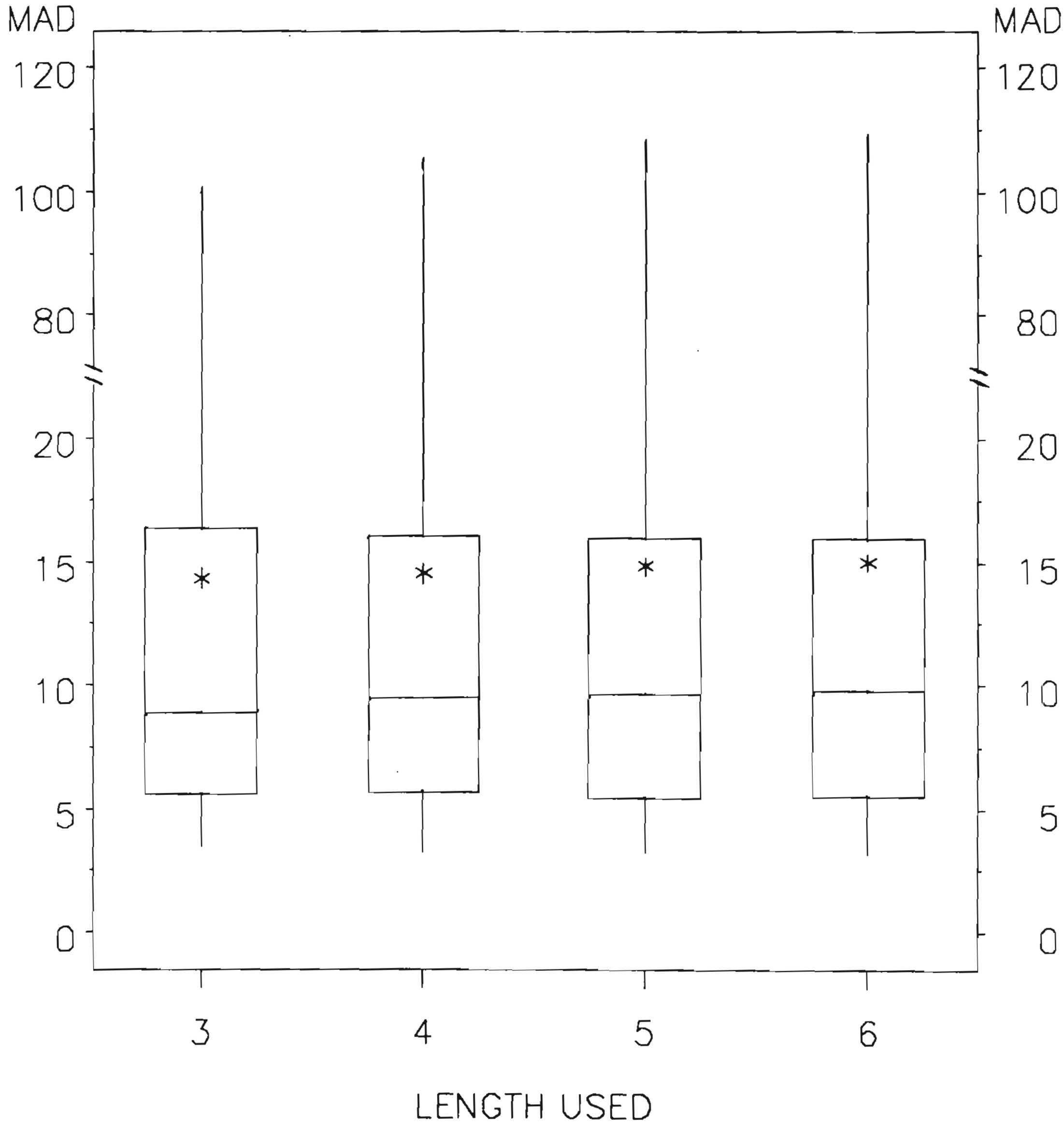

Figure 2Ø. Sample Box Plots of Simulation MAD Results using Moving Averages on a Series with Cauchy Errors, and Phi = .3

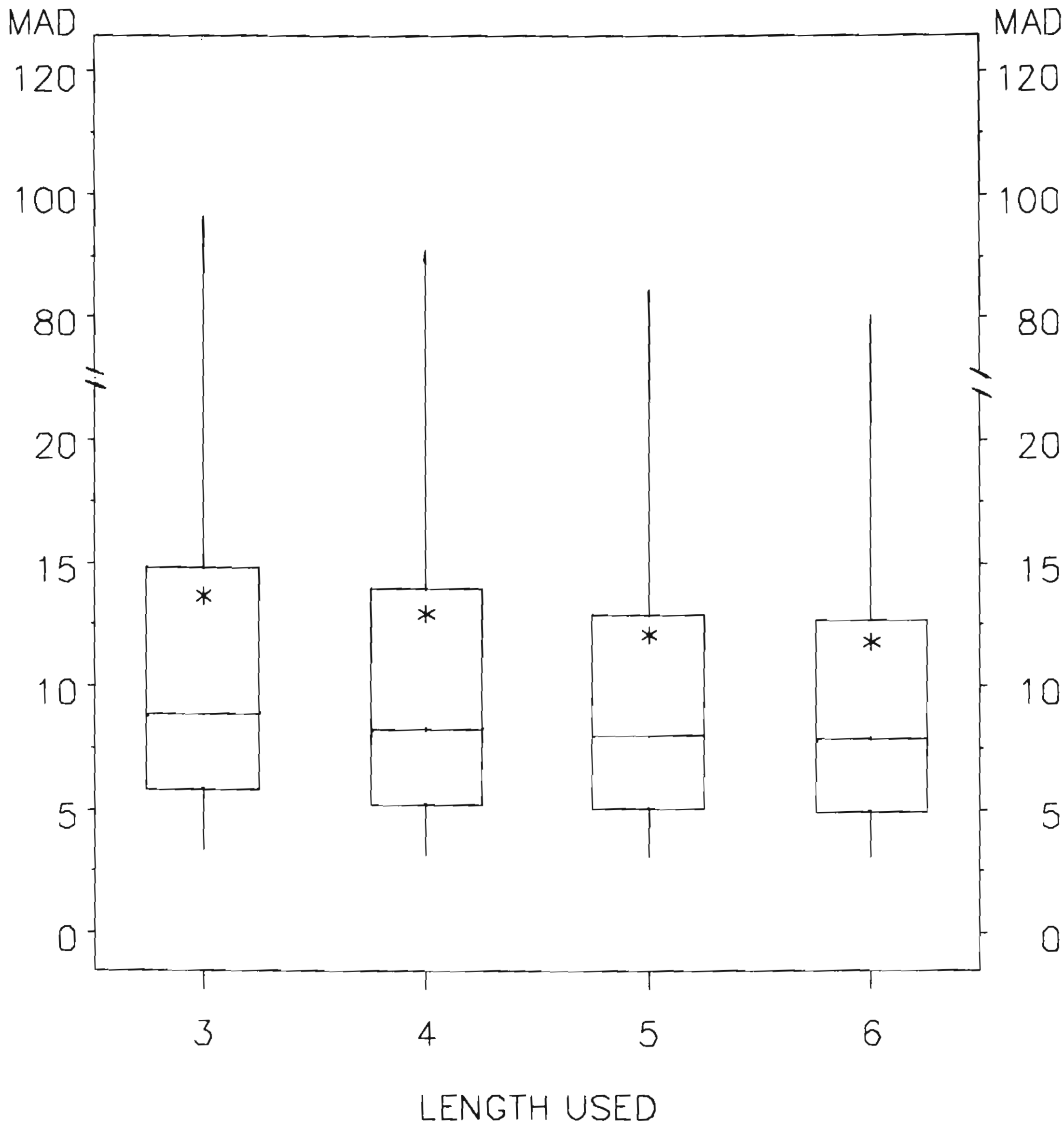

Figure 21. Sample Box Plots of Simulation MAD Results using Running Medians on a Series with Cauchy Errors, and Phi = .3

exponential smoothing is uniformly superior, the difference between the best exponential smoothing result and the best nonparametric result is not necessarily large, as can be seen in Table 9. One conclusion to be drawn from Table 9 is that use of nonparametric techniques on these stationary time series would result in errors that are 5% to 20% larger than necessary, depending upon the error measure chosen.

Normal Errors, Φ=.3

More detailed analysis of Tables 23 - 26 (of Appendix C) indicates that when Φ = .3, the best forecasting method leans toward stability. That is smaller smoothing constants and longer length averages or medians have lower fitting values. Additional results (not tabled) indicated that for this case, use of the population mean as the forecast resulted in a lower error measure value (MAD, MSE, MAPE, or Theil's U) than did the use of any other technique considered, including Naive One. Of course, one problem with proposing the use of the population mean as a forecasting method is that the mean is unknown in most forecasting situations. A second problem is that even if the true mean were known (or an excellent estimate could be made), its use for forecasting would be ill advised. Lawrence, Edmundson, and O'Connor (1986) noted many cases of the failure of the assumption of constancy in a subset of the M-competition data. Thus the forecaster should use a technique that will adjust the forecasts automatically if the mean experienced a minor or major change. For this reason, the simulation study restricted the average and median lengths and the values used for the smoothing constants. It is clear that a better fit exists outside the grid range, but it was felt that the grids chosen represented the lengths and smoothing constants most likely to be utilized by forecasters. Reference to the $\Phi = .3$ portions of Tables 23 to 26 indicates that in general

Table 9

Best Nonparametric Results
versus Best Parametric Results:
Stationary Series, Normal Errors
Average Results over 100 Simulated Series

Phi	Error Measure	Ratio*	Best Nonparametric	Best Parametric
.3	MAD	1.05	2.62	2.50
	MSE	1.10	10.74	9.77
	MAPE	1.05	13.69	13.09
	Theil's U	1.05	.89	.84
.5	MAD	1.08	2.79	2.58
	MSE	1.18	12.28	10.44
	MAPE	1.08	14.57	13.48
	Theil's U	1.08	1.01	.93
.7	MAD	1.09	2.85	2.61
	MSE	1.21	12.86	10.65
	MAPE	1.09	15.09	13.83
	Theil's U	1.10	1.09	.99

* Ratio is the value derived by dividing the best average nonparametric result by the best average parametric result. When Ratio is more than one then the best nonparametric technique performed worse than the best parametric.

the more stable parameter values yield lower error measures. Reference to graphs such as those in Figures 19 and 20 result in the same conclusion.

To some extent, the magnitude of the ratio values in the Φ = .3 portion of the Tables 23 - 26 results from the rather poor job done by Naive One on series with this autoregressive coefficient. Numbers are rounded to two decimal places. The best nonparametric results were always obtained from the smoothed median technique with the smallest smoothing constant used for this technique. The use of longer length medians (5 or 6 period versus 3 or 4 period) had surprisingly little effect as long as the smoothing constant was kept at .2. Moving average did not perform as well as exponential smoothing. In the case where Φ = .3 (stability desired), the moving average of length nine (included in the study, but not tabled) was not as good as exponential smoothing with a .1 smoothing constant.

<u>Normal Errors, Φ=.5</u>

When Φ = .5 the use of the mean as the forecast, or the use of Naive One results in essentially the same error measure values. In this case, the only technique that improves upon Naive One is exponential smoothing (see Tables 23 - 26). While the exponential smoothing results lean toward a model that is more responsive, the smoothed medians technique achieves its best results with a small smoothing constant yet the shortest length medians. Here the best performing nonparametric method is the Walsh average. It maintains much of the median's resistance to outliers (not tested in the simulation), yet it is more responsive.

Normal Errors, Φ=.7

With normally distributed errors and with Φ = .7, the Naive One technique outperforms all others tested (see Tables 23 - 26). The most responsive exponential smoothing model used was still worse than the simple Naive One. By looking at the Φ = .7 portions of Tables 23 - 26, the techniques can be ranked by responsiveness or stability. Consider the column with n = 3. The moving average is most responsive, followed closely by the Walsh average, then by the running median technique. The need for responsiveness even has the smoothed median results properly ordered with lower error measures for shorter period medians and for larger smoothing constants.

6.1.1.2 Cauchy Errors

The Cauchy distribution is usually represented as a long tailed distribution that looks somewhat like a normal distribution. It is often used as an alternative distribution when parametric and nonparametric measures and tests are evaluated, see Hollander and Wolfe (1973). Parzen (1960), indicates that the Cauchy distribution can be expressed as:

$$f(x) = 1/(\pi * \beta * (1+((x-\alpha)/\beta)^2)) \qquad (6.3)$$

where alpha is the location parameter, and beta is the scale parameter. Johnson and Kotz (1970) indicate that while the Cauchy distribution does not have a finite expected mean or standard deviation, the location and scale parameters may be regarded as being analogous to the mean and standard deviation. For this study, the errors are constrained to be zero on average, and the location parameter is thus set to zero. The SAS Users Guide: Basics, Version 5 (1985) specifies that their RANCAU function can be used to generate a Cauchy variate as:

x = ALPHA + BETA*RANCAU(seed) (6.4)

where ALPHA is the location parameter, BETA is the scale parameter, and seed is a positive real number used to intitialize the random number string. A scale parameter of slightly less than two was required to equate the Cauchy deviates' midrange to that of the normal. A comparison of the best average results of parametric and nonparametric techniques are presented in Table 10, while best median results are in Table 11.

Cauchy Errors, Φ=.3

When Φ = .3, average results (Tables 27 - 30) and median results (Tables 31 - 34) favor the nonparametric techniques. Average improvements range from 1% for MSE to 13% for MAD, while median improvements range from less than 1% for MAPE to almost 16% for MSE. The best nonparametric results are from smoothed medians for MAD, MSE, and Theil's U; while running medians barely surpasses smoothed medians

Table 10

Best Nonparametric Results
versus Best Parametric Results:
Stationary Series, Cauchy Errors
Average Results over 100 Simulated Series

Phi	Error Measure	Ratio*	Best Nonparametric	Best Parametric
.3	MAD	.89	11.64	13.17
	MSE	.99	3144.09	3167.67
	MAPE	.90	72.10	80.19
	Theil's U	.89	.83	.93
.5	MAD	1.16	15.74	13.56
	MSE	1.16	3928.30	3382.24
	MAPE	.70	123.96	177.47
	Theil's U	1.05	1.19	1.13
.7	MAD	1.18	17.54	14.88
	MSE	1.23	4279.36	3479.19
	MAPE	1.14	183.77	161.82
	Theil's U	1.15	1.37	1.19

* Ratio is the value derived by dividing the best average nonparametric result by the best average parametric result. If Ratio is less than one then the best nonparametric technique performed better than the best parametric technique.

for MAPE. The best median results are from running medians for MSE, MAPE, and Theil's U; while smoothed medians are better for MAD, and very close for MSE.

Cauchy Errors, Φ=.5

When Φ = .5 the results are mixed. Average MAD, MSE, and Theil's U all indicate that the parametric technique of exponential smoothing is favored (see Tables 27, 28, and 30). Average MAPE, however, indicates that the nonparametric technique of running medians is preferable (see Table 29). This seems inexplicable. If medians rather than means are considered, all measures indicate that exponential smoothing is the best technique (see Tables 31 - 34).

Since the MAPE results are unexpected, some further analysis was performed on the initial data. For the original data, the exponential smoothing MAPE had a maximum value of 63% when the smoothing constant was .10, versus a maximum MAPE of 16% for the best length running medians. While changes in average results were usually quite regular with changes in the smoothing constant, for this instance the average MAPE is very sensitive to the choice of parameter. Simulating results from these two techniques on a new data set resulted in a lower average MAPE from the exponential average than from smoothed medians. This raises the question of the stability or reliability of the

Table 11

Best Nonparametric Results
versus Best Parametric Results:
Stationary Series, Cauchy Errors
Median Results over 100 Simulated Series

Phi	Error Measure	Ratio*	Best Nonparametric	Best Parametric
.3	MAD	.92	7.82	8.53
	MSE	.86	250.10	289.49
	MAPE	.99	53.46	53.81
	Theil's U	.98	.86	.88
.5	MAD	1.09	9.53	8.78
	MSE	1.16	347.17	300.90
	MAPE	1.16	73.28	63.33
	Theil's U	1.03	1.02	.99
.7	MAD	1.16	11.06	9.54
	MSE	1.26	401.55	318.42
	MAPE	1.23	91.24	73.92
	Theil's U	1.10	1.18	1.08

* Ratio is the value derived by dividing the best median nonparametric result by the best median parametric result. If Ratio is less than one then the best nonparametric technique performed better than the best parametric technique.

use of MAPE since none of the other error measures under consideration reflected any problem. Perhaps the combination of MAPE and Cauchy errors drove this peculiar result. While this anomaly is intriguing, it is not seen as critical to this research and is not pursued further at this point.

Review of Tables 27 - 34 indicates that the dominant nonparametric technique is running medians. It had the lowest median MAD, MSE, MAPE, and Theil's U value, and also the lowest average MAD and Theil's U, although Walsh average and even smoothed median had lower average MSE values. It is interesting to note that for Φ = .3 and .5, the error measures generally indicate that running medians would be superior to moving averages (although worse than exponential smoothing) when errors are Cauchy. Since single moving averages did not do badly in the Makridakis and Hibon (1979) study, hope remains that running medians may perform well in the empirical portion of this study.

Cauchy Errors, Φ=.7

When Φ = .7, the results favor the parametric techniques. Since the first order autocorrelation here is .7, the value of an observation can be reasonably well forecast by the Naive One technique. None of the other techniques can match the Naive One results. The best nonparametric technique is Walsh Averages based upon all average measures as well as median MAD and median MAPE. Running medians is best based upon median MSE and median Theil's U.

6.1.1.3 Conclusions from Stationary Simulations

The best parametric technique is exponential smoothing. For the smoothing constants and lengths considered in this study, it is consistently a better performer than moving averages for Φ= .3, .5, or .7. If Brown's (1963) suggested limit of .3 is imposed on the exponential smoothing constant (that is, rows corresponding to smoothing constants higher than .3 are ignored in Tables 23 - 34), then the difference in results between the two parametric techniques is not large for cases with normal errors. The difference is at most 6% for MSE and about 3% for the other error measures used. When errors are from a Cauchy distribution and the smoothing constant is limited to .3, the moving average yields marginally better average results for MAD and Theil's U; yet somewhat worse for MSE and MAPE. Exponential smoothing is superior otherwise for average and median results. This is consistent with earlier reported work that favors exponential smoothing over moving averages.

The best nonparametric technique varies by type of series. With Φ = .3 and normally distributed errors, the smoothed median technique results in the lowest error measures, followed closely by Walsh average and then running median. When Φ = .3 and errors are from a Cauchy distribution, smoothed medians remains superior based upon average errors, followed closely by running medians. In this case Walsh averages perform considerably worse, probably because the technique is inherently more responsive than the medians. Based upon median error measure results, the running median is best for MSE, MAPE, and Theil's U; losing only to smoothed medians for MAD.

When the series has Φ = .5 and normal errors, the best nonparametric technique is Walsh averages, followed by smoothed medians, and then running medians. When the errors are Cauchy, then the running median technique is better based upon MAD, MAPE, and Theil's U. Average MSE favors Walsh averages, but the median error measures all favor the running median.

The best nonparametric technique for series with the highest autocorrelation (Φ = .7) is Walsh averages. Running medians is better only for median MSE and median Theil's U. The need for responsiveness placed smoothed medians at a marked disadvantage versus the more responsive Walsh average and running median.

It appears that stationary series with high levels of autocorrelation will not be well forecast by nonparametric techniques. The question remains as to whether high levels of autocorrelation are common. Previous good performance by exponential smoothing versus Box-Jenkins seems to imply that such levels of autocorrelation are not all that common, since exponential smoothing itself is not designed for such series while Box-Jenkins is so designed.

Brown (1963) indicated that when the best fit exponential smoothing model has a smoothing constant greater than .3, the series probably is not appropriate for use of the technique. Discounting the obvious case where a stationary technique is being used for a nonstationary series (for example, a series with linear trend), this raises a separate research question. Brown merely indicates the symptom and does not indicate how much autocorrelation is too much. This simulation indicates that for these types of series, a first order autocorrelation of .3 gives series that are appropriate for exponential smoothing. First order autocorrelation of .5 or higher results in series that are inappropriate. That is, best fitting results in a model with a smoothing constant that is too large. While it might be interesting to determine the autocorrelation levels at which exponential smoothing becomes inappropriate, this is not seen as central to the nonparametric forecasting question and is thus deferred.

6.1.2 Performance on Linear Trend Series

The extrapolative techniques of double moving average, linear exponential smoothing, double running median, double smoothed median, and robust regression are run for six cases (Φ = .3, .5, .7; normal and Cauchy errors). The same grid approach of parameter values is used for the trend as was used on the stationary simulations.

6.1.2.1 Normal Errors

When errors are drawn from a normal distribution, results again always favor the parametric technique of exponential smoothing. Table 12 compares the best parametric and nonparametric results achieved. Detailed results are presented in Tables 35 - 38. For exponential smoothing and Φ = .5, MAD and MSE are minimized at a smoothing parameter value of .1, MAPE is minimized at a value of .3, and Theil's U is minimized at .2. If Φ = .7, MAD and MSE are minimized at smoothing constant values of .4, MAPE at a value of .5, and Theil's U at a value of .3.

When Φ = .3 or .5 the smoothed median yields the best nonparametric results, with the choice of a smoothing constant having a more pronounced effect than the length of the medians being smoothed. When Φ = .7, the robust regression technique becomes much more competitive. Robust regression is quite competitive with double moving averages in terms of MAD, MSE, and MAPE.

6.1.2.2 Cauchy Errors

Simulation results when the error around the trend line was Cauchy are in Tables 39 - 46. As with the stationary case, the nonparametric technique based on smoothed medians is superior to all other techniques when Φ = .3. When the first order autocorrelation rises to .5 and above, average error measures favor the more responsive parametric technique. Comparison of best average parametric and nonparametric results is displayed in Table 13 and for median results in Table 14. It is somewhat interesting to note that the double smoothed median seems

Table 12

Best Nonparametric Results
versus Best Parametric Results:
Trend Series, Normal Errors
Average Results over 100 Simulated Series

Phi	Error Measure	Ratio*	Best Nonparametric	Best Parametric
.3	MAD	1.05	2.69	2.57
	MSE	1.10	11.34	10.34
	MAPE	1.05	15.92	15.19
	Theil's U	1.06	.87	.82
.5	MAD	1.08	2.93	2.72
	MSE	1.16	13.51	11.63
	MAPE	1.10	180.35	163.93
	Theil's U	1.10	1.02	.93
.7	MAD	1.20	3.35	2.80
	MSE	1.45	17.86	12.35
	MAPE	1.19	200.41	168.07
	Theil's U	1.21	1.23	1.02

* Ratio is the value derived by dividing the best nonparametric result by the best parametric result. A Ratio value above one indicates that the best nonparametric method did not perform as well as the best parametric method.

to be more competitive with linear exponential smoothing than the smoothed median is with single exponential smoothing. When median error measures are considered, double smoothed medians is the dominant nonparametric technique for Φ = .3 and .5, while robust regression is dominant when Φ = .7.

Table 13

Best Nonparametric Results
versus Best Parametric Results:
Trend Series, Cauchy Errors
Average Results over 100 Simulated Series

Phi	Error Measure	Ratio*	Best Nonparametric	Best Parametric
.3	MAD	.81	12.06	14.81
	MSE	.94	3159.09	3365.98
	MAPE	.79	95.93	120.71
	Theil's U	.82	.93	1.14
.5	MAD	1.03	16.97	16.41
	MSE	1.06	4033.95	3807.84
	MAPE	1.04	107.88	103.88
	Theil's U	1.11	1.29	1.16
.7	MAD	1.06	17.58	16.61
	MSE	1.09	4353.85	3998.69
	MAPE	1.04	126.20	121.62
	Theil's U	1.14	1.36	1.19

* Ratio is the value derived by dividing the best nonparametric result by the best parametric result. A Ratio value less than one indicates that the best nonparametric technique is better than the best parametric technique.

When Φ = .3, reference to Tables 39 - 42 indicates that the minimum exponential smoothing mean MAD, MSE, and MAPE is achieved with $\underline{c}$= .1; yet the minimum mean Theil's U is achieved with $\underline{c}$= .4 (the minimum median Theil's U is with $\underline{c}$= .1). For Φ = .7 the average MAPE and average Theil's U are minimized with $\underline{c}$ = .5 while average MAD and average MSE are minimized with $\underline{c}$ = .4. This, along with the similar results when errors are normal underscore that one should not expect a technique fit to minimize one error measure to also minimize the other error measures. Further analysis of the detailed results in the Appendix C tables is left to the interested reader.

Table 14

Best Nonparametric Results
versus Best Parametric Results:
Trend Series, Cauchy Errors
Median Results over 100 Simulated Series

Phi	Error Measure	Ratio*	Best Nonparametric	Best Parametric
.3	MAD	.88	8.18	9.25
	MSE	.82	256.51	312.06
	MAPE	.91	56.04	61.39
	Theil's U	.93	.89	.96
.5	MAD	1.03	10.68	10.41
	MSE	1.04	357.36	345.25
	MAPE	1.07	70.67	66.09
	Theil's U	.99	1.02	1.02
.7	MAD	1.09	11.48	10.56
	MSE	1.15	413.84	361.03
	MAPE	1.05	75.52	72.04
	Theil's U	1.01	1.07	1.06

* Ratio is the value derived by dividing the best nonparametric result by the best parametric result. A Ratio value less than one indicates that the best nonparametric technique is better than the best parametric technique.

6.1.2.3 Conclusions of Linear Trend Simulation

The best parametric technique is linear exponential smoothing. It provides better forecast statistics in almost every case even if restricted to a smoothing constant of .3. Double moving averages performs worse relative to linear exponential smoothing than the analogous techniques in the stationary case.

The best nonparametric technique again depends upon the combination of series type and error measure. With normally distributed errors, all average error measures favor the double smoothed median technique, except for MAPE when Φ = .7. For this one combination, robust regression provides the best result. When errors are from a Cauchy distribution, optimal results are dependent on Φ value.

When Φ = .3 and errors are Cauchy, the double smoothed median is the clear choice, while double running median comes in second and robust regression last. When Φ = .5 robust regression outperforms double running medians, although double smoothed medians generally still provides the best results. When Φ = .7 robust regression becomes the dominant nonparametric technique.

If the exponential smoothing constant is restricted to .3 for Brown's linear exponential smoothing, it is interesting to note that robust regression is quite competitive with the best parametric technique for Φ = .7 and Cauchy errors. With average measures (Tables 39 - 42), they tie based upon the MAD, while exponential smoothing is preferable based on MSE, and robust regression wins on MAPE and Theil's U. With median measures (Table 43 - 46), exponential smoothing is better for the raw measures of MSE and MAD, while robust regression is better for the relative measures of MAPE and Theil's U.

6.1.3 Simulation Conclusions

Results indicate that nonparametric forecasting techniques show promise for series with nonnormal errors and lower levels of autocorrelation. Series with normally distributed errors, or with high levels of autocorrelation seem to be better forecasted using the parametric methods.

As an example of the value of nonparametric techniques when the autocorrelation is low, consider the trend series, with $\Phi = .3$. If double smoothed medians (the best nonparametric technique based upon mean MAD from Table 39) is used instead of double exponential smoothing (the best parametric technique) on a series with normally distributed errors, then the resulting average MAD is about 5% higher than necessary (see Table 12). If on the other hand, the series has errors from a Cauchy distribution, then the smoothed median average MAD is some 22% better than that achieved by exponential smoothing (see Table 13). The 5% applies to the rather small MAD of 2.94 when the errors are normal, while the larger percentage applies to a MAD of 14.77 when the errors are Cauchy.

The simulation study has allowed an evaluation of the proposed nonparametric techniques on generated time series under a few limited conditions. This provides greater insight for when the use of nonparametric extrapolative techniques might be worthwhile. The empirical study which is discussed next further examines their usefulness.

6.2 Empirical Study

The time series used are the 111 series subset of the M-competition data used by Makridakis, et. al. (1982). When received from Dr. Makridakis, these were loaded directly onto a VAX 11/785 super minicomputer, and transformed into a SAS data set. (The empirical study used VMS SAS Releases 5.03 and 5.16). Each of the series consists of a number (NFIT) of values used for fitting, a number (NOF) of additional hold-out values used to evaluate the techniques' forecasting performance, seasonal factors, and some descriptive codes. NOF ranges from 6 to 18 depending upon the type of series. One important descriptive code is the series identification number. The series are numbered from 1 to 1001. The series identification number (SID) uniquely identifies each series. The 111 series subset starts with the series with SID = 4, and every seventh series thereafter. These 111 series were used in the M-competition to evaluate the performance of the more complex, time intensive techniques. Before testing the techniques used in this study on these series, two adjustments were made. First, the series are deseasonalized using the seasonal factors included with the data, then the series are separated into two groups through use of a Rank-Spearman test.

The data are deseasonalized since none of the techniques being tested are designed to cope with the additional complexities of seasonality. Makridakis and Hibon (1979) and Makridakis, et. al. (1982) found that deseasonalizing time series allowed simpler techniques to be very competitive with those more complicated approaches designed to handle seasonality. Use of the deseasonalized series for fitting and forecasting is deemed an adequate and appropriate test of the techniques.

The M-competition applied techniques indiscriminately to all series despite the fact that some of the techniques, such as moving average, and exponential smoothing, will be biased in cases of linear or monotonic trend series. For example, they will, on average, forecast low if there is an increasing trend. Use of a stationary technique for a monotonic trend series is clearly inappropriate, while use of a linear trend technique on a stationary series violates the principle of parsimony and very likely would drive larger forecast errors, particularly for longer horizons. Some division of the series into those more appropriate for particular classes of techniques is therefore necessary. This division is made based upon a Rank-Spearman test for monotonic trend on the first NFIT points. Basically, this test indicates that if the absolute value of the correlation between the ranks of the values and their associated time period is sufficiently large, then we likely have monotonic trend which presumably continues into the NOF values. If this correlation is low, then we conclude that we do not have monotonic trend. This group of series without monotonic trend would consist of stationary series, series with a nonlinear trend that display a changing pattern (increasing to decreasing, or decreasing to increasing) near the middle of the series, and perhaps series where the variation in the data overpowers the trend. This division of the series into two groups provides more logical series for testing of the stationary and linear trend techniques used in this study. At a 90% confidence level, ninety-two of the series test to have monotonic trend, while nineteen do not.

6.2.1 Stationary Techniques

The stationary techniques of moving average, exponential smoothing, running median, smoothed median, and Walsh average are tested on the subset of nineteen series. All techniques were programmed in the SAS data step. Starting values for moving average, running median, and Walsh average are clear from Section 5.1. Exponential smoothing, however, requires a starting exponential average which can affect the fitting measures. Since observations are limited in an empirical study, the first twenty observations cannot be discarded to eliminate initialization effects. A good initialization procedure must be used. The procedure chosen was suggested by Montgomery and Johnson (1976) where the mean of the first six values is used to initialize exponential smoothing. The smoothed median technique is initialized by using the first running median of the appropriate length as the initial value. Where possible, the program results were checked versus results of other software. The moving average program gave identical results, while the exponential smoothing results only differed due to variations in starting procedures. Fitting was done using a grid search technique over NFIT observations. Grid width was set at .05 for all smoothing parameters, medians up to length six, and moving averages up to length 10 were considered. (For one series where NFIT is equal to twelve, the moving average length is restricted to six periods to assure an adequate number of points for fitting evaluation. More extensive discussion is deferred until Section 6.2.2 where the problem is much more prevalent.) Four different error measures are used with the five different techniques, resulting in a possibility of twenty different forecasting models for each series. Forecasts are made from time period NFIT for times NFIT+1 through NFIT+6. Thus, horizons ranged from one to six periods. Results are presented by error measure.

6.2.1.1 Results Using MAD

The results using MAD are displayed in Table 15. They indicate that the smoothed median technique not only has a lower fitted MAD, but also is most accurate on average over the nineteen series for horizons of one, four, and five periods. The lowest average MAD over all six horizons is achieved by the moving average technique, followed closely by Walsh average, and smoothed median. Since the MAD is a raw error measure, a quick look at the magnitudes of the fitted MAD values is worthwhile. The largest fitted MAD for a series was over 8000, while the smallest was less

Table 15

Empirical Results
using Mean Absolute Deviation:
Stationary Techniques'
Average MAD Values over 19 Series

		Forecast Horizon						Avg
Technique	Fitting	1	2	3	4	5	6	1-6
Moving Average	693	1038	1375	672	1378	525	664	939
Exponential Smoothing	714	1094	1468	597	1433	602	574	961
Running Median	696	984	1190	834	1325	510	850	949
Smoothed Median	690	952	1236	860	1292	504	823	944
Walsh Average	695	1042	1411	631	1382	567	629	943

Tabled values are the average MAD for the indicated technique over the 19 series that did not test to have monotonic trend.

than one. The performance of a technique on the series with the largest values thus carries much larger weight than the technique's performance on series with smaller values.

6.2.1.2 Results Using MSE

The results when using MSE are similar to those for the MAD. The results in Table 16 indicate that smoothed median again was most accurate over the fitted data and for horizons 1, 2, and 4. In this case the running median technique had the best average MSE over the six horizons, followed closely by the smoothed median technique. Here,

Table 16

Empirical Results
using Mean Squared Error:
Stationary Techniques'
Average MSE Values (Millions) over 19 Series

Technique	Fitting	Forecast Horizon 1	2	3	4	5	6	Avg 1-6
Moving Average	6.24	8.62	14.37	2.01	19.52	1.20	2.12	7.97
Exponential Smoothing	6.36	9.67	16.84	1.85	22.06	1.72	1.47	8.94
Running Median	6.05	6.53	9.36	2.90	14.14	1.02	4.58	6.42
Smoothed Median	5.98	6.49	9.21	2.68	14.14	1.24	4.80	6.43
Walsh Average	6.15	7.75	14.25	2.49	18.56	1.02	2.09	7.69

Tabled values are the average MSE for the indicated technique over the 19 series that did not test to have monotonic trend.

magnitudes ranged from around 100 million for one series to almost zero for another, so the techniques that fit the larger series well are especially favored.

6.2.1.3 Results Using MAPE

The MAPE results displayed in Table 17 indicate relative small differences in the mean absolute percentage error between these techniques. The parametric techniques were slightly better for fitting. The best technique over

Table 17

Empirical Results
using Mean Absolute Percentage Error:
Stationary Techniques
Average MAPE Values over 19 Series

		Forecast Horizon						Avg
Technique	Fitting	1	2	3	4	5	6	1-6
Moving Average	11.1	9.4	16.6	19.8	20.9	15.5	20.6	17.1
Exponential Smoothing	11.1	9.9	16.3	19.4	19.9	14.6	20.3	16.7
Running Median	11.4	8.6	15.0	19.9	21.4	15.7	21.3	17.0
Smoothed Median	11.6	9.8	15.0	20.2	20.8	15.0	20.4	16.9
Walsh Average	11.5	9.6	16.0	19.0	21.0	15.7	20.0	16.9

Tabled values are the average MAPE for the indicated technique over the 19 series that did not test to have monotonic trend.

all six horizons was exponential smoothing, with a 1.2% advantage over both smoothed median and Walsh average.

6.2.1.4 Results Using Theil's U

Results for use of the Theil's U measure are more difficult to portray. The extreme skewness of the measure causes the average Theil's U per horizon to be large. Due to this problem, Makridakis and Hibon (1979) limited their Theil's U values to a maximum of two. In this study, an alternative approach was used. That is, median values were calculated and used in Table 18 for each horizon. Even so, this is not highly informative since the medians in almost every case are one. A simple interpretation of these

Table 18

Empirical Results
using Theil's U:
Stationary Techniques
Theil's U Values over 19 Series

		Forecast Horizon						
Technique	Fitting	1	2	3	4	5	6	1-6
Moving Average	.807	1.00	1.00	1.00	1.00	1.00	1.00	1.017
Exponential Smoothing	.812	1.00	1.00	1.00	1.00	1.00	1.00	1.022
Running Median	.818	1.00	1.00	1.00	1.00	1.00	1.00	1.049
Smoothed Median	.866	1.50	1.01	.95	1.08	1.10	1.11	1.046
Walsh Average	.816	1.00	1.00	1.00	1.00	1.00	1.00	1.035

The fitting column is the average fitted Theil's U value over the 19 series that did not test to have monotonic trend. The numbers for horizons 1 through 6 are medians. The 1-6 column in this table is not the average of the median Theil's U values for horizons one through six. Rather, the numerator and denominator terms for the Theil's U measure were summed over these six horizons, then the Theil's U computation was performed.

results is that none of the techniques improved upon use of a simple naive estimate that future values would be the same as the values at time period NFIT. Perhaps the last column is the most informative in this table. Summing the numerator and denominator components of the Theil's U measure over several observations seems to reduce it's variability enough to provide useful information. In this case, summing the components over the six forecasts for the six different horizons results in the reported values. Since only a rare forecaster would be content to use a naive forecast, this column indicates that the parametric techniques would have been close to the naive results.

6.2.1.5 Stationary Techniques Conclusion

The proportion of the series that failed to display monotonic trend indicates that opportunities to use these stationary techniques are somewhat limited. The nonparametric techniques of running median and smoothed median are very competitive on these series when raw error measures were used. When relative error measures were used, the parametric techniques appear marginally superior. The next section deals with the application of linear trend techniques on the considerably larger set of monotonic trend series.

6.2.2 Linear Trend Techniques

The linear trend techniques of double moving average, linear exponential smoothing, double running median, double smoothed median, and robust regression are tested on the subset of 92 series that did test to have monotonic trend. All techniques are programmed in the SAS data step. Sample programs appear in Appendix D. Essential elements in these programs include the initialization of the smoothing techniques and control of averaging lengths.

When dealing with long series, the initialization of exponential smoothing may not be critical. If the smoothing constant is subjectively selected by the analyst, then the initialization procedure has little effect. Forecasts are only slightly affected because the exponential smoothing process rapidly reduces the effect of the initialization. If the smoothing constant is best fit, then the initialization can have a significant effect, especially on shorter series. The selection of a smoothing constant could be driven by rather large fitting errors on the first few data points, errors that were in fact more a function of the poor initialization procedure than the smoothing constant. The series in this sample ranged in length from 9 to 126 fitted values, with 14 series ranging in length from 10 to 20 fitted values. The selection of a good initialization procedure is therefore necessary.

The procedure selected was recommended by Montgomery and Johnson (1976). This procedure uses a regression line through the first six data points to initialize the single exponential average and double exponential average. The procedure applies directly for linear exponential smoothing, while a slight modification is necessary for initialization of the double smoothed median technique. An alternative initialization procedure, not selected, uses regression on all NFIT points to start the exponential averaging process. A comparison of the fitting errors using each procedure indicated the clear superiority of the chosen procedure. Although the details are not included in this report, the use of regression results on all fitting data for initialization did not work well. It resulted in large fitting errors for the first few observations when the series displayed a nonlinear monotonic trend.

The double moving average and double running median programs require control statements to keep the length of the average or median within reason given a particular series size. This problem was briefly discussed in Section 6.2.1, although only one series was affected. In this case, a much larger number of series is affected. For example, a ten period double moving average does not generate an error term until time period 20. Recalling the lengths of the series in the sample, the longer length averages or medians are not always possible. A further constraint is that when a best coefficient is being selected using a grid search, it is better to compare the average of several error terms. A series of length (NFIT) 22 would result in only three errors contributing to the fitting error measure for the ten period double moving average, while the three period double moving average would be evaluated on sixteen errors. To avoid this disparity and the concurrent increased risk of a poor choice of length, control statements were inserted into the program to keep the considered lengths short enough to assure that approximately NFIT/2 errors were available for the longest considered length.

Forecasts are again best fit using the grid search technique. Forecasts are made from time period NFIT for horizons of length 1 through 6. Results are considered by the type of error measure used for fitting and evaluation, first for the raw error measures of MAD and MSE, then for the relative error measures of MAPE and Theil's U.

6.2.2.1 Results Using MAD

The results when using the Mean Absolute Deviation are displayed in Table 19. They indicate that the exponential smoothing approaches of linear exponential smoothing and double smoothed median fit better. In contrast, the best forecasts for most horizons and overall were from double moving averages. Parametric techniques thus yielded lower MAD values for fitting and forecasting.

Table 19

Empirical Results
using Mean Absolute Deviation:
Linear Trend Techniques
Average MAD Values over 92 Series

		Forecast Horizon						Avg
Technique	Fitting	1	2	3	4	5	6	1-6
Double Moving Average	1070	947	1130	2296	3084	3517	3704	2446
Double Exponential Smoothing	847	1313	1699	2886	3645	4141	3547	2872
Double Running Median	1168	1641	2270	3728	4643	5346	4848	3745
Double Smoothed Median	846	1351	1869	3197	4173	4674	4208	3870
Robust Regression	1123	1864	2314	3497	4579	4957	4419	3605

Tabled values are the average MAD for the indicated technique over the 92 monotonic trend series.

6.2.2.2 Results Using MSE

The results using MSE are shown in Table 20. In this case the trend technique of linear exponential smoothing yielded both the best fit and the best forecast over all horizons. While double smoothed medians matched the fitting ability, its forecasts were not as accurate using this error criteria. It is well to keep in mind the cautions from various researchers that average raw error measures over series that vary widely in magnitude are of questionable value.

Table 20

Empirical Results
using Mean Squared Error:
Linear Trend Techniques
Average MSE Values (Millions) over 92 Series

Technique	Fitting	Forecast Horizon 1	2	3	4	5	6	Avg 1-6
Double Moving Average	24.7	33.5	41.0	187.2	327.2	396.0	147.1	188.7
Double Exponential Smoothing	18.1	27.6	34.9	161.4	284.7	341.9	165.0	169.8
Double Running Median	31.7	52.1	75.6	260.5	431.2	533.0	328.5	280.2
Double Smoothed Median	18.1	27.9	45.7	211.8	391.4	504.8	293.5	245.9
Robust Regression	31.7	112.4	144.2	393.2	625.8	760.6	442.9	413.2

Tabled values are average MSE for the indicated technique over the 92 monotonic trend series.

6.2.2.3 Results Using MAPE

The results when using the mean absolute percentage error are in Table 21. The fitting and forecasting advantage went to the double smoothed median technique. For fitting, the double exponential smoothing technique was

Table 21

Empirical Results
using Mean Absolute Percentage Error:
Linear Trend Techniques
Average MAPE Values over 92 Series

Technique	Fitting	Forecast Horizon 1	2	3	4	5	6	Avg 1-6
Double Moving Average	7.7	8.4	9.5	11.8	13.5	18.5	19.3	13.5
Double Exponential Smoothing	7.3	8.5	9.7	12.00	13.6	17.7	18.4	13.3
Double Running Median	8.6	9.4	12.3	15.2	17.3	22.6	23.5	16.7
Double Smoothed Median	7.2	8.3	9.2	11.3	12.5	16.1	16.8	12.4
Robust Regression	8.7	8.6	9.7	11.2	12.7	15.9	16.2	12.4

Tabled values are average MAPE for the indicated technique over the 92 monotonic trend series.

a close second, while robust regression fit poorest. For forecasting, robust regression essentially tied for best performance on average over the six horizons considered. robust regression was competitive over all horizons, and seemed to be superior for the longer horizons. It is interesting to note that double smoothed median was superior to linear exponential smoothing for fitting and for every horizon considered, and that the difference between them seems to be increasing with longer horizons. Although the double moving average did not fit as well as linear exponential smoothing, its forecasts were superior to those from linear exponential smoothing for horizons 1 through 4.

6.2.2.4 Results Using Theil's U

The results from use of Theil's U are presented in Table 22. The techniques that included exponential smoothing procedures achieved the best fit, with double smoothed medians fitting somewhat better than linear exponential smoothing. Techniques that forecast best over the separate horizons were double moving averages (lowest median three times), double smoothed median (lowest median twice), and double running median (lowest median once).

Table 22

Empirical Results
using Theil's U:
Linear Trend Techniques
Theil's U Values over 91 Series

		Forecast Horizon						
Technique	Fitting	1	2	3	4	5	6	1-6
Double Moving Average	.892	.884	.881	.849	.809	.957	.811	.997
Double Exponential Smoothing	.825	.900	.906	.897	.814	.902	.845	.992
Double Running Median	.949	1.048	.839	.939	.857	1.012	.945	1.027
Double Smoothed Median	.817	.892	.947	.880	.791	.862	.905	.979
Robust Regression	.972	1.091	.987	.887	.933	.898	.948	.974

The fitting column represents the average Theil's U value for each technique. The columns for the different horizons contain median values for 91 series. The final column is developed by summing the numerator and denominator of the Theil's U equation, then calculating one Theil's U value as a measure of each technique's forecasting ability over the set of forecasts.

Considering the techniques' average ability over all horizons, the nonparametric techniques of double smoothed median and robust regression were superior.

6.2.2.5 Linear Trend Techniques Conclusion

Comparison of the forecasting performance of these five techniques over the subset of ninety-two series leads to the conclusion that nonparametric techniques indeed hold promise. The nonparametric techniques of robust regression and double smoothed median provided superior forecasts on average over the six horizons for the more meaningful relative measures of MAPE and Theil's U.

While the parametric techniques performed better when raw errors are considered, these results should be discounted due to the great difference in the levels of the series considered. MAD and MSE performance of a technique is greatly influenced by its performance on the series with a high level and a large variability. For example, series 4 (SID = 4) has a last fitted value of 350,000 and an MSE fitting value (using linear exponential smoothing) of 1.17 billion. This series has over four times the influence of the second ranked series (SID = 391) with a last fitted value of about 113 thousand, yet an MSE fitting value of 241 million. The third most influential series (SID = 49) has respective values of 80,389 and 144 million. On the other hand, the MAPE values for these three series are 21%, 7% and 18%. Comparison of results using relative terms is much more equitable and less prone to dominance by particular series.

6.2.3 Empirical Study Summary

The empirical study was performed to see how these techniques would compare on real series. As displayed in Tables 15 - 22, parametric techniques forecast best on average five times out of eight. They had lower errors for MAD, MAPE and Theil's U on the smaller set of no monotonic trend series. They also had lower MAD and MSE values for the monotonic trend series.

The best parametric technique appears to be exponential smoothing. Exponential smoothing is superior in terms of MAPE for both subsets, and for MSE and Theil's U on the larger subset. The moving average technique is superior based on MAD, MSE and Theil's U on the smaller subset of series, and for the MAD on the larger subset. Discounting the raw error measures of MAD and MSE leaves exponential smoothing with stronger credentials.

The best nonparametric techniques appear to be Walsh averages for no monotonic trend series, and robust regression when monotonic trend series are used. However, if MSE is chosen as the error measure, then smoothed or running median techniques perform better.

6.3 Summary of Performance

The tests of the performance of nonparametric extrapolative techniques indicate that these techniques would not perform well on series where there was a high level of autocorrelation, or where the true error term distribution was normal. They do, however, perform well on empirical series with monotonic trend. Smoothed medians was the leading contender through the simulation study. The empirical study found that robust regression would be the best nonparametric technique for horizons 1-6 on average. The double smoothed median technique performed better for horizons 1 and 2, while robust regression gained the overall advantage on the longer horizons tested.

CHAPTER VII

SUMMARY AND CONCLUSIONS

7.1 Summary

The objective of this research was to explore the extension of nonparametrics into the time series arena, and to evaluate the relative performance of nonparametric and parametric extrapolative techniques. Four parametric techniques were selected for inclusion based upon their solid performance in prior studies and their tractability; two techniques were designed for stationary series, and two were designed for linear trend series. Since no references to nonparametric extrapolative techniques were found, six techniques were developed or modified for time series application. Three techniques were for stationary series and three were for linear trend series. Since results can be affected by the error measure used, a study was performed to select the appropriate error measure or measures. Four error measures were selected based upon usage by researchers, the structure of the error measure, and the joint behavior of the error measures. A simulation study disclosed that when the series error terms were in fact normally distributed, the parametric techniques outperformed the nonparametric techniques. The study also illustrated that when the errors were Cauchy, the nonparametric techniques outperformed the parametric techniques so long as there was not a high level of first order autocorrelation. The simulation study did not attempt to test for performance when outliers were present or when there were shifts in level.

The empirical study used the 111 series subset of the M-competition data. All series were deseasonalized. Series were divided into two subgroups based on a Rank-Spearman test for monotonic trend. Stationary techniques were used on the subset that did not test to have monotonic trend, while linear trend techniques were used on those that did test to have monotonic trend. On the stationary subset, the nonparametric techniques were superior using the raw error measures, while the parametric techniques were superior using the relative error measures. On the larger monotonic trend subset, the nonparametric techniques were superior using the relative error measures, while the parametric techniques were superior using the raw error measures.

7.2 Conclusions and Recommendations for Further Study

While the raw error measures of MAD and MSE are commonly used in forecasting, their value lies in comparison of forecasting techniques on the same series. When different series have different levels, different inherent coefficients of variation, and different units of measure, the use of average MAD or average MSE cannot be justified. Their use in simulation studies where levels and variation are controlled is appropriate. Their average results in comparisons over empirical series must be discounted. This is probably the reason why Makridakis, et. al. (1982) leaned so strongly toward MAPE, although they did report MSE results as well.

Considering the average relative error results from the empirical study, the nonparametric techniques were superior for the larger subset of monotonic trend series. The difference was about a 7% improvement based upon MAPE and over a 2% improvement based upon Theil's U values. The nonparametric techniques did not fare so well on the series without monotonic trend. The relative differences were about a negative 1% for MAPE and almost a negative 2% for Theil's U.

To help reconcile these results, it is interesting to compare exponential smoothing results between the M-competition and this study. Although both used the same series, the only comparison must be made based upon the exponential smoothing results when measured using MAPE. The double exponential smoothing MAPE from the M-competition for the 111 series was 13.9%. Double exponential smoothing in this study yielded 13.3% on the 92 monotonic trend series, while single exponential smoothing yielded 16.7% on the 19 series that did not test to have monotonic trend. These results are consistent. Weighting this study's results by the number of series in each subset results in the same overall average.

The principal conclusion from this research is that nonparametric techniques hold promise in extrapolative forecasting. The simulation study indicated their superiority only with nonnormal errors and low levels of autocorrelation. The empirical study identified robust regression as a technique superior (on average, for horizons 1 - 6) to exponential smoothing for monotonic trend series. This level of performance places it among the best extrapolative techniques available. Additional advantages of the robust regression technique are that it can be completely automated and that it is insensitive to outliers. While it will require more computation time than exponential smoothing or moving averages, it probably would not require as much time as methods such as Bayesian forecasting or Box-Jenkins.

Further study is needed in several areas. First robust regression and double smoothed medians should be tested on both the entire 111 series subset, and on all 1001 series used in the M-competition. While these techniques would strictly not satisfy the principle of parsimony when applied to series without trend, their MAPE results could then be compared directly to MAPEs of other techniques used in the M-competition. Only further testing can validate the situations when robust regression, or other nonparametric techniques such as double smoothed medians, can help provide better forecasts and better information to the decision maker. The proposed procedures for this study parallel the procedures used in the monotonic trend portion of the empirical study. Use of all 111 series would require a modest increase in computer memory and processing time. Use of all 1001 series would require a larger increase in memory, and a much larger increase in processing time.

Second, if such testing yields results consistent with this study, the robust regression technique should be written in an alternate language (such as FORTRAN or Basic) so that it can be provided for those interested in further study and testing. While SAS is very common on mainframes, a FORTRAN or Basic program would be more useful to those who work with smaller computers. A skilled programmer could doubtlessly improve the efficiency of the SAS programs used in this research, and the programs rewritten into any alternate languages.

Third, forecasting competitions should divide the series under consideration into categories that more nearly match the design capabilities of the techniques. It is doubtful that any forecaster would apply a technique designed for cases without trend to series with known monotonic trend, this would assure bias in the forecasts.

It is not difficult to perform a Rank Spearman (or comparable) test to segregate the deseasonalized series for more appropriate forecasting. The portion of the M-competition that dealt with simple extrapolative techniques on deseasonalized series should be redone with this segregation as an early step.

Fourth, further work is needed on error measures for comparison of technique results over nonhomogeneous series, i.e. series with different variables, scales, levels, and dispersion. All error measures identified have flaws under certain conditions. Raw errors have serious logical problems. MAPE is insensitive to scale yet sensitive to level (and is particularly sensitive when series values are near zero). Theil's U is reasonable over a series of forecasts, but may not even exist for a single forecast and is extremely skewed. Development of a relative error measure that avoided the sensitivity to small series values of MAPE, and the skewness inherent in Theil's U type measures could be of great help in evaluating techniques over dissimilar series.

Fifth, the question of whether error measures display the same relationships across forecasting models needs to be answered. This study used an exponential smoothing model to generate forecast errors when the relationships between different error measures was considered. The assumption was made that, for example, the relationship between MAD and MAPE depended more upon the structure of the error measures themselves than upon the forecasting model or technique used. The use of different forecasting models on different types of simulated series could be used to answer this research question.

Sixth, the procedure of selecting parameters based upon one period ahead forecasts should be further considered. If researchers and practitioners report better success using different techniques for different horizons, the use of the same model for all horizons seems very ill advised.

Seventh, the anomaly discovered when the MAPE is used to evaluate forecasts of a series with Cauchy errors needs to be researched. This could be a weakness of the MAPE, or a result of the combination.

Eighth, the relationship between autocorrelation levels and exponential average smoothing constant could be established. In particular, where within the first order autocorrelation range of .3 to .5 does the exponential smoothing coefficient break the .3 level identified by Brown (1963).

Ninth, the autocorrelation levels within deseasonalized business type series should be determined. The simulation study indicated that the nonparametric techniques would not do well when there was a high level of autocorrelation. Their good performance on empirical series tends to imply that at least the first order autocorrelation is not high, but only further research can provide a definitive answer.

Tenth, research should be performed to see how the choice of an error measure affects the resulting model. Several examples were given where smoothing constants (therefore forecasting models) were different when best fit using different error measures. If all error measures result in the choice of the same forecasting model, then the choice of an error measure carries much less significance. If different error measures result in the choice of different forecasting models, then the choice is very critical. A preliminary study (not reported in this study) indicates the effect lies somewhere between these extremes, with different error measures frequently, but not always, providing similar models.

APPENDIX A

SAMPLE ERROR MEASURE EVALUATION PROGRAMS

These programs were written and run using the VMS Version of SAS, Release 5.03 and Release 5.16. Reference SAS Institute Inc. (1982 and 1985).

SAMPLE PROGRAM FOR THE EVALUATION OF ERROR MEASURES. THIS EXAMPLE GENERATES THE SERIES WITH NORMAL ERRORS, THE LARGEST VARIATION, AND NO OUTLIERS. AN EXPONENTIAL SMOOTHING MODEL IS USED TO GENERATE ERRORS, WHICH ARE THEN SUMMARIZED BY USE OF MAD, MSE, MAPE AND THEIL'S U.

```
DATA VALUES;                  This data step generates
OPTIONS NOCENTER;                   the series and the
DO I = 1 TO 20;                   exponential average.
  DO J = 1 TO 520;
    IF RANUNI(86) LT 1.1 THEN K = 2.2;
    ELSE K = 10;
    V = 10. + K*RANNOR(46);
    IF J = 1 THEN EA = V;
    ELSE EA = .2*V + .8*EA;
    OUTPUT VALUES;
  END;
END;

DATA NORMAL;                 This data step calculates
SET VALUES;                         the error measures.
DEV = V-LAG(EA);
LAGV = LAG(V);
IF _N_ = 1 THEN REP = 1;                  There are 20
RETAIN REP;                               replications.
N+1;  M+1;
IF M LT 2 THEN GO TO NEXTOBS;
MADI = ABS(DEV);                       The instant MAD.
MADS + MADI;                   The sum of the absolute
MSEI = (DEV)**2;                            deviations.
MSES + MSEI;
MAPEI = (MADI/ABS(V))*100;
MAPES + MAPEI;
THDI = ((V-LAGV)/LAGV)**2;           Theil's U has a
THDS + THDI;                         denominator and a
THNI = (DEV/LAGV)**2;        numerator, which must
THNS + THNI;                 be summed seperately.
IF N = 20 THEN DO;
  MAD = MADS/N;              Calculate error measures
  MSE = MSES/N;              for twenty observations
  MAPE = MAPES/N;                      and then reset
  TH = (THNS/THDS)**.5;                for the next 20.
  KEEP REP K MAD MSE MAPE TH;
  N=0; MADS=0; MSES=0; MAPES=0; THDS=0; THNS=0;
  IF M = 20 THEN RETURN;                 Note that REP
  ELSE OUTPUT;               (replication) is output.
END;                                Each of the twenty
IF M = 520 THEN DO;                  replications uses
  REP + 1;                     520 observations, the
  M = 0;               first 20 observations are for
END;                                   initialization.
IF N GT 22 THEN ABORT RETURN;
```

```
IF M GT 522 THEN ABORT RETURN;
NEXTOBS:;

DATA _NULL_;               This data step calculates
SET WORK.NORMAL;            Spearman's correlations.
FILE CKDIS22;
PUT REP K MAD MSE MAPE TH;
PROC CORR DATA=NORMAL SPEARMAN NOSIMPLE NOPROB
         NOPRINT OUTS=CORRS;
VAR MAD MSE MAPE TH;
BY REP;                Note, results by replication.
TITLE 'CORRELATION RESULTS, DISPERSION = 2.2 *
      NOMINAL';

DATA NEXT;                 This data step sorts and
SET CORRS;                     prints the results.
PROC SORT;
BY _NAME_;
PROC PRINT;

DATA _NULL_;               This data step calculates
SET WORK.NEXT;                the values for the box
PROC UNIVARIATE;                              plots.
VAR MAD MSE MAPE TH;
BY _NAME_;
```

SAMPLE PROGRAMS FOR PLOTTING. THESE TYPE PROGRAMS WERE USED TO GET SIMPLE SCATTERPLOTS DISPLAYING THE JOINT BEHAVIOR OF SELECTED ERROR MEASURES.

```
DATA PLOTT;                          Simple scatterplot, dot
INFILE CKDIS22;                              matrix results.
INPUT REP K MAD MSE MAPE TH;
TITLE 'PLOT OF ERROR MEASURES, DISPERSION = 2.2
      * NOMINAL';
PROC PLOT;
PLOT MAD*MSE/HAXIS=0 TO 40 BY 4 VAXIS=0 TO 3.25
     BY .25;
```

THIS CODE RESULTS IN A PLOT COMBINING THE NOMINAL DISPERSION (SIGMA = 1) AND LARGEST DISPERSION (SIGMA = 2.2) RESULTS.

```
DATA ONE;                            Scatterplot combining
INFILE CKDIS10;                         the results of two
INPUT REP K MAD MSE MAPE THEILS;                 different
INFILE CKDIS22;                               dispersions.
INPUT REP2 K2 MAD2 MSE2 MAPE2 THEILS2;
PROC PLOT;
PLOT MAD*MSE='O' MAD2*MSE2='+' / OVERLAY;
```

THIS THIRD EXAMPLE RESULTS IN A SASGRAPH GRAPHICS STRING FILE THAT WAS THEN RUN THROUGH A LASER PRINTER FOR PUBLICATION QUALITY PLOTS. PLOT SIZE AND SYMBOL SIZE ARE REDUCED FOR REPRODUCTION AND LEGIBILITY. AXES ARE DEFINED AND LEGENDS FOR THE PLOT SYMBOLS ARE INCLUDED WITH THE PLOT. THIS PROGRAM PRODUCES THE PLOTS TO ILLUSTRATE THE EFFECT OF OUTLIERS.

```
LIBNAME ERREVAL '[RKANKEY.ERREVAL]';
GOPTIONS DEVICE=TEK4010
    HSIZE=5.5 VSIZE=6.5          Reducing the plot
    DISPLAY                             dimensions.
    GSFMODE=REPLACE
    GSFNAME=GPOUTA;
FILENAME VECTOR ' [RKANKEY.ERREVAL]PLOTOUT';
DATA A;
SET ERREVAL.PLOTOUT;
SYMBOL1 V=PLUS H=.2;             Defining the plot
SYMBOL2 V=TRIANGLE H=.5;                  symbols.
SYMBOL3 V=NONE;
SYMBOL4 V=NONE;
PROC GPLOT;
     AXIS1 LABEL=(F=SIMPLEX H=1.5 'MAD')
           ORDER=(0 TO 4 BY 1)
           VALUE=(F=SIMPLEX)
           OFFSET=(1,1);
     AXIS2 LABEL=(F=SIMPLEX H=1.5 'MSE')
           ORDER=(0 TO 50 BY 10)
```

```
                VALUE=(F=SIMPLEX)
                OFFSET=(1,1);
          AXIS3 LABEL=(F=SIMPLEX H=1.5 'MAPE')
                ORDER=(0 TO 400 BY 100)
                VALUE=(F=SIMPLEX)
                OFFSET=(1,1);
          AXIS4 LABEL=(F=SIMPLEX H=1.5 'U')
                ORDER=(0 TO 1.2 BY .2)
                VALUE=(F=SIMPLEX)
                OFFSET=(1,1);
          LEGEND1 LABEL=(F=TITALIC  'PROB. OF
                        OUTLIER:: ')
             VALUE=(F=SIMPLEX '=0' '=.02')
             ACROSS=2
             FRAME;
         PLOT MAD*MSE=SIGMA / VAXIS=AXIS1 HAXIS=AXIS2
                              LEGEND=LEGEND1 FRAME;
         PLOT2 MAD*MSE=3 / VAXIS=AXIS1 HAXIS=AXIS2
                           NOLEGEND;
         PLOT MAD*MAPE=SIGMA / VAXIS=AXIS1
                           HAXIS=AXIS3 LEGEND=LEGEND1
                           FRAME;
         PLOT2 MAD*MAPE=3 / VAXIS=AXIS1 HAXIS=AXIS3
                           NOLEGEND;
         PLOT MAD*THEIL=SIGMA / VAXIS=AXIS1
                           HAXIS=AXIS4 LEGEND=LEGEND1
                           FRAME;
         PLOT2 MAD*THEIL=3 / VAXIS=AXIS1 HAXIS=AXIS4
                           NOLEGEND;
         PLOT MSE*MAPE=SIGMA / VAXIS=AXIS2
                           HAXIS=AXIS3 LEGEND=LEGEND1
                           FRAME;
         PLOT2 MSE*MAPE=3 / VAXIS=AXIS2 HAXIS=AXIS3
                           NOLEGEND;
         PLOT MSE*THEIL=SIGMA / VAXIS=AXIS2
                           HAXIS=AXIS4 LEGEND=LEGEND1
                           FRAME;
         PLOT2 MSE*THEIL=3 / VAXIS=AXIS2 HAXIS=AXIS4
                           NOLEGEND;
         PLOT MAPE*THEIL=SIGMA / VAXIS=AXIS3
                           HAXIS=AXIS4 LEGEND=LEGEND1
                           FRAME;
         PLOT2 MAPE*THEIL=SIGMA / VAXIS=AXIS3
                           HAXIS=AXIS4 NOLEGEND;
     RUN;
```

THIS FOURTH EXAMPLE PROGRAM PRODUCES THE PLOTS TO ILLUSTRATE THE EFFECT OF INCREASED DISPERSION.

```
LIBNAME ERREVAL '[RKANKEY.ERREVAL]';
GOPTIONS DEVICE=TEK4010
    HSIZE=5.5 VSIZE=6.5
    DISPLAY
    GSFMODE=REPLACE
    GSFNAME=GPDIS3M;
FILENAME VECTOR ' [RKANKEY.ERREVAL]PLOTDIS';
DATA A;
SET ERREVAL.PLOTDIS;
SYMBOL1 V=PLUS H=.2;
SYMBOL2 V=TRIANGLE H=.5;
SYMBOL3 V=NONE;
SYMBOL4 V=NONE;
PROC GPLOT;
     AXIS1 LABEL=(F=SIMPLEX H=1.5 'MAD')
           ORDER=(0 TO 3 BY 1)
           VALUE=(F=SIMPLEX)
           OFFSET=(1,1);
     AXIS2 LABEL=(F=SIMPLEX H=1.5 'MSE')
           ORDER=(0 TO 14 BY 2)
           VALUE=(F=SIMPLEX)
           OFFSET=(1,1);
     AXIS3 LABEL=(F=SIMPLEX H=1.5 'MAPE')
           ORDER=(0 TO 50 BY 10)
           VALUE=(F=SIMPLEX)
           OFFSET=(1,1);
     AXIS4 LABEL=(F=SIMPLEX H=1.5 'U')
           ORDER=(0 TO 1.2 BY .2)
           VALUE=(F=SIMPLEX)
           OFFSET=(1,1);
     LEGEND1 LABEL=(F=TITALIC  'SIGMA: ')
        VALUE=(F=SIMPLEX '=1' '=2.2')
        ACROSS=2
        FRAME;
    PLOT MAD*MSE=SIGMA / VAXIS=AXIS1 HAXIS=AXIS2
                         LEGEND=LEGEND1 FRAME;
    PLOT2 MAD*MSE=3 / VAXIS=AXIS1 HAXIS=AXIS2
                    NOLEGEND;
    PLOT MAD*MAPE=SIGMA / VAXIS=AXIS1
                    HAXIS=AXIS3 LEGEND=LEGEND1
                    FRAME;
    PLOT2 MAD*MAPE=3 / VAXIS=AXIS1 HAXIS=AXIS3
                    NOLEGEND;
    PLOT MAD*THEIL=SIGMA / VAXIS=AXIS1
                    HAXIS=AXIS4 LEGEND=LEGEND1
                    FRAME;
    PLOT2 MAD*THEIL=3 / VAXIS=AXIS1 HAXIS=AXIS4
                    NOLEGEND;
```

```
        PLOT MSE*MAPE=SIGMA / VAXIS=AXIS2
                        HAXIS=AXIS3 LEGEND=LEGEND1
                        FRAME;
        PLOT2 MSE*MAPE=3 / VAXIS=AXIS2 HAXIS=AXIS3
                        NOLEGEND;
        PLOT MSE*THEIL=SIGMA / VAXIS=AXIS2
                        HAXIS=AXIS4 LEGEND=LEGEND1
                        FRAME;
        PLOT2 MSE*THEIL=3 / VAXIS=AXIS2 HAXIS=AXIS4
                        NOLEGEND;
        PLOT MAPE*THEIL=SIGMA / VAXIS=AXIS3
                        HAXIS=AXIS4 LEGEND=LEGEND1
                        FRAME;
        PLOT2 MAPE*THEIL=SIGMA / VAXIS=AXIS3
                        HAXIS=AXIS4 NOLEGEND;
    RUN;
```

APPENDIX B

SAMPLE SIMULATION PROGRAMS

These programs were written and run using the VMS Version of SAS, Release 5.03 and Release 5.16. Reference SAS Institute Inc. (1982 and 1985).

SAMPLE PROGRAM FOR SERIES GENERATION. THE SERIES GENERATED BELOW IS AUTOREGRESSIVE OF ORDER ONE, WITH A LEVEL OF 20 AND COEFFICIENT OF VARIATION OF 15%. HERE THE PHI VALUE IS .5. WHEN THE CAUCHY SERIES WERE GENERATED, A SCALE PARAMETER OF 1.9393 WAS USED (VS 3) SO THAT THE MIDRANGE OF THE CAUCHY AND THE MIDRANGE OF THE NORMAL WERE APPROXIMATELY EQUAL INTERVALS.

```
DATA _NULL_;
DO T = 1 TO 6000;
  IF T EQ 1 THEN X = 20;
  ELSE X = .5*X + 10 + 3*RANNOR(75);
  FILE NEWN5;
  PUT T X;
END;
```

THIS PROGRAM GENERATES A GRAPHICS STREAM FILE TO DISPLAY THE PLOT OF A FIRST ORDER AUTOREGRESSIVE SERIES WITH A PHI VALUE OF .3, ALREADY FILED IN AR1PHI3N. BY PLOTTING A NULL CHARACTER WITH THE PLOT2 STATEMENT, A SCALED RIGHT HAND AXIS IS ADDED TO THE PLOT.

```
GOPTIONS DEVICE=TEK4010
    HSIZE=5.5 VSIZE=5.5
    DISPLAY
    GSFMODE=REPLACE
    GSFNAME=GRAFPHI5;
FILENAME VECTOR ' [RKANKEY.STAT]AR1PHI3N';
DATA A;
     INFILE VECTOR;
     INPUT T Y;
SYMBOL1 L=1 V=STAR I=JOIN H=.8;
SYMBOL2;
PROC GPLOT;
     PLOT Y*T=1 / FRAME VAXIS=0 TO 30 BY 10;
     PLOT2 Y*T=2 /VAXIS=0 TO 30 BY 10;
RUN;
```

MOVING AVERAGE PROGRAM. THIS PROGRAM UTILIZES THE LAG FUNCTIONS AVAILABLE IN SAS TO FIT THE SERIES USING SINGLE MOVING AVERAGES OF LENGTHS FROM ONE TO NINE, FORECASTS ARE FOR ONE PERIOD AHEAD.

```
DATA _NULL_;            Infile sixkc7 is a file with
INFILE SIXKC7;       6000 observations and a Cauchy
INPUT T X;              error term, Phi equal to .7.
RETAIN;
IF T = 1 THEN DO;
  K = 1;
  COUNTER = 0;                The V array was used for
END;                           checking and debugging.
ARRAY V{18} V1-V18;
ARRAY MA{9} MA1-MA9;           Nine moving averages.
ARRAY SUMS{11,11} S1-S121;
ERRTYPE = 2;         Errtype = 1 when normal errors
PHI = .7;               are used.  Errtype = 2 when
LAGX = LAG(X);             Cauchy errors are used.
COUNTER + 1;
L1=LAGX;                   Renaming the lag values
L2=LAG2(X);                 for easier programming
L3=LAG3(X);                             statements.
L4=LAG4(X);
L5=LAG5(X);
L6=LAG6(X);
L7=LAG7(X);
L8=LAG8(X);
SUM=X;
MA1=SUM;
SUM=SUM+L1;
MA2=SUM/2;
SUM=SUM+L2;
MA3=SUM/3;
SUM=SUM+L3;
MA4=SUM/4;
SUM=SUM+L4;
MA5=SUM/5;
SUM=SUM+L5;
MA6=SUM/6;                    Moving averages up to
SUM=SUM+L6;               length 9 are calculated.
MA7=SUM/7;
SUM=SUM+L7;
MA8=SUM/8;
SUM = SUM+L8;
MA9=SUM/9;
DO I=1 TO 9;
  N = I;
  F = LAG(MA{I});
  E = X - F;
  V1 = T;
  V2 = K;
```

```
  V3 = X;                     Many of these elements of
  V4 = N;                     the V array are duplicate
  V5 = MA{I};             variables used in debugging.
  V6 = F;
  V7 = E;
  IF COUNTER LT 21 THEN GO TO CUT;
  V8 = ABS(E);                 The error is manipulated
  V9 = E**2;                       as necessary for the
  V10 = (V8/X)**2;            various error measures.
  V11 = ABS(E/X);
  V12 = (X-LAGX)**2;
  V13 = ABS(E/LAGX);
  V14 = ABS ((X-LAGX)/LAGX);
  V15 = ABS(E);
  V16 = ABS(X-LAGX);
  V17 = (E/LAGX)**2;
  V18 = ((X-LAGX)/LAGX)**2;
  SUMS{I,1} + V8;                   These various items
  SUMS{I,2} + V9;                            are summed.
  SUMS{I,3} + V10;
  SUMS{I,4} + V11;
  SUMS{I,5} + V12;
  SUMS{I,6} + V13;
  SUMS{I,7} + V14;
  SUMS{I,8} + V15;
  SUMS{I,9} + V16;
  SUMS{I,10} + V17;
  SUMS{I,11} + V18;                                Eight
  IF COUNTER EQ 60 THEN DO;               error measures
    MAD = SUMS{I,1}/40;                 are calculated,
    MSE = SUMS{I,2}/40;              although only four
    MSPE = SUMS{I,3}/40;              are ultimately
    MAPE = SUMS{I,4}/40;                       reported.
    TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
    TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
    TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
    TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
    FILE SUM6KC7M;
    PUT T K ERRTYPE N PHI MAD MSE MAPE MSPE
        TH1 TH2 TH3 TH4;
  END;
CUT:END;

IF COUNTER = 60 THEN DO;                   Resetting the
  COUNTER = 0;                        counters and arrays
  K+1;                                       for the next
                                               iteration.
  DO M = 1 TO 11;
    DO N = 1 TO 11;
      SUMS{M,N} = 0;
    END;
  END;
END;
```

EXPONENTIAL AVERAGE PROGRAM. THIS SAMPLE PROGRAM OPERATES ON THE SERIES WITH CAUCHY ERRORS AND PHI = .7. FOR SINGLE EXPONENTIAL SMOOTHING THE SMOOTHING CONSTANTS EVALUATED RANGE FROM ZERO TO .5. THE KEY TO USE OF RECURSIVE EQUATIONS IN SAS IS THE SMALL TERM ON LINE 4 OF THE PROGRAM, WITHOUT WHICH SAS SETS VALUES TO MISSING WHEN IT READS THE NEXT OBSERVATION'S VALUES.

```
 1      DATA _NULL_;
 2      INFILE SIXKC7;
 3      INPUT T X;
 4      RETAIN;
 5      IF T=1 THEN DO;
 6           K=1;
 7           COUNTER=0;
 8      END;
 9      ARRAY V{18} V1-V18;
10      ARRAY EX{11} EX1-EX11;
11      ARRAY SUMS{11,11} S1-S121;
12      ERRTYPE = 2;
13      PHI = .7;
14      LAGX = LAG(X);                                   For
15      COUNTER +1;                              exponential
16      DO I=1 TO 11;                              smoothing
17        C=(I-1)*.05;                         the smoothing
18        IF COUNTER=1 THEN EX{I} = 20;  constant ranges
19        EX{I} = C*X+(1-C)*EX{I};       from zero to .5.
20        F = LAG(EX{I});
21        IF COUNTER = 1 THEN F=20;             The initial
22        E=X-F;                                 exponential
23        V1 = T;                            average is set to
24        V2 = K;                              equal the true
25        V3 = X;                            population mean.
26        V4 = C;                         Initialization is not
27        V5 = EX{I};                       critical here since
28        V6 = F;                      no errors are collected
29        V7 = E;                                 until t = 21.
30        IF COUNTER LT 21 THEN GO TO CUT;
31        V8 = ABS(E);
32        V9 = E**2;
33        V10 = (V8/X)**2;
34        V11 = ABS(E/X);
35        V12 = (X-LAGX)**2;
36        V13 = ABS(E/LAGX);
37        V14 = ABS ((X-LAGX)/LAGX);
38        V15 = ABS(E);
39        V16 = ABS(X-LAGX);
40        V17 = (E/LAGX)**2;
41        V18 = ((X-LAGX)/LAGX)**2;
42        SUMS{I,1} + V8;
43        SUMS{I,2} + V9;
44        SUMS{I,3} + V10;
```

```
          SUMS{I,4} + V11;
          SUMS{I,5} + V12;
          SUMS{I,6} + V13;
          SUMS{I,7} + V14;
          SUMS{I,8} + V15;
          SUMS{I,9} + V16;
          SUMS{I,10} + V17;
          SUMS{I,11} + V18;
          IF COUNTER EQ 60 THEN DO;                   Eight
            MAD = SUMS{I,1}/40;              error measures
            MSE = SUMS{I,2}/40;             are calculated.
            MSPE = SUMS{I,3}/40;
            MAPE = SUMS{I,4}/40;
            TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
            TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
            TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
            TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
            FILE SUM6KC7E;
            PUT T K ERRTYPE C PHI MAD MSE MAPE MSPE
                TH1 TH2 TH3 TH4;
          END;
        CUT:END;

        IF COUNTER = 60 THEN DO;      Counters and arrays
          COUNTER = 0;                    are set for the
          K+1;                            next iteration.
          DO M = 1 TO 11;
            DO N = 1 TO 11;
              SUMS{M,N} = 0;
            END;
          END;
        END;
```

RUNNING MEDIAN PROGRAM. THIS PROGRAM USES LAG FUNCTIONS AND BUBBLE SORTS TO FIND THE RUNNING MEDIANS OF LENGTH THREE THROUGH SIX. ONE KEY TO THIS PROGRAM IS RENAMING THE LAG VALUES BEFORE SORTING, OTHERWISE THE LAG VALUES THEMSELVES BECOME JUMBLED DURING THE PROCESS.

```
DATA _NULL_;                      Infile sixkc5 is a file
INFILE SIXKC5;                     with 6000 observations
INPUT T XI;                         generated with Cauchy
IF T=1 THEN DO;                       errors and Phi = .5.
     K=1;
     COUNTER=0;
END;
ARRAY V{18} V1-V18;
ARRAY SUMS{4,11} S1-S44;
ERRTYPE = 2;
PHI = .5;
X = XI;
LAGX = LAG(X);                      Defining the arrays as
COUNTER +1;                           shown below and then
L1=LAGX;              calculating the running medians
L2=LAG2(X);              in the order indicated takes
L3=LAG3(X);              advantage of the sorting for
L4=LAG4(X);     the three period when calculating
L5=LAG5(X);              the four period median, etc.
IF COUNTER LE 14 THEN GO TO SKIP;
ARRAY TEMP3{3} X L1 L2;
ARRAY TEMP4{4} X L1 L2 L3;
ARRAY TEMP5{5} X L1 L2 L3 L4;
ARRAY TEMP6{6} X L1 L2 L3 L4 L5;

DO H=1 TO 3;                      Calculate running median
MORE=0;                                   of length three.
DO I = 1 TO 2;
     IF TEMP3{I+1} LT TEMP3{I} THEN DO;
          TEMP = TEMP3{I};
          TEMP3{I} = TEMP3{I+1};
          TEMP3{I+1} = TEMP;
          MORE = 1;
          END;
     END;
     IF MORE = 0 THEN GO TO DONE3;
END;
DONE3:;
RM3 = TEMP3{2};

DO H = 1 TO 4;                    Calculate running median
MORE = 0;                                  of length four.
DO I = 1 TO 3;
  IF TEMP4{I+1} LT TEMP4{I} THEN DO;
    TEMP = TEMP4{I};
    TEMP4{I} = TEMP4{I+1};
```

```
          TEMP4{I+1} = TEMP;
          MORE = 1;
          END;
        END;
        IF MORE = 0 THEN GO TO DONE4;
      END;
      DONE4:;
      RM4 = (TEMP4{2} + TEMP4{3})/2;

      DO H = 1 TO 5;          Calculate running median
      MORE = 0;                         of length five.
        DO I = 1 TO 4;
          IF TEMP5{I+1} LT TEMP5{I} THEN DO;
          TEMP = TEMP5{I};
          TEMP5{I} = TEMP5{I+1};
          TEMP5{I+1} = TEMP;
          MORE = 1;
          END;
        END;
        IF MORE = 0 THEN GO TO DONE5;
      END;
      DONE5:;
      RM5 = TEMP5{3};

      DO H = 1 TO 6;          Calculate running median
      MORE = 0;                          of length six.
        DO I = 1 TO 5;
          IF TEMP6{I+1} LT TEMP6{I} THEN DO;
          TEMP = TEMP6{I};
          TEMP6{I} = TEMP6{I+1};
          TEMP6{I+1} = TEMP;
          MORE = 1;
          END;
        END;
        IF MORE = 0 THEN GO TO DONE6;
      END;
      DONE6:;
      RM6 = (TEMP6{3} + TEMP6{4})/2;

      SKIP:;
      ARRAY RM{4} RM3 RM4 RM5 RM6;
      LAGXI = LAG(XI);
      DO I=1 TO 4;
        F = LAG(RM{I});
        N = I+2;
        E=XI-F;
        V1 = T;
        V2 = K;
        V3 = XI;
        V4 = N;
        V5 = RM{I};
        V6 = F;
```

```
    V7 = E;
    IF COUNTER LT 21 THEN GO TO CUT;
    V8 = ABS(E);
    V9 = E**2;                               Manipulating
    V10 = (V8/XI)**2;                          the error.
    V11 = ABS(E/XI);
    V12 = (XI-LAGXI)**2;
    V13 = ABS(E/LAGXI);
    V14 = ABS ((XI-LAGXI)/LAGXI);
    V15 = ABS(E);
    V16 = ABS(XI-LAGXI);
    V17 = (E/LAGXI)**2;
    V18 = ((XI-LAGXI)/LAGXI)**2;
    SUMS{I,1} + V8;                              Summing.
    SUMS{I,2} + V9;
    SUMS{I,3} + V10;
    SUMS{I,4} + V11;
    SUMS{I,5} + V12;
    SUMS{I,6} + V13;
    SUMS{I,7} + V14;
    SUMS{I,8} + V15;
    SUMS{I,9} + V16;
    SUMS{I,10} + V17;
    SUMS{I,11} + V18;
    IF COUNTER EQ 60 THEN DO;
      MAD = SUMS{I,1}/40;    Calculating the error
      MSE = SUMS{I,2}/40;                 measures.
      MSPE = SUMS{I,3}/40;
      MAPE = SUMS{I,4}/40;
      TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
      TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
      TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
      TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
      FILE SUM6KC5R;
      PUT T K ERRTYPE N PHI MAD MSE MAPE MSPE
          TH1 TH2 TH3 TH4;
    END;
  CUT:END;

  IF COUNTER = 60 THEN DO;                    Resetting.
    COUNTER = 0;
    K+1;
    DO M = 1 TO 4;
      DO N = 1 TO 11;
        SUMS{M,N} = 0;
      END;
    END;
  END;
```

WALSH AVERAGE PROGRAM - FIRST HALF. THIS HALF OF THE PROGRAM PRODUCES MOVING WALSH AVERAGES OF LENGTHS THREE THROUGH SIX. THE VALUE OF THE LATEST WALSH AVERAGE IS THE FORECAST VALUE FOR THE FUTURE.

```
DATA _NULL_;
INFILE SIXKC7;
INPUT T XI;
IF T EQ 1 THEN COUNTER = 0;
COUNTER+1;
W1 = XI;                          Renaming for ease of
W2 = LAG(XI);                      programming, and to
W3 = LAG2(XI);                         avoid jumbling
W4 = LAG3(XI);                       during any sorts.
W5 = LAG4(XI);
W6 = LAG5(XI);
WA1=0;
WA2=0;
WA3=0;
WA4=0;
IF COUNTER LT 20 THEN GO TO SKIP;
IF COUNTER EQ 60 THEN GO TO SKIP2;
ARRAY VALUES{6} W1-W6;
ARRAY Z{6,6} Z1-Z36;               The array of sums.
ARRAY ORDERING{21} UNORD1-UNORD21;
ARRAY WAL{4} WA1-WA4;  Array of Walsh Averages.
DO I = 1 TO 6;              Calculating the 6 X 6
  DO J = I TO 6;                  matrix of sums.
    Z{I,J} = (VALUES{I} + VALUES{J});
  END;
END;
Q = 1;
DO N = 3 TO 6 BY 1;               N is the number of
  K=1;                         data points considered
  DO I = 1 TO N;                in the Walsh Average.
    DO J = I TO N;
      ORDERING{K} = Z{I,J};          Pulling out the
      K+1;                         appropriate sums,
    END;                        i.e. upper triangular
  END;                               matrix elements.
  IF N EQ 3 THEN MAX = 6;
  IF N EQ 4 THEN MAX = 10;         Number of sums to
  IF N EQ 5 THEN MAX = 15;            to consider in
  IF N EQ 6 THEN MAX = 21;                each case.
                                  Walsh averages are
  DO L = 1 TO MAX;             calculated for lengths
    MORE = 0;                        four through six
    DO I = 1 TO (MAX-1);
      IF ORDERING{I+1} LT ORDERING{I} THEN DO;
        TEMP = ORDERING{I};
        ORDERING{I} = ORDERING{I+1};
        ORDERING {I+1} = TEMP;
```

```
            MORE = 1;                        This step uses a
          END;                                    bubble sort.
        END;
        IF MORE EQ 0 THEN GO TO NEXTN;
      END;
      NEXTN: IF MOD(MAX,2) EQ 1 THEN WAL{Q} =
           ORDERING{(K+1)/2}/2;                   If an odd
      ELSE WAL{Q} = (ORDERING{K/2} +          number of
           ORDERING{K/2+1})/4;    sums is considered
      Q+1;       the median is the center value after
    END;       ranking, otherwise the average of the
           two center values.  Note division by two.
    SKIP2:;                             File the results for
    FILE WAVGSC7;                      access by the second
    PUT T COUNTER XI WA1-WA4;                  program and
    SKIP:;                               reset the counter.
    IF COUNTER = 60 THEN COUNTER = 0;
```

WALSH AVERAGE PROGRAM - SECOND HALF. THIS HALF OF THE PROGRAM TAKES THE WALSH AVERAGES CALCULATED FROM THE PREVIOUS PROGRAM AND USES THEM TO FIT THE SERIES. EIGHT ERROR MEASURES ARE CALCULATED FOR EACH OF THE WALSH AVERAGE SIZES USED.

```
    DATA _NULL_;                        Note that for this
    INFILE WAVGSC7;                program, counter is an
    INPUT T COUNTER XI WA3 WA4 WA5 WA6;          input
    ERRTYPE = 2;                                variable.
    PHI = .7;
    ARRAY WAL{4} WA3 WA4 WA5 WA6;
    ARRAY SUMS{4,11} S1-S44;
    LAGXI = LAG(XI);
    DO I=1 TO 4;
      F = LAG(WAL{I});
      N = I+2;     N is the number of observations
      E=XI-F;     used in calculating the Walsh Avg.
      V1 = T;
      V2 = K;
      V3 = XI;
      V4 = N;
      V5 = WAL{I};
      V6 = F;
      V7 = E;
      IF COUNTER LT 21 THEN GO TO CUT;
      V8 = ABS(E);
      V9 = E**2;                 Manipulating the error.
      V10 = (V8/XI)**2;
      V11 = ABS(E/XI);
      V12 = (XI-LAGXI)**2;
      V13 = ABS(E/LAGXI);
      V14 = ABS ((XI-LAGXI)/LAGXI);
```

```
      V15 = ABS(E);
      V16 = ABS(XI-LAGXI);
      V17 = (E/LAGXI)**2;
      V18 = ((XI-LAGXI)/LAGXI)**2;
      SUMS{I,1} + V8;
      SUMS{I,2} + V9;                      Summing up the
      SUMS{I,3} + V10;                    terms needed for
      SUMS{I,4} + V11;                  the various error
      SUMS{I,5} + V12;                          measures.
      SUMS{I,6} + V13;
      SUMS{I,7} + V14;
      SUMS{I,8} + V15;
      SUMS{I,9} + V16;
      SUMS{I,10} + V17;
      SUMS{I,11} + V18;
      IF COUNTER EQ 60 THEN DO;
        MAD = SUMS{I,1}/40;              Calculating the
        MSE = SUMS{I,2}/40;              error measures.
        MSPE = SUMS{I,3}/40;
        MAPE = SUMS{I,4}/40;
        TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
        TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
        TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
        TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
        FILE SUM6KC7W;
        PUT T K ERRTYPE N PHI MAD MSE MAPE MSPE
            TH1 TH2 TH3 TH4;
      END;
    CUT:END;

    IF COUNTER = 60 THEN DO;
      K+1;                                     Resetting.
      DO M = 1 TO 4;
        DO N = 1 TO 11;
          SUMS{M,N} = 0;
        END;
      END;
    END;
```

SAMPLE ANALYSIS PROGRAM. THIS SAMPLE PROGRAM TAKES THE RESULTS IN TESTC5E.DAT (RESULTS ON A CAUCHY ERROR SERIES WITH PHI = .5 USING EXPONENTIAL SMOOTHING), RUNS THE SAS PROCEDURE UNIVARIATE, AND PRODUCES AN OUTPUT DATA SET CONTAINING THE MEANS, MEDIANS, MAX AND MIN, AND FIRST AND THIRD QUARTILE VALUES FOR EACH OF THE EIGHT ERROR MEASURES INITIALLY CONSIDERED. THIS DATA SET IS THEN AVAILABLE FOR PLOTTING.

```
LIBNAME STAT '[RKANKEY.STAT]';
DATA RESULTS;
INFILE TESTC5E;
INPUT T K ERRTYPE C PHI MAD MSE MAPE MSPE
      TH1 TH2 TH3 TH4;
PROC SORT;
BY C;
PROC UNIVARIATE NOPRINT;
VAR MAD MSE MAPE MSPE TH1 TH2 TH3 TH4;
BY C;
OUTPUT OUT=STAT.TESTCE MEAN=AV1-AV8 Q3=QU1-QU8
       Q1=QL1-QL8 MEDIAN=MD1-MD8 MAX=MX1-MX8
       MIN=MI1-MI8;
```

THIS PROGRAM USES A FILE SIMILAR TO THAT ABOVE TO PRODUCE BOX PLOTS AS IN FIGURES 18-21. THIS PARTICULAR EXAMPLE PRODUCES BOX PLOTS OF RUNNING MEDIAN RESULTS ON A STATIONARY SERIES WITH PHI = .3 AND NORMAL ERRORS. THE PROGRAM PRODUCES A GRAPHICS STRING FILE WHICH IS THEN RUN ON A LASER PRINTER. UNFORTUNATELY THE MISSING VALUES WOULD NOT BREAK THE LINE, SO A CERTAIN AMOUNT OF WHITEOUT WAS USED TO REMOVE UNDESIRED CONNECTING LINES.

```
LIBNAME STAT '[RKANKEY.STAT]';
GOPTIONS DEVICE=TEK4010
    HSIZE=5.5 VSIZE=6.5
    DISPLAY
    GSFMODE=REPLACE
    GSFNAME=GPBOX7;
DATA ONE;
SET STAT.AR3N6KM;
KEEP N X Y AV1 MX1 MI1 QL1 MD1 QU1;
SYMBOL1 L=1 I=JOIN;
SYMBOL2 V=NONE;
SYMBOL3 V=STAR H=1;
DELTA=.25;
Y=QU1;
X=N-DELTA; OUTPUT;
X=N+DELTA; OUTPUT;
Y=QL1;
X=N+DELTA; OUTPUT;
X=N-DELTA; OUTPUT;
```

```
        Y=QU1; OUTPUT;
        X=.; Y=.; OUTPUT;
        X=N; Y=MX1; OUTPUT;
        X=N; Y=QU1; OUTPUT;
        X=.; Y=.; OUTPUT;
        X=N; Y=MI1; OUTPUT;
        X=N; Y=QL1; OUTPUT;
        X=.; Y=.; OUTPUT;
        X=N-DELTA; Y=MD1; OUTPUT;
        X=N+DELTA; Y=MD1; OUTPUT;
        X=.; Y=.; OUTPUT;
        PROC PRINT;
        PROC GPLOT;
        AXIS1 LABEL=(F=SIMPLEX H=1 'LENGTH USED')
              ORDER=(3 TO 6 BY 1)
              VALUE=(F=SIMPLEX)
              MINOR=NONE
              OFFSET=(10 PCT,10 PCT);
        AXIS2 LABEL=(F=SIMPLEX H=1 'MAD')
              ORDER=(0 TO 20 BY 5, 80 TO 120 BY 20)
              VALUE=(F=SIMPLEX)
              OFFSET=(1,1);
        AXIS3  LABEL=(F=SIMPLEX H=1 'MAD')
               ORDER=(1 TO 4 BY 1)
               VALUE=(F=SIMPLEX)
               OFFSET=(1,1);
        PLOT Y*X=1 AV1*N=3 / OVERLAY VAXIS=AXIS3
                             HAXIS=AXIS1 FRAME;
        PLOT2 Y*X=2 / VAXIS=AXIS3 HAXIS=AXIS1;
```

DOUBLE MOVING AVERAGE PROGRAM. THE TREND SIMULATIONS OPERATE FROM THE SAME DATA FILES USED IN THE STATIONARY SIMULATIONS. THE STATIONARY SERIES IS ESSENTIALLY SPREAD AROUND A TREND LINE WITH A .5 SLOPE. THE INTERCEPT VALUE OF ZERO IS SELECTED SO THAT THE EXPECTED VALUE OF THE FORTY OBSERVATIONS USED FOR THE MEASURES IS CONSISTENT WITH THE STATIONARY SIMULATION.

```
 1        DATA _NULL_;
 2        INFILE SIXKC7;
 3        INPUT T XI;
 4        IF T = 1 THEN K = 1;
 5        COUNTER + 1;
 6        X = XI -20 +.5*COUNTER;
 7        ARRAY V{18} V1-V18;
 8        ARRAY MA{9} MA1-MA9;
 9        ARRAY DMA{9} DMA1-DMA9;
10        ARRAY FC{9} FC1-FC9;
11        ARRAY SUMS{11,11} S1-S121;
12        ERRTYPE = 2;
13        PHI = .7;
14        LAGX = LAG(X);
15        L1 = LAGX;
16        L2 = LAG2(X);
17        L3 = LAG3(X);
18        L4 = LAG4(X);
19        L5 = LAG5(X);
20        L6 = LAG6(X);
21        L7 = LAG7(X);
22        L8 = LAG8(X);
23        SUM = X;
24        MA1 = SUM;
25        SUM = SUM+L1;
26        MA2 = SUM/2;
27        SUM = SUM+L2;
28        MA3 = SUM/3;
29        SUM = SUM+L3;
30        MA4 = SUM/4;
31        SUM = SUM+L4;
32        MA5 = SUM/5;
33        SUM = SUM+L5;
34        MA6 = SUM/6;
35        SUM = SUM+L6;
36        MA7 = SUM/7;
37        SUM = SUM+L7;
38        MA8 = SUM/8;
39        SUM = SUM+L8;
40        MA9 = SUM/9;
41        SUM = 0;
42        DMA{1} = MA{1};
43        DMA{2} = (MA{2} + LAG(MA{2}))/2;
44        DMA{3} = (MA{3} + LAG(MA{3}) + LAG2(MA{3}))/3;
```

Recall that the stationary series had a level of 20, line six thus results in X values around the line from the origin to the point (60,30).

Moving averages (MA1-MA9) are calculated in this section, double moving averages (DMA{1}-DMA{9}) are calculated below.

```
        DMA{4} = (MA{4} + LAG(MA{4}) + LAG2(MA{4}) +
                  LAG3(MA{4}))/4;
        DMA{5} = (MA{5} + LAG(MA{5}) + LAG2(MA{5}) +
                  LAG3(MA{5}) + LAG4(MA{5}))/5;
        DMA{6} = (MA{6} + LAG(MA{6}) + LAG2(MA{6}) +
                  LAG3(MA{6}) + LAG4(MA{6}) +
                  LAG5(MA{6}))/6;
        DMA{7} = (MA{7} + LAG(MA{7}) + LAG2(MA{7}) +
                  LAG3(MA{7})  + LAG4(MA{7}) +
                  LAG5(MA{7}) + LAG6(MA{7}))/7;
        DMA{8} = (MA{8} + LAG(MA{8}) + LAG2(MA{8}) +
                  LAG3(MA{8}) + LAG4(MA{8}) +
                  LAG5(MA{8}) + LAG6(MA{8}) +
                  LAG7(MA{8}))/8;
        DMA{9} = (MA{9} + LAG(MA{9}) + LAG2(MA{9}) +
                  LAG3(MA{9}) + LAG4(MA{9}) +
                  LAG5(MA{9}) + LAG6(MA{9}) +
                  LAG7(MA{9}) + LAG8(MA{9}))/9;
        DO I=1 TO 9;
          N = I;
          AT = 2*MA{I} - DMA{I};
          IF N = 1 THEN BT=0;
          ELSE BT = (2/(N-1))*(MA{I}-DMA{I});
          FC{I} = AT + BT;
          F = LAG(FC{I});
          E = X-F;
          V1 = T;
          V2 = X;
          V3 = AT;
          V4 = BT;
          V5 = N;
          V6 = F;
          V7 = E;
          IF T GT 9 THEN DO;         Example of checking/
            IF T LT 20 THEN DO;                 debugging.
              FILE CHECK;
              PUT V1 V2 N MA{I} DMA{I} V3 V4 V6 V7;
            END;
          END;
          IF COUNTER LT 21 THEN GO TO CUT;
          V8 = ABS(E);
          V9 = E**2;             Manipulating the error.
          V10 = (V8/X)**2;
          V11 = ABS(E/X);
          V12 = (X-LAGX)**2;
          V13 = ABS(E/LAGX);
          V14 = ABS ((X-LAGX)/LAGX);
          V15 = ABS(E);
          V16 = ABS(X-LAGX);
          V17 = (E/LAGX)**2;
          V18 = ((X-LAGX)/LAGX)**2;
          SUMS{I,1} + V8;                          Summing.
```

```
          SUMS{I,2} + V9;
          SUMS{I,3} + V10;
          SUMS{I,4} + V11;
          SUMS{I,5} + V12;
          SUMS{I,6} + V13;
          SUMS{I,7} + V14;
          SUMS{I,8} + V15;
          SUMS{I,9} + V16;
          SUMS{I,10} + V17;
          SUMS{I,11} + V18;
          IF COUNTER EQ 60 THEN DO;
            MAD = SUMS{I,1}/40;                 Calculating.
            MSE = SUMS{I,2}/40;
            MSPE = SUMS{I,3}/40;
            MAPE = SUMS{I,4}/40;
            TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
            TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
            TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
            TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
            FILE TC7M;                                   Filing.
            PUT ERRTYPE N PHI MAD MSE MAPE MSPE TH1
            TH2 TH3 TH4;
          END;
        CUT:END;

        IF COUNTER = 60 THEN DO;
          COUNTER = 0;
          K+1;
          DO M = 1 TO 11;
            DO N = 1 TO 11;                            Resetting.
              SUMS{M,N} = 0;
            END;
          END;
        END;
```

LINEAR EXPONENTIAL AVERAGE PROGRAM. THIS SAMPLE DEALS WITH THE CAUCHY ERROR DISTRIBUTION WITH PHI = .7. THE SERIES IS IDENTICAL WITH THAT DISCUSSED FOR THE DOUBLE MOVING AVERAGE.

```
DATA _NULL_;
INFILE SIXKC7;
INPUT T XI;
RETAIN;
COUNTER + 1;
X = XI - 20 + .5*COUNTER;
ARRAY V{18} V1-V18;
ARRAY EX{10} EX1-EX10;
ARRAY E2X{10} E2X1-E2X10;
ARRAY FA{10} F1-F10;
ARRAY SUMS{11,11} S1-S121;
ERRTYPE = 2;
PHI = .7;
LAGX = LAG(X);
DO I = 1 TO 10;
  C = (I)*.05;
  IF COUNTER = 1 THEN DO;
    EX0 = -((1-C)/C)*.5;
    E2X0 = 2*EX0;
    EX{I} = (C*X)+(1-C)*EX0;
    E2X{I} = C*EX{I} + (1-C)*E2X0;
    END;
  ELSE DO;
    EX{I} = C*X + (1-C) *EX{I};
    E2X{I} = C*EX{I} + (1-C) * E2X{I};
  END;
  AT = 2*EX{I} - E2X{I};
  BT = (C/(1-C))*(EX{I} - E2X{I});
  FA{I} = AT + BT;
  F = LAG(FA{I});
  E = X-F;
  V1 = T;
  V2 = X;
  V3 = C;
  V4 = AT;
  V5 = BT;
  V6 = F;
  V7 = E;
  IF COUNTER LT 21 THEN GO TO CUT;
  V8 = ABS(E);
  V9 = E**2;
  V10 = (V8/X)**2;
  V11 = ABS(E/X);
  V12 = (X-LAGX)**2;
  V13 = ABS(E/LAGX);
  V14 = ABS ((X-LAGX)/LAGX);
  V15 = ABS(E);
  V16 = ABS(X-LAGX);
```

Since no errors are accumulated for the first 20 observations, the initialization scheme is not critical for the simulation. The single and double exponential averages are initialized to match the known population value at time t = 1.

Calculating the various error term elements.

```
    V17 = (E/LAGX)**2;
    V18 = ((X-LAGX)/LAGX)**2;
    SUMS{I,1} + V8;
    SUMS{I,2} + V9;              Summing the necessary
    SUMS{I,3} + V10;              error term elements.
    SUMS{I,4} + V11;
    SUMS{I,5} + V12;
    SUMS{I,6} + V13;
    SUMS{I,7} + V14;
    SUMS{I,8} + V15;
    SUMS{I,9} + V16;
    SUMS{I,10} + V17;
    SUMS{I,11} + V18;
    IF COUNTER EQ 60 THEN DO;
      MAD = SUMS{I,1}/40;              Calculating the
      MSE = SUMS{I,2}/40;              error measures.
      MSPE = SUMS{I,3}/40;
      MAPE = SUMS{I,4}/40;
      TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
      TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
      TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
      TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
      FILE TC7E;
      PUT T ERRTYPE C PHI MAD MSE MAPE MSPE TH1
          TH2 TH3 TH4;
    END;
  CUT:END;

  IF COUNTER = 60 THEN DO;                   Resetting.
    COUNTER = 0;
    DO M = 1 TO 11;
      DO N = 1 TO 11;
        SUMS{M,N} = 0;
      END;
    END;
  END;
```

DOUBLE RUNNING MEDIAN PROGRAM - FIRST HALF. DOUBLE RUNNING MEDIANS ARE MUCH LIKE DOUBLE MOVING AVERAGES, EXCEPT THEY ARE MUCH LESS SENSITIVE TO OUTLIERS. THIS FIRST HALF OF THE PROGRAM CALCULATES THE SINGLE RUNNING MEDIANS AND STORES THEM IN A DATA SET FOR LATER USE.

```
DATA _NULL_;
INFILE SIXKC7;
INPUT T XI;
IF T=1 THEN DO;
     K=1;
END;
ARRAY V{18} V1-V18;
ERRTYPE = 2;
PHI = .7;
COUNTER + 1;
XI = XI - 20 + .5*COUNTER;
X = XI;
LAGX = LAG(X);
L1=LAGX;                                Renaming the lags
L2=LAG2(X);                             makes programming
L3=LAG3(X);                        easier and precludes the
L4=LAG4(X);                          jumbling problems when
L5=LAG5(X);                                        sorting.
IF COUNTER LT 9 THEN GO TO SKIP;                  Skipping
ARRAY TEMP3{3} X L1 L2;                           unneeded
ARRAY TEMP4{4} X L1 L2 L3;                   calculations.
ARRAY TEMP5{5} X L1 L2 L3 L4;
ARRAY TEMP6{6} X L1 L2 L3 L4 L5;
DO H=1 TO 3;
MORE=0;
DO I = 1 TO 2;
     IF TEMP3{I+1} LT TEMP3{I} THEN DO;
          TEMP = TEMP3{I};
          TEMP3{I} = TEMP3{I+1};
          TEMP3{I+1} = TEMP;
          MORE = 1;
          END;
     END;
     IF MORE = 0 THEN GO TO DONE3;
END;                               If all values are sorted
DONE3:;                         properly, then MORE remains
RM3 = TEMP3{2};                    equal to zero and there
DO H = 1 TO 4;                     is no reason to continue
MORE = 0;                                   in the do loop.
DO I = 1 TO 3;
  IF TEMP4{I+1} LT TEMP4{I} THEN DO;
    TEMP = TEMP4{I};
    TEMP4{I} = TEMP4{I+1};
    TEMP4{I+1} = TEMP;
    MORE = 1;
    END;
```

```
      END;
      IF MORE = 0 THEN GO TO DONE4;
    END;
    DONE4:;
    RM4 = (TEMP4{2} + TEMP4{3})/2;  Median of length
    DO H = 1 TO 5;                             four.
    MORE = 0;
      DO I = 1 TO 4;
        IF TEMP5{I+1} LT TEMP5{I} THEN DO;
        TEMP = TEMP5{I};
        TEMP5{I} = TEMP5{I+1};
        TEMP5{I+1} = TEMP;
        MORE = 1;
        END;
      END;
      IF MORE = 0 THEN GO TO DONE5;
    END;
    DONE5:;
    RM5 = TEMP5{3};           Median of length five.
    DO H = 1 TO 6;
    MORE = 0;
      DO I = 1 TO 5;
        IF TEMP6{I+1} LT TEMP6{I} THEN DO;
        TEMP = TEMP6{I};
        TEMP6{I} = TEMP6{I+1};
        TEMP6{I+1} = TEMP;
        MORE = 1;
        END;
      END;
      IF MORE = 0 THEN GO TO DONE6;
    END;
    DONE6:;                         Median of length
    RM6 = (TEMP6{3} + TEMP6{4})/2;              six.
    SKIP:;
    ARRAY RM{4} RM3 RM4 RM5 RM6;
    FILE MEDSC7;                             Filing.
    PUT T XI RM3 RM4 RM5 RM6;
    IF COUNTER = 60 THEN COUNTER =0;      Resetting.
```

DOUBLE RUNNING MEDIAN PROGRAM - SECOND HALF. HERE THE SINGLE RUNNING MEDIANS ARE READ, THE DOUBLE MOVING MEDIANS ARE CALCULATED, AND ONE PERIOD AHEAD FORECASTS ARE EVALUATED FOR FIT USING EACH OF THE FOUR LENGTHS OF MEDIANS CONSIDERED.

```
    DATA _NULL_;
    INFILE MEDSC7;
    INPUT T XI RMI3 RMI4 RMI5 RMI6;
    RM3 = RMI3;                     Renamed again to
    RM4 = RMI4;                       avoid jumbling
    RM5 = RMI5;                      during sorting.
```

```
RM6 = RMI6;
COUNTER + 1;
ARRAY RM{4} RMI3 RMI4 RMI5 RMI6;
ARRAY FC{4} FC1-FC4;
ERRTYPE = 2;
PHI =.7;
LT1 = LAG(RM3);            The three period double
LT2 = LAG2(RM3);       running median is the median
LF1 = LAG(RM4);      of the last three three-period
LF2 = LAG2(RM4);          (single) running medians.
LF3 = LAG3(RM4);            In this section all the
LFI1 = LAG(RM5);            appropriate lags of the
LFI2 = LAG2(RM5);           medians are defined for
LFI3 = LAG3(RM5);             the following arrays.
LFI4 = LAG4(RM5);         LT - lags for Three period
LS1 = LAG(RM6);           LF - lags for Four period
LS2 = LAG2(RM6);          LFI - lags for FIve period
LS3 = LAG3(RM6);          LS  - lags for Six period
LS4 = LAG4(RM6);
LS5 = LAG5(RM6);
ARRAY SC3{3} RM3 LT1 LT2;
ARRAY SC4{4} RM4 LF1-LF3;
ARRAY SC5{5} RM5 LFI1-LFI4;
ARRAY SC6{6} RM6 LS1-LS5;
DO H = 1 TO 3;
MORE = 0;
DO I = 1 TO 2;
  IF SC3{I+1} LT SC3{I} THEN DO;
    TEMP = SC3{I};
    SC3{I} = SC3{I+1};
    SC3{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE3;
END;
DONE3:;
DRM3 = SC3{2};           Three period double running
DO H = 1 TO 4;                                median.
MORE = 0;
DO I = 1 TO 3;
  IF SC4{I+1} LT SC4{I} THEN DO;
    TEMP = SC4{I};
    SC4{I} = SC4{I+1};
    SC4{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE4;
END;
DONE4:;
DRM4 = (SC4{2} + SC4{3})/2;     Four period DRM.
```

```
DO H = 1 TO 5;
MORE = 0;
DO I = 1 TO 4;
  IF SC5{I+1} LT SC5{I} THEN DO;
    TEMP = SC5{I};
    SC5{I} = SC5{I+1};
    SC5{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE5;
END;
DONE5:;
DRM5 = SC5{3};                    Five period DRM.
DO H = 1 TO 6;
MORE = 0;
DO I = 1 TO 5;
  IF SC6{I+1} LT SC6{I} THEN DO;
    TEMP = SC6{I};
    SC6{I} = SC6{I+1};
    SC6{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE6;
END;
DONE6:;
DRM6 = (SC6{3} + SC6{4})/2;       Six period DRM.
ARRAY DRM{4} DRM3 DRM4 DRM5 DRM6;
ARRAY SUMS{4,11} S1-S44;
LAGXI = LAG(XI);
IF T LE 25 THEN DO;     An example of a checking/
  FILE CHECKDRM;                  debugging routine.
  PUT T COUNTER XI RMI3 RMI4 RMI5 RMI6 DRM3 DRM4
      DRM5 DRM6;
END;
DO I=1 TO 4;
  N = I+2;                   N is the length of the
                              running medians used.
  AT = 2*RM{I} - DRM{I};
  BT = (2/(N-1))*(RM{I} -DRM{I});
  FC{I} = AT +BT;
  F = LAG(FC{I});
  N = I+2;
  E=XI-F;
  V1 = T;
  V2 = XI;
  V3 = AT;
  V4 = BT;
  V5 = N;
  V6 = F;
  V7 = E;
```

```
          IF COUNTER LT 21 THEN GO TO CUT;
          V8 = ABS(E);                              Manipulating.
          V9 = E**2;
          V10 = (V8/XI)**2;
          V11 = ABS(E/XI);
          V12 = (XI-LAGXI)**2;
          V13 = ABS(E/LAGXI);
          V14 = ABS ((XI-LAGXI)/LAGXI);
          V15 = ABS(E);
          V16 = ABS(XI-LAGXI);
          V17 = (E/LAGXI)**2;
          V18 = ((XI-LAGXI)/LAGXI)**2;
          SUMS{I,1} + V8;                              Summing.
          SUMS{I,2} + V9;
          SUMS{I,3} + V10;
          SUMS{I,4} + V11;
          SUMS{I,5} + V12;
          SUMS{I,6} + V13;
          SUMS{I,7} + V14;
          SUMS{I,8} + V15;
          SUMS{I,9} + V16;
          SUMS{I,10} + V17;
          SUMS{I,11} + V18;
          IF COUNTER EQ 60 THEN DO;
            MAD = SUMS{I,1}/40;                    Calculating.
            MSE = SUMS{I,2}/40;
            MSPE = SUMS{I,3}/40;
            MAPE = SUMS{I,4}/40;
            TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
            TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
            TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
            TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
            FILE TC7R;                                  Filing.
            PUT ERRTYPE N PHI MAD MSE MAPE MSPE TH1 TH2
                TH3 TH4;
          END;
        CUT:END;

        IF COUNTER = 60 THEN DO;                     Resetting.
          COUNTER = 0;
          K+1;
          DO M = 1 TO 4;
            DO N = 1 TO 11;
              SUMS{M,N} = 0;
            END;
          END;
        END;
```

DOUBLE SMOOTHED MEDIAN PROGRAM. HERE RUNNING MEDIANS ARE DEVELOPED AND USED TO CALCULATE SMOOTHED AND DOUBLE SMOOTHED MEDIANS. THESE ARE THEN USED IN THE DOUBLE SMOOTHED MEDIAN TECHNIQUE.

```
 1      DATA _NULL_;
 2      INFILE SIXKC7;
 3      INPUT T XI;
 4      RETAIN;
 5      COUNTER + 1;
 6      XI = XI - 20 + .5*COUNTER;
 7      ARRAY V{18} V1-V18;
 8      ARRAY SUMS{4,6,11} S1-S264;
 9      ERRTYPE = 2;
10      PHI = .7;
11      X = XI;
12      LAGX = LAG(X);
13      L1=LAGX;
14      L2=LAG2(X);
15      L3=LAG3(X);
16      L4=LAG4(X);
17      L5=LAG5(X);
18      IF COUNTER LE 5 THEN GO TO SKIP;
19      ARRAY TEMP3{3} X L1 L2;
20      ARRAY TEMP4{4} X L1 L2 L3;
21      ARRAY TEMP5{5} X L1 L2 L3 L4;
22      ARRAY TEMP6{6} X L1 L2 L3 L4 L5;
23      DO H=1 TO 3;
24      MORE=0;
25      DO I = 1 TO 2;
26           IF TEMP3{I+1} LT TEMP3{I} THEN DO;
27                TEMP = TEMP3{I};
28                TEMP3{I} = TEMP3{I+1};
29                TEMP3{I+1} = TEMP;
30                MORE = 1;
31                END;
32           END;
33           IF MORE = 0 THEN GO TO DONE3;
34      END;
35      DONE3:;
36      RM3 = TEMP3{2};
37      DO H = 1 TO 4;
38      MORE = 0;
39      DO I = 1 TO 3;
40        IF TEMP4{I+1} LT TEMP4{I} THEN DO;
41          TEMP = TEMP4{I};
42          TEMP4{I} = TEMP4{I+1};
43          TEMP4{I+1} = TEMP;
44          MORE = 1;
45          END;
46        END;
47        IF MORE = 0 THEN GO TO DONE4;
```

For the simulation all smoothed medians are initialized at time counter equals six, the first period the longest median considered exists (see line 18).

```
END;
DONE4:;
RM4 = (TEMP4{2} + TEMP4{3})/2;
DO H = 1 TO 5;
MORE = 0;
  DO I = 1 TO 4;
    IF TEMP5{I+1} LT TEMP5{I} THEN DO;
    TEMP = TEMP5{I};
    TEMP5{I} = TEMP5{I+1};
    TEMP5{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE5;
END;
DONE5:;
RM5 = TEMP5{3};
DO H = 1 TO 6;
MORE = 0;
  DO I = 1 TO 5;
    IF TEMP6{I+1} LT TEMP6{I} THEN DO;
    TEMP = TEMP6{I};
    TEMP6{I} = TEMP6{I+1};
    TEMP6{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE6;
END;
DONE6:;
RM6 = (TEMP6{3} + TEMP6{4})/2;
SKIP:;
ARRAY RM{4} RM3 RM4 RM5 RM6;
ARRAY SM{4,6} SM1-SM24;
ARRAY DSM{4,6} DSM1-DSM24;
ARRAY FI{4,6} F1-F24;
LAGXI = LAG(XI);           For the trend simulation
DO I=1 TO 4;               smoothing constants were
  N = I+2;                    varied from .1 to .6.
  DO J = 1 TO 6;
  C = .1*J;
  IF COUNTER EQ 6 THEN DO;          Initializing.
    SM{I,J} = 3 - ((N-1)/4) -.5*((1-C)/C);
    DSM{I,J} = 3 - ((N-1)/4) - ((1-C)/C);
    END;
  ELSE DO;                       Once Initialized.
    SM{I,J} = C*RM{I} + (1-C)*SM{I,J};
    DSM{I,J} = C*SM{I,J} + (1-C)*DSM{I,J};
  END;
  AT = 2*SM{I,J} - DSM{I,J};
  BT = (C/(1-C))*(SM{I,J} - DSM{I,J});
  FI{I,J} = AT + ((N+1)/2)*BT;
```

```
    F = LAG(FI{I,J});
    E=XI-F;
    V1 = T;
    V2 = XI;
    V3 = N;
    V4 = C;
    V5 = RM{I};
    V6 = F;
    V7 = E;
    IF COUNTER LT 21 THEN GO TO CUT;
    V8 = ABS(E);                               Manipulating.
    V9 = E**2;
    V10 = (V8/XI)**2;
    V11 = ABS(E/XI);
    V12 = (XI-LAGXI)**2;
    V13 = ABS(E/LAGXI);
    V14 = ABS ((XI-LAGXI)/LAGXI);
    V15 = ABS(E);
    V16 = ABS(XI-LAGXI);
    V17 = (E/LAGXI)**2;
    V18 = ((XI-LAGXI)/LAGXI)**2;
    SUMS{I,J,1} + V8;                             Summing.
    SUMS{I,J,2} + V9;
    SUMS{I,J,3} + V10;
    SUMS{I,J,4} + V11;
    SUMS{I,J,5} + V12;
    SUMS{I,J,6} + V13;
    SUMS{I,J,7} + V14;
    SUMS{I,J,8} + V15;
    SUMS{I,J,9} + V16;
    SUMS{I,J,10} + V17;
    SUMS{I,J,11} + V18;
    IF COUNTER EQ 60 THEN DO;
      MAD = SUMS{I,J,1}/40;                   Calculating.
      MSE = SUMS{I,J,2}/40;
      MSPE = SUMS{I,J,3}/40;
      MAPE = SUMS{I,J,4}/40;
      TH1 = SQRT(SUMS{I,J,2}/SUMS{I,J,5});
      TH2 = SQRT(SUMS{I,J,8}/SUMS{I,J,9});
      TH3 = SQRT(SUMS{I,J,10}/SUMS{I,J,11});
      TH4 = SQRT(SUMS{I,J,6}/SUMS{I,J,7});
      FILE TC7DSM;                                 Filing.
      PUT ERRTYPE N C PHI MAD MSE MAPE MSPE TH1
          TH2 TH3 TH4;
    END;
  CUT:END;
  END;

  IF COUNTER = 60 THEN DO;                      Resetting.
    COUNTER = 0;
    K+1;
    DO M = 1 TO 4;
```

```
          DO P = 1 TO 6;
           DO N = 1 TO 11;
             SUMS{M,P,N} = 0;
           END;
          END;
         END;
       END;
```

ROBUST REGRESSION PROGRAM - FIRST HALF. THIS PROGRAM PERFORMS A ROBUST REGRESSION SIMILAR TO THAT DESCRIBED BY HOLLANDER AND WOLFE (1973). THE DIFFICULTY IS THAT FOUR REGRESSIONS ARE CONSIDERED FOR EACH ADVANCED OBSERVATION, ONE USING THE LAST THREE OBSERVATIONS, ONE USING THE LAST FOUR, ETC. EACH REGRESSION DETERMINES A LINE, WITH A SLOPE AND A POINT. EACH OF THESE FOUR LINES IS USED TO ESTIMATE THE NEXT OBSERVATION'S VALUE, GENERATING FITTING ERRORS. THE FIRST HALF OF THE PROGRAM CALCULATES THE PARAMETERS OF THE LINES, USES THE LINES TO ESTIMATE ONE PERIOD AHEAD, AND FILES THE RESULTS FOR THE SECOND HALF OF THE PROGRAM.

```
DATA _NULL_;
INFILE SIXKC7;
INPUT T XI;
COUNTER + 1;
XI = XI - 20 + .5*COUNTER;
X = XI;                                         Renaming.
L1 = LAG(X);
L2 = LAG2(X);
L3 = LAG3(X);                   Since no initialization
L4 = LAG4(X);               is required, no computations
L5 = LAG5(X);                are performed until counter
IF COUNTER LT 15 THEN GO TO SHORT;              is equal
ARRAY LAGS{6} L5 L4 L3 L2 L1 X;             to fifteen.
ARRAY SLOPE{15} S1-S15;
L = 0;
DO I = 1 TO 5;                      Generating the set of
  DO J = I+1 TO 6;                        pairwise slopes.
    L+1;
    SLOPE{L} = (LAGS{J} - LAGS{I})/(J-I);
  END;
END;
ARRAY SL3{3} S13 S14 S15;              Finding the median
DO I = 1 TO 3;                         of the slopes from
REDO = 0;                             the last three pts.
  DO J = 1 TO 2;
    IF SL3{J+1} LT SL3{J} THEN DO;
      TEMP = SL3{J};
      SL3{J} = SL3{J+1};
      SL3{J+1} = TEMP;
      REDO = 1;
    END;
  END;
  IF REDO = 0 THEN GO TO DONESL3;
END;
DONESL3:;                      The slope statistic from
SLOPE3 = SL3{2};                   use of the last three
ARRAY SL4{6} S10-S15;                            points.
DO I = 1 TO 6;
REDO = 0;
  DO J = 1 TO 5;
```

```
          IF SL4{J+1} LT SL4{J} THEN DO;
            TEMP = SL4{J};
            SL4{J} = SL4{J+1};
            SL4{J+1} = TEMP;
            REDO = 1;
          END;
        END;
        IF REDO = 0 THEN GO TO DONESL4;
      END;
      DONESL4:;                     The slope statistic
      SLOPE4 = (SL4{3} + SL4{4})/2;    using four pts.
      ARRAY SL5{10} S6-S15;
      DO I = 1 TO 10;
      REDO = 0;
        DO J = 1 TO 9;
          IF SL5{J+1} LT SL5{J} THEN DO;
            TEMP = SL5{J};
            SL5{J} = SL5{J+1};
            SL5{J+1} = TEMP;
            REDO = 1;
          END;
        END;
        IF REDO = 0 THEN GO TO DONESL5;
      END;
      DONESL5:;                     The slope statistic
      SLOPE5 = (SL5{5} + SL5{6})/2;    using five pts.
      DO I = 1 TO 15;
        DO J = 1 TO 14;
          IF SLOPE{J+1} LT SLOPE{J} THEN DO;
            TEMP = SLOPE{J};
            SLOPE{J} = SLOPE{J+1};
            SLOPE{J+1} = TEMP;
            REDO = 1;
          END;
        END;
        IF REDO = 0 THEN GO TO DONESL6;
      END;
      DONESL6:;
      SLOPE6 = (SLOPE{8});          The slope statistic
      ARRAY B{4} SLOPE3-SLOPE6;        using six pts.
      ARRAY TEMP3{3} L2 L1 X;
      DO I = 1 TO 3;
        REDO = 0;
        DO J = 1 TO 2;
          IF TEMP3{J+1} LT TEMP3{J} THEN DO;
            TEMP = TEMP3{J};
            TEMP3{J} = TEMP3{J+1};          Finding the
            TEMP3{J+1} = TEMP;           y-value of the
            REDO = 1;                point, the median of
          END;              the last three observations.
        END;
        IF REDO = 0 THEN GO TO DONE3;
```

```
END;
DONE3:;                       A3 is the y-value for the
A3 = TEMP3{2};                            three period
ARRAY TEMP4{4} L3 L2 L1 X;    robust regression.
DO I = 1 TO 4;
  REDO = 0;
  DO J = 1 TO 3;
    IF TEMP4{J+1} LT TEMP4{J} THEN DO;
      TEMP = TEMP4{J};
      TEMP4{J} = TEMP4{J+1};
      TEMP4{J+1} = TEMP;
      REDO = 1;
    END;
  END;
  IF REDO = 0 THEN GO TO DONE4;
END;
DONE4:;
A4 = (TEMP4{2} + TEMP4{3})/2;      Point for four
ARRAY TEMP5{5} L4 L3 L2 L1 X;            periods.
DO I = 1 TO 5;
  REDO = 0;
  DO J = 1 TO 4;
    IF TEMP5{J+1} LT TEMP5{J} THEN DO;
      TEMP = TEMP5{J};
      TEMP5{J} = TEMP5{J+1};
      TEMP5{J+1} = TEMP;
      REDO = 1;
    END;
    IF REDO = 0 THEN GO TO DONE5;
  END;
  IF REDO = 0 THEN GO TO DONE5;
END;
DONE5:;                    Point for five periods.
A5 = TEMP5{3};
ARRAY TEMP6{6} L5 L4 L3 L2 L1 X;
DO I = 1 TO 6;
  REDO = 0;
  DO J = 1 TO 5;
    IF TEMP6{J+1} LT TEMP6{J} THEN DO;
      TEMP = TEMP6{J};
      TEMP6{J} = TEMP6{J+1};
      TEMP6{J+1} = TEMP;
      REDO = 1;
    END;
    IF REDO = 0 THEN GO TO DONE6;
  END;
  IF REDO = 0 THEN GO TO DONE6;
END;
DONE6:;                              Point for six
A6 = (TEMP6{3} + TEMP6{4})/2;             periods.
ARRAY A{4} A3-A6;
ARRAY FI{4} FI1-FI4;
```

```
DO I = 1 TO 4;
  H = I+3;                                   Forecasts for each
  FI{I} = A{I} + (H/2)*B{I};                      observation.
END;
SHORT:;
FILE RRC7;                                             Filing.
PUT T XI FI1-FI4;
IF COUNTER = 60 THEN COUNTER = 0;
```

ROBUST REGRESSION PROGRAM - SECOND HALF. THIS HALF OF THE PROGRAM READS IN THE ESTIMATES AND CALCULATES THE ERROR MEASURES FOR EACH MODEL.

```
DATA _NULL_;
INFILE RRC7;                                File created by the
INPUT T XI FI1-FI4;                           first half of the
COUNTER + 1;                                           program.
ARRAY FI{4} FI1-FI4;
ARRAY SUMS{4,11} S1-S44;
ERRTYPE = 2;
PHI = .7;
LAGXI = LAG(XI);
DO I = 1 TO 4;
  N = I+2;
  F = LAG(FI{I});
  E = XI-F;
  V1 = T;
  V2 = XI;
  V3 = AT;
  V4 = BT;
  V5 = N;
  V6 = F;
  V7 = E;
  IF COUNTER LT 21 THEN GO TO CUT;
  V8 = ABS(E);
  V9 = E**2;                                      Manipulating.
  V10 = (V8/XI)**2;
  V11 = ABS(E/XI);
  V12 = (XI-LAGXI)**2;
  V13 = ABS(E/LAGXI);
  V14 = ABS ((XI-LAGXI)/LAGXI);
  V15 = ABS(E);
  V16 = ABS(XI-LAGXI);
  V17 = (E/LAGXI)**2;
  V18 = ((XI-LAGXI)/LAGXI)**2;
  SUMS{I,1} + V8;
  SUMS{I,2} + V9;                                     Summing.
  SUMS{I,3} + V10;
  SUMS{I,4} + V11;
  SUMS{I,5} + V12;
  SUMS{I,6} + V13;
```

```
    SUMS{I,7} + V14;
    SUMS{I,8} + V15;
    SUMS{I,9} + V16;
    SUMS{I,10} + V17;
    SUMS{I,11} + V18;
    IF COUNTER EQ 60 THEN DO;
      MAD = SUMS{I,1}/40;                         Calculating.
      MSE = SUMS{I,2}/40;
      MSPE = SUMS{I,3}/40;
      MAPE = SUMS{I,4}/40;
      TH1 = SQRT(SUMS{I,2}/SUMS{I,5});
      TH2 = SQRT(SUMS{I,8}/SUMS{I,9});
      TH3 = SQRT(SUMS{I,10}/SUMS{I,11});
      TH4 = SQRT(SUMS{I,6}/SUMS{I,7});
      FILE TC7REG;                                     Filing.
      PUT ERRTYPE N PHI MAD MSE MAPE MSPE TH1
          TH2 TH3 TH4;
    END;
  CUT:END;

  IF COUNTER = 60 THEN DO;                          Resetting.
    COUNTER = 0;
    K+1;
    DO M = 1 TO 4;
      DO N = 1 TO 11;
        SUMS{M,N} = 0;
      END;
    END;
  END;
```

SAMPLE PLOT PROGRAMS. THIS FIRST PROGRAM READS A SAS DATA SET CREATED THROUGH PROC UNIVARIATE. IT THEN PLOTS THE AVERAGE (AV), UPPER QUARTILE (QU), LOWER QUARTILE (QL), MEDIAN (MD), MAXIMUM (MX), AND MINIMUM (MI) VALUES OVER THE TWENTY REPLICATIONS.

```
LIBNAME STAT '[RKANKEY.TREND]';
DATA DISPLAY;
SET STAT.AR3NDM;
```

All trend simulations files are in a subdirectory. This data set has results from the autoregressive simulation with Phi = .3 and normal errors when double moving averages was used as the technique.

```
PROC PLOT;
PLOT AV1*N='A' QU1*N='U' QL1*N='L' MD1*N='M'
     MX1*N='G' MI1*N='S' / OVERLAY;
PLOT AV2*N='A' QU2*N='U' QL2*N='L' MD2*N='M'
     MX2*N='G' MI2*N='S' / OVERLAY;
PLOT AV3*N='A' QU3*N='U' QL3*N='L' MD3*N='M'
     MX3*N='G' MI3*N='S' / OVERLAY;
PLOT AV4*N='A' QU4*N='U' QL4*N='L' MD4*N='M'
     MX4*N='G' MI4*N='S' / OVERLAY;
PLOT AV5*N='A' QU5*N='U' QL5*N='L' MD5*N='M'
     MX5*N='G' MI5*N='S' / OVERLAY;
PLOT AV6*N='A' QU6*N='U' QL6*N='L' MD6*N='M'
     MX6*N='G' MI6*N='S' / OVERLAY;
PLOT AV7*N='A' QU7*N='U' QL7*N='L' MD7*N='M'
     MX7*N='G' MI7*N='S' / OVERLAY;
PLOT AV8*N='A' QU8*N='U' QL8*N='L' MD8*N='M'
     MX8*N='G' MI8*N='S' / OVERLAY;
```

APPENDIX C

DETAILED SIMULATION RESULTS

Table 23

Mean MAD Ratios To Naive One MAD
for Stationary Series, Normal Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.86	.99	1.20
	.2	.90	.90	.90	.90	1.04	1.05	1.06	1.06	1.24	1.28	1.31	1.32	.86	.97	1.13
SMOOTHED	.3	.92	.92	.92	.91	1.05	1.07	1.07	1.07	1.23	1.27	1.30	1.32	.87	.96	1.08
MEDIAN	.4	.93	.93	.93	.92	1.06	1.07	1.08	1.08	1.21	1.25	1.29	1.31	.88	.95	1.05
	.5	.94	.93	.94	.93	1.06	1.08	1.09	1.08	1.20	1.24	1.28	1.30	.89	.95	1.02
	.6	.95	.94	.94	.93	1.06	1.07	1.09	1.08	1.19	1.23	1.27	1.28			
WALSH AVERAGE		.94	.94	.93	.92	1.03	1.05	1.05	1.05	1.12	1.18	1.21	1.22			
RUNNING MEDIAN		.98	.95	.95	.93	1.06	1.06	1.08	1.07	1.15	1.19	1.24	1.25			
MOVING AVERAGE		.94	.93	.92	.91	1.02	1.04	1.04	1.04	1.11	1.16	1.20	1.21			

Table 24

Mean MSE Ratios To Naive One MSE
for Stationary Series, Normal Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.73	.97	1.43
	.2	.80	.80	.81	.80	1.08	1.10	1.12	1.13	1.55	1.64	1.71	1.76	.74	.94	1.28
SMOOTHED	.3	.83	.83	.84	.82	1.10	1.12	1.15	1.15	1.51	1.61	1.70	1.75	.75	.91	1.17
MEDIAN	.4	.86	.85	.86	.84	1.11	1.14	1.17	1.16	1.48	1.58	1.68	1.73	.77	.90	1.09
	.5	.88	.87	.87	.85	1.12	1.15	1.17	1.16	1.45	1.55	1.66	1.70	.79	.90	1.03
	.6	.90	.88	.88	.86	1.13	1.15	1.18	1.16	1.42	1.52	1.63	1.67			
WALSH AVERAGE		.88	.88	.85	.82	1.05	1.10	1.09	1.09	1.25	1.38	1.45	1.50			
RUNNING MEDIAN		.95	.90	.90	.86	1.14	1.13	1.16	1.13	1.34	1.41	1.55	1.57			
MOVING AVERAGE		.87	.86	.84	.82	1.04	1.07	1.08	1.07	1.24	1.35	1.42	1.47			

Table 25

Mean MAPE Ratios To Naive One MAPE
for Stationary Series, Normal Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.87	1.01	1.23
	.2	.91	.91	.91	.91	1.06	1.07	1.08	1.08	1.27	1.31	1.34	1.36	.88	.99	1.15
SMOOTHED	.3	.93	.93	.93	.92	1.07	1.08	1.09	1.09	1.25	1.29	1.33	1.35	.88	.98	1.10
MEDIAN	.4	.94	.94	.94	.93	1.07	1.09	1.10	1.10	1.23	1.28	1.32	1.34	.89	.97	1.06
	.5	.95	.95	.95	.94	1.07	1.09	1.10	1.10	1.22	1.27	1.31	1.32	.89	.96	1.03
	.6	.96	.95	.95	.94	1.07	1.09	1.11	1.10	1.20	1.25	1.30	1.31			
WALSH AVERAGE		.95	.94	.94	.93	1.04	1.06	1.07	1.07	1.13	1.19	1.22	1.24			
RUNNING MEDIAN		.98	.95	.96	.94	1.07	1.07	1.10	1.09	1.16	1.20	1.26	1.27			
MOVING AVERAGE		.94	.94	.93	.92	1.03	1.05	1.06	1.06	1.12	1.18	1.21	1.24			

Table 26

Mean Theil's U Values
for Stationary Series, Normal Errors

Technique	C	Φ=.3 n=3	Φ=.3 n=4	Φ=.3 n=5	Φ=.3 n=6	Φ=.5 n=3	Φ=.5 n=4	Φ=.5 n=5	Φ=.5 n=6	Φ=.7 n=3	Φ=.7 n=4	Φ=.7 n=5	Φ=.7 n=6	Exponential Smoothing Φ .3	Exponential Smoothing Φ .5	Exponential Smoothing Φ .7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.84	.97	1.18
	.2	.89	.89	.89	.89	1.02	1.04	1.05	1.05	1.23	1.26	1.29	1.31	.85	.95	1.11
SMOOTHED	.3	.90	.90	.90	.90	1.03	1.05	1.06	1.06	1.21	1.25	1.28	1.31	.85	.94	1.05
MEDIAN	.4	.92	.91	.92	.91	1.04	1.05	1.07	1.06	1.19	1.23	1.27	1.29	.86	.93	1.02
	.5	.93	.92	.92	.91	1.04	1.05	1.07	1.06	1.17	1.22	1.26	1.28	.88	.93	.99
	.6	.94	.93	.93	.91	1.05	1.05	1.07	1.06	1.16	1.21	1.25	1.27			
WALSH AVERAGE		.93	.93	.91	.90	1.01	1.03	1.03	1.03	1.09	1.14	1.17	1.20			
RUNNING MEDIAN		.97	.94	.94	.92	1.05	1.04	1.06	1.05	1.13	1.16	1.22	1.23			
MOVING AVERAGE		.92	.91	.90	.89	1.00	1.00	1.02	1.02	1.08	1.13	1.16	1.18			

Table 27

Mean MAD Ratios to Naive One MAD
for Stationary Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING Φ		
		n				n				n						
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.92	1.16	1.62
	.2	.86	.85	.80	.79	1.18	1.18	1.15	1.13	1.66	1.74	1.78	1.80	.91	1.09	1.43
SMOOTHED	.3	.87	.86	.81	.80	1.18	1.19	1.15	1.14	1.59	1.69	1.74	1.77	.89	1.05	1.29
MEDIAN	.4	.89	.87	.81	.80	1.17	1.19	1.15	1.14	1.55	1.64	1.72	1.75	.91	1.01	1.18
	.5	.90	.87	.82	.80	1.15	1.19	1.15	1.14	1.51	1.61	1.69	1.73	.93	.96	1.12
	.6	.90	.87	.82	.80	1.16	1.18	1.15	1.14	1.47	1.59	1.67	1.71			
WALSH AVERAGE		.94	.94	.89	.90	1.12	1.21	1.17	1.16	1.32	1.48	1.54	1.57			
RUNNING MEDIAN		.93	.88	.82	.92	1.17	1.12	1.15	1.11	1.43	1.49	1.64	1.64			
MOVING AVERAGE		.97	.99	.90	1.01	1.11	1.16	1.19	1.21	1.28	1.40	1.50	1.56			

Table 28

Mean MSE Ratios To Naive One MSE
for Stationary Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.73	.97	1.42
	.2	.74	.74	.72	.72	1.08	1.07	1.04	1.04	1.60	1.68	1.73	1.74	.74	.94	1.29
	.3	.75	.75	.73	.72	1.07	1.08	1.08	1.05	1.59	1.68	1.76	1.77	.75	.92	1.18
SMOOTHED	.4	.76	.75	.73	.72	1.08	1.09	1.08	1.06	1.57	1.67	1.77	1.79	.77	.90	1.10
MEDIAN	.5	.76	.75	.73	.73	1.09	1.10	1.08	1.06	1.55	1.66	1.78	1.79	.79	.89	1.04
	.6	.77	.76	.73	.73	1.10	1.10	1.09	1.07	1.54	1.64	1.78	1.79			
WALSH AVERAGE		.82	.79	.77	.77	1.03	1.16	1.10	1.10	1.28	1.58	1.62	1.66			
RUNNING MEDIAN		.79	.76	.73	.73	1.10	1.09	1.09	1.07	1.52	1.57	1.76	1.76			
MOVING AVERAGE		.86	.84	.84	.83	1.03	1.09	1.10	1.23	1.23	1.35	1.46	1.53			

Table 29

Mean MAPE Ratios To Naive One MAPE
for Stationary Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.80	.58	2.11
	.2	.77	.77	.72	.72	.58	.68	.73	.78	2.01	2.14	2.26	2.35	.82	.60	1.74
SMOOTHED	.3	.79	.78	.73	.72	.46	.55	.63	.70	1.72	1.89	2.06	2.19	.80	.88	1.40
MEDIAN	.4	.80	.78	.73	.73	.62	.45	.55	.64	1.61	1.67	1.89	2.03	.84	1.05	1.30
	.5	.82	.79	.73	.73	.75	.56	.49	.59	1.62	1.72	1.76	1.91	.88	1.16	1.24
	.6	.83	.79	.73	.72	.86	.66	.43	.55	1.59	1.74	1.69	1.81			
WALSH AVERAGE		.87	.87	.79	.81	1.20	1.08	.70	.54	1.42	1.70	1.64	1.64			
RUNNING MEDIAN		.86	.81	.73	.72	1.11	.90	.58	.41	1.42	1.67	1.70	1.70			
MOVING AVERAGE		.91	.91	.89	.93	1.24	1.18	.93	.76	1.43	1.67	1.66	1.66			

Table 30

Mean Theil's U Values
for Stationary Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.93	1.23	2.07
	.2	.92	.90	.85	.83	1.31	1.30	1.23	1.22	2.06	2.17	2.21	2.29	.99	1.26	1.77
SMOOTHED	.3	.94	.91	.86	.84	1.35	1.33	1.24	1.24	1.85	1.99	2.07	2.17	1.01	1.23	1.50
MEDIAN	.4	.96	.94	.87	.85	1.36	1.34	1.25	1.24	1.73	1.87	1.94	2.06	1.02	1.18	1.32
	.5	.97	.95	.88	.85	1.36	1.34	1.25	1.24	1.63	1.78	1.86	1.97	1.01	1.13	1.19
	.6	.98	.96	.89	.85	1.36	1.32	1.25	1.23	1.56	1.71	1.80	1.91			
WALSH AVERAGE		1.09	1.01	1.01	.98	1.33	1.35	1.32	1.31	1.37	1.56	1.70	1.78			
RUNNING MEDIAN		.99	.94	.90	.86	1.36	1.27	1.22	1.19	1.42	1.56	1.66	1.74			
MOVING AVERAGE		1.13	1.14	1.24	1.17	1.33	1.37	1.43	1.45	1.37	1.52	1.67	1.78			

Table 31

Median MAD Ratios To Median Naive One MAD
for Stationary Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.88	1.07	1.45
	.2	.86	.83	.82	.81	1.10	1.09	1.07	1.07	1.51	1.56	1.61	1.62	.89	1.06	1.30
SMOOTHED	.3	.88	.84	.83	.81	1.09	1.09	1.05	1.05	1.47	1.53	1.57	1.62	.88	1.04	1.22
MEDIAN	.4	.90	.86	.84	.81	1.09	1.09	1.06	1.04	1.43	1.51	1.54	1.58	.89	1.00	1.12
	.5	.91	.87	.84	.82	1.11	1.09	1.07	1.05	1.35	1.48	1.53	1.56	.92	.96	1.08
	.6	.92	.87	.84	.82	1.12	1.09	1.07	1.05	1.31	1.44	1.50	1.54			
WALSH AVERAGE		.90	.89	.92	.89	1.13	1.15	1.11	1.11	1.26	1.30	1.36	1.41			
RUNNING MEDIAN		.92	.85	.82	.81	1.16	1.08	1.08	1.04	1.32	1.31	1.45	1.46			
MOVING AVERAGE		.92	.98	.99	1.01	1.11	1.13	1.16	1.17	1.24	1.28	1.37	1.45			

Table 32

Median MSE Ratios To Median Naive One MSE
for Stationary Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	1.01	1.15	1.57
	.2	.89	.89	.84	.84	1.17	1.18	1.15	1.15	1.58	1.68	1.75	1.78	1.00	1.07	1.33
SMOOTHED	.3	.90	.90	.84	.84	1.17	1.17	1.16	1.15	1.43	1.58	1.63	1.60	.98	1.05	1.19
MEDIAN	.4	.93	.91	.84	.84	1.18	1.17	1.16	1.15	1.43	1.51	1.64	1.59	.97	1.03	1.11
	.5	.97	.92	.85	.84	1.20	1.18	1.15	1.15	1.45	1.52	1.65	1.58	.99	.99	1.07
	.6	.96	.93	.85	.85	1.23	1.20	1.15	1.15	1.44	1.52	1.66	1.58			
WALSH AVERAGE		1.02	1.01	.94	1.00	1.20	1.31	1.21	1.17	1.35	1.47	1.48	1.49			
RUNNING MEDIAN		.93	.93	.83	.85	1.21	1.23	1.17	1.14	1.39	1.46	1.56	1.52			
MOVING AVERAGE		1.12	1.19	1.16	1.09	1.21	1.26	1.26	1.22	1.33	1.40	1.44	1.49			

Table 33

Median MAPE Ratios To Median Naive One MAPE
for Stationary Series, Cauchy Errors

Technique	C	Φ=.3 n 3	4	5	6	Φ=.5 n 3	4	5	6	Φ=.7 n 3	4	5	6	EXPONENTIAL SMOOTHING Φ .3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.90	1.08	1.80
	.2	.93	.96	.88	.88	1.15	1.16	1.16	1.19	1.93	1.99	1.99	2.05	.91	1.06	1.57
SMOOTHED	.3	.92	.92	.88	.88	1.15	1.18	1.15	1.14	1.82	1.92	1.89	1.95	.87	.98	1.28
MEDIAN	.4	.91	.91	.89	.88	1.12	1.18	1.15	1.14	1.69	1.85	1.89	1.96	.88	.96	1.22
	.5	.94	.91	.89	.88	1.14	1.16	1.15	1.15	1.65	1.77	1.86	1.92	.93	.95	1.20
	.6	.95	.92	.89	.88	1.16	1.14	1.15	1.14	1.61	1.80	1.89	1.94			
WALSH AVERAGE		.91	1.00	.94	.90	1.13	1.19	1.13	1.13	1.49	1.62	1.65	1.69			
RUNNING MEDIAN		.96	.92	.88	.86	1.16	1.14	1.15	1.10	1.56	1.66	1.78	1.84			
MOVING AVERAGE		.95	.96	1.01	.97	1.13	1.14	1.14	1.17	1.48	1.48	1.63	1.70			

Table 34

Median Theil's U Ratios
for Stationary Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	-	-	-	-	-	-	-	-	-	-	-	-	.88	1.03	1.19
	.2	.89	.88	.87	.87	1.05	1.06	1.04	1.04	1.33	1.35	1.35	1.40	.88	1.01	1.17
SMOOTHED	.3	.90	.89	.86	.87	1.06	1.04	1.04	1.04	1.25	1.34	1.34	1.34	.92	1.00	1.14
MEDIAN	.4	.92	.90	.87	.87	1.05	1.06	1.02	1.04	1.32	1.38	1.33	1.33	.93	1.00	1.10
	.5	.92	.91	.87	.87	1.07	1.06	1.03	1.03	1.31	1.38	1.38	1.36	.94	.99	1.08
	.6	.93	.92	.87	.86	1.09	1.07	1.04	1.03	1.26	1.33	1.38	1.38			
WALSH AVERAGE		1.01	.94	.92	.90	1.05	1.08	1.07	1.05	1.22	1.30	1.27	1.31			
RUNNING MEDIAN		.98	.92	.89	.86	1.07	1.08	1.03	1.02	1.18	1.30	1.33	1.39			
MOVING AVERAGE		1.02	1.02	.98	.96	1.05	1.05	1.07	1.04	1.21	1.27	1.29	1.32			

Table 35

Mean MAD Ratios To Naive One MAD
for Trend Series, Normal Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.91	.92	.92	.92	1.07	1.09	1.10	1.11	1.30	1.34	1.37	1.40	.87	.99	1.17
	.2	.98	.98	1.00	.99	1.14	1.16	1.18	1.19	1.33	1.39	1.44	1.47	.90	1.00	1.12
DOUBLE	.3	1.04	1.04	1.05	1.04	1.18	1.21	1.24	1.24	1.34	1.41	1.48	1.50	.94	1.02	1.09
SMOOTHED	.4	1.09	1.08	1.10	1.08	1.23	1.24	1.28	1.28	1.35	1.42	1.51	1.52	.99	1.04	1.08
MEDIAN	.5	1.14	1.11	1.14	1.11	1.27	1.26	1.32	1.30	1.37	1.42	1.52	1.51	1.05	1.08	1.10
	.6	1.20	1.15	1.18	1.13	1.31	1.29	1.35	1.31	1.40	1.42	1.53	1.50			
ROBUST REGRESSION		1.33	1.25	1.21	1.15	1.34	1.30	1.29	1.26	1.33	1.32	1.34	1.35			
DOUBLE RUNNING MEDIAN		1.42	1.22	1.26	1.16	1.52	1.35	1.43	1.34	1.57	1.44	1.60	1.53			
DOUBLE MOVING AVERAGE		1.19	1.13	1.07	1.04	1.26	1.24	1.22	1.20	1.30	1.33	1.36	1.37			

Table 36

Mean MSE Ratios To Naive One MSE
for Trend Series, Normal Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING Φ		
		n				n				n						
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.83	.84	.85	.85	1.14	1.17	1.21	1.22	1.69	1.80	1.90	1.97	.76	.98	1.38
	.2	.95	.96	.98	.97	1.27	1.32	1.39	1.40	1.75	1.91	2.08	2.17	.82	.99	1.24
DOUBLE	.3	1.08	1.07	1.11	1.07	1.39	1.44	1.52	1.52	1.80	1.97	2.17	2.25	.89	1.02	1.18
SMOOTHED	.4	1.20	1.17	1.21	1.15	1.51	1.54	1.63	1.60	1.85	2.00	2.22	2.26	.99	1.08	1.17
MEDIAN	.5	1.32	1.25	1.31	1.22	1.62	1.61	1.72	1.65	1.90	2.01	2.27	2.25	1.11	1.17	1.20
	.6	1.44	1.32	1.41	1.27	1.74	1.66	1.81	1.68	1.97	2.02	2.32	2.22			
ROBUST REGRESSION		1.80	1.55	1.46	1.35	1.82	1.68	1.66	1.61	1.79	1.74	1.82	1.86			
DOUBLE RUNNING MEDIAN		2.13	1.49	1.66	1.37	2.44	1.81	2.09	1.80	2.61	2.09	2.64	2.34			
DOUBLE MOVING AVERAGE		1.41	1.28	1.17	1.07	1.58	1.54	1.47	1.40	1.69	1.78	1.82	1.83			

Table 37

Mean MAPE Ratios To Naive One MAPE
for Trend Series, Normal Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING Φ		
		n				n				n						
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.96	.97	.98	.97	1.16	1.17	1.19	1.19	1.43	1.48	1.52	1.54	.92	1.07	1.29
	.2	1.02	1.03	1.04	1.04	1.21	1.23	1.26	1.27	1.43	1.50	1.56	1.59	.94	1.05	1.19
DOUBLE	.3	1.07	1.08	1.10	1.08	1.24	1.27	1.32	1.31	1.42	1.50	1.58	1.60	.97	1.05	1.13
SMOOTHED	.4	1.12	1.12	1.14	1.12	1.28	1.30	1.35	1.34	1.41	1.49	1.59	1.59	1.01	1.06	1.10
MEDIAN	.5	1.17	1.15	1.18	1.14	1.31	1.32	1.37	1.35	1.41	1.47	1.59	1.58	1.07	1.09	1.10
	.6	1.22	1.17	1.21	1.16	1.34	1.33	1.39	1.36	1.42	1.46	1.59	1.56			
ROBUST REGRESSION		1.34	1.26	1.23	1.17	1.35	1.31	1.30	1.28	1.32	1.31	1.35	1.38			
DOUBLE RUNNING MEDIAN		1.43	1.23	1.29	1.19	1.54	1.37	1.47	1.39	1.57	1.46	1.64	1.58			
DOUBLE MOVING AVERAGE		1.20	1.15	1.10	1.08	1.27	1.26	1.26	1.25	1.30	1.35	1.40	1.42			

Table 38

Mean Theil's U Values
for Trend Series, Normal Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING Φ		
		n				n				n						
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.87	.88	.88	.88	1.02	1.04	1.06	1.06	1.24	1.28	1.32	1.35	.82	.94	1.10
	.2	.93	.93	.94	.93	1.06	1.09	1.12	1.12	1.23	1.29	1.35	1.38	.85	.93	1.03
DOUBLE	.3	.98	.98	.99	.97	1.10	1.12	1.15	1.15	1.23	1.29	1.36	1.39	.90	.96	1.02
SMOOTHED	.4	1.04	1.02	1.04	1.00	1.15	1.16	1.19	1.17	1.25	1.29	1.37	1.38	.96	1.00	1.04
MEDIAN	.5	1.09	1.06	1.07	1.03	1.20	1.18	1.22	1.18	1.28	1.30	1.38	1.37	1.03	1.06	1.09
	.6	1.14	1.09	1.11	1.05	1.24	1.21	1.24	1.19	1.31	1.31	1.39	1.36			
ROBUST REGRESSION		1.33	1.22	1.16	1.11	1.35	1.27	1.23	1.20	1.35	1.30	1.29	1.28			
DOUBLE RUNNING MEDIAN		1.38	1.16	1.19	1.08	1.49	1.27	1.33	1.22	1.55	1.36	1.47	1.38			
DOUBLE MOVING AVERAGE		1.14	1.07	1.02	.98	1.21	1.17	1.13	1.10	1.26	1.24	1.24	1.25			

Table 39

Mean MAD Ratios To Naive One MAD
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.91	.90	.83	.82	1.29	1.29	1.23	1.21	1.87	1.96	1.99	2.02	1.00	1.22	
	.2	.97	.95	.87	.85	1.34	1.37	1.30	1.29	1.86	1.97	2.05	2.10	1.03	1.18	1.46
DOUBLE	.3	1.03	.99	.90	.88	1.39	1.41	1.37	1.33	1.81	1.99	2.09	2.14	1.09	1.16	1.32
SMOOTHED	.4	1.08	1.04	.93	.90	1.44	1.42	1.42	1.36	1.83	1.95	2.13	2.14	1.14	1.19	1.25
MEDIAN	.5	1.12	1.07	.95	.91	1.49	1.46	1.45	1.39	1.86	1.89	2.16	2.13	1.19	1.22	1.25
	.6	1.16	1.10	.97	.92	1.53	1.50	1.48	1.40	1.88	1.90	2.18	2.14			
ROBUST REGRESSION		1.49	1.15	1.15	1.07	1.46	1.20	1.40	1.37	1.41	1.32	1.54	1.60			
DOUBLE RUNNING MEDIAN		1.34	1.18	1.02	.94	1.72	1.57	1.56	1.43	2.02	1.97	2.26	2.19			
DOUBLE MOVING AVERAGE		1.53	1.49	1.46	1.43	1.53	1.55	1.60	1.64	1.50	1.68	1.86	1.93			

Table 40

Mean MSE Ratios To Naive One MSE
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING Φ		
		n				n				n						
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.76	.75	.73	.73	1.11	1.12	1.08	1.07	1.76	1.84	1.88	1.90	.77	1.01	1.42
	.2	.79	.78	.74	.73	1.20	1.21	1.16	1.13	1.87	2.02	2.12	2.14	.83	1.02	1.29
DOUBLE	.3	.83	.80	.75	.74	1.28	1.29	1.22	1.18	1.94	2.11	2.29	2.31	.91	1.04	1.21
SMOOTHED	.4	.86	.82	.76	.75	1.36	1.34	1.28	1.22	2.01	2.17	2.43	2.42	1.01	1.10	1.20
MEDIAN	.5	.90	.84	.77	.75	1.44	1.39	1.33	1.25	2.11	2.21	2.54	2.50	1.13	1.19	1.23
	.6	.95	.86	.77	.75	1.55	1.43	1.37	1.26	2.26	2.25	2.66	2.54			
ROBUST REGRESSION		1.42	.92	.91	.83	1.51	1.10	1.27	1.21	1.69	1.30	1.65	1.65			
DOUBLE RUNNING MEDIAN		1.27	.92	.79	.76	2.30	1.56	1.50	1.31	3.64	2.48	3.08	2.74			
DOUBLE MOVING AVERAGE		1.42	1.26	1.20	1.13	1.59	1.51	1.53	1.51	1.69	1.75	1.92	2.01			

Table 41

Mean MAPE Ratios To Naive One MAPE
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.96	.95	.87	.86	1.44	1.44	1.34	1.31	2.42	2.45	2.31	2.24	1.08	1.44	2.20
	.2	1.04	1.02	.92	.90	1.55	1.56	1.45	1.42	2.63	2.77	2.67	2.58	1.12	1.40	2.04
DOUBLE	.3	1.07	1.04	.95	.93	1.63	1.64	1.53	1.48	2.61	2.85	2.82	2.73	1.13	1.35	1.75
SMOOTHED	.4	1.08	1.05	.96	.95	1.66	1.68	1.60	1.55	2.55	2.85	2.86	2.72	1.15	1.31	1.51
MEDIAN	.5	1.11	1.06	.97	.94	1.69	1.71	1.65	1.58	2.47	2.86	2.90	2.68	1.18	1.27	1.28
	.6	1.16	1.08	.97	.94	1.70	1.74	1.68	1.60	2.30	2.80	2.95	2.65			
ROBUST REGRESSION		1.56	1.18	1.13	1.03	1.50	1.36	1.53	1.60	1.33	1.38	1.71	1.87			
DOUBLE RUNNING MEDIAN		1.33	1.19	.98	.94	1.78	1.76	1.78	1.64	1.68	2.61	3.18	2.71			
DOUBLE MOVING AVERAGE		1.55	1.50	1.47	1.41	1.69	1.90	1.99	2.03	1.50	2.50	2.68	2.87			

Table 42

Mean Theil's U Values
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	1.09	1.07	.95	.93	1.39	1.39	1.30	1.29	2.30	2.43	2.44	2.48	1.22	1.36	2.02
	.2	1.17	1.16	1.12	.98	1.39	1.43	1.35	1.35	1.96	2.12	2.19	2.25	1.22	1.23	1.57
DOUBLE	.3	1.21	1.21	1.06	1.01	1.37	1.41	1.33	1.33	1.98	2.19	2.31	2.37	1.15	1.16	1.41
SMOOTHED	.4	1.21	1.23	1.10	1.03	1.43	1.46	1.35	1.31	1.87	2.10	2.27	2.32	1.14	1.17	1.24
MEDIAN	.5	1.23	1.23	1.12	1.04	1.48	1.51	1.41	1.32	1.75	1.98	2.20	2.18	1.22	1.17	1.19
	.6	1.33	1.23	1.13	1.03	1.52	1.53	1.46	1.34	1.72	1.87	2.15	2.08			
ROBUST REGRESSION		1.66	1.52	1.68	1.33	1.50	1.36	1.43	1.39	1.52	1.36	1.58	1.69			
DOUBLE RUNNING MEDIAN		1.65	1.37	1.15	1.05	1.81	1.53	1.53	1.40	1.95	1.79	2.17	2.08			
DOUBLE MOVING AVERAGE		1.89	1.74	1.78	1.75	1.55	1.65	1.77	1.75	1.65	1.63	2.01	2.14			

Table 43

Median MAD Ratios To Median Naive One MAD
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING Φ		
		n				n				n						
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.88	.85	.84	.84	1.25	1.25	1.17	1.17	1.68	1.74	1.79	1.81	.95	1.16	1.46
DOUBLE	.2	.97	.92	.88	.87	1.22	1.21	1.16	1.19	1.71	1.78	1.78	1.78	1.03	1.19	1.39
SMOOTHED	.3	1.01	.97	.92	.90	1.31	1.27	1.24	1.22	1.62	1.74	1.82	1.84	1.06	1.13	1.28
MEDIAN	.4	1.06	.99	.94	.93	1.36	1.28	1.28	1.26	1.64	1.72	1.85	1.83	1.13	1.17	1.20
	.5	1.08	1.02	.96	.95	1.41	1.31	1.32	1.28	1.69	1.64	1.84	1.85	1.19	1.23	1.21
	.6	1.12	1.04	.97	.95	1.48	1.33	1.37	1.29	1.76	1.65	1.85	1.86			
ROBUST REGRESSION		1.45	1.20	1.19	1.09	1.47	1.27	1.37	1.28	1.43	1.31	1.49	1.47			
DOUBLE RUNNING MEDIAN		1.31	1.11	1.03	.97	1.74	1.44	1.41	1.32	1.98	1.77	1.93	1.90			
DOUBLE MOVING AVERAGE		1.50	1.44	1.34	1.34	1.54	1.53	1.51	1.51	1.49	1.55	1.65	1.74			

Table 44

Median MSE Ratios To Median Naive One MSE
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.94	.91	.86	.86	1.23	1.23	1.18	1.18	1.83	1.93	1.95	1.97	1.04	1.14	1.51
DOUBLE	.2	1.04	1.03	.86	.86	1.34	1.33	1.24	1.20	1.85	1.94	1.96	1.94	1.10	1.19	1.34
SMOOTHED	.3	1.10	1.11	.93	.90	1.50	1.43	1.28	1.25	1.92	1.97	1.97	1.94	1.13	1.22	1.30
MEDIAN	.4	1.19	1.16	.99	.97	1.68	1.53	1.30	1.27	2.11	2.02	1.98	1.99	1.27	1.24	1.27
	.5	1.24	1.21	1.02	.99	1.82	1.64	1.33	1.28	2.29	2.15	2.05	2.00	1.46	1.29	1.21
	.6	1.42	1.25	1.03	1.00	1.90	1.78	1.36	1.29	2.50	2.29	2.10	2.05			
ROBUST REGRESSION		1.96	1.41	1.43	1.15	1.85	1.38	1.61	1.43	1.85	1.39	1.75	1.68			
DOUBLE RUNNING MEDIAN		1.86	1.33	1.15	1.01	2.72	1.96	1.62	1.32	3.78	2.63	2.51	2.19			
DOUBLE MOVING AVERAGE		2.12	1.90	1.64	1.55	1.99	1.92	1.78	1.73	1.89	1.95	1.95	1.92			

Table 45

Median MAPE Ratios To Median Naive One MAPE
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.89	.88	.83	.82	1.30	1.28	1.25	1.20	1.82	1.95	1.98	1.99	.94	1.21	1.66
DOUBLE	.2	.95	.93	.87	.86	1.33	1.38	1.31	1.29	1.93	2.14	2.32	2.32	.90	1.12	1.49
SMOOTHED	.3	1.01	.99	.90	.89	1.40	1.44	1.40	1.36	1.74	1.97	2.16	2.40	.99	1.14	1.30
MEDIAN	.4	1.05	1.05	.93	.90	1.46	1.49	1.48	1.43	1.77	1.91	2.14	2.27	1.17	1.23	1.23
	.5	1.09	1.07	.98	.91	1.56	1.55	1.52	1.50	1.83	1.83	2.14	2.15	1.13	1.29	1.22
	.6	1.15	1.09	1.01	.90	1.64	1.54	1.49	1.41	1.69	1.90	2.14	2.05			
ROBUST REGRESSION		1.46	1.09	1.12	1.01	1.60	1.29	1.46	1.51	1.31	1.27	1.44	1.65			
DOUBLE RUNNING MEDIAN		1.30	1.15	1.02	.91	1.84	1.63	1.59	1.42	1.82	1.86	2.30	2.11			
DOUBLE MOVING AVERAGE		1.41	1.36	1.26	1.22	1.69	1.58	1.71	1.73	1.44	1.78	1.81	2.03			

Table 46

Median Theil's U Values
for Trend Series, Cauchy Errors

		Φ=.3				Φ=.5				Φ=.7				LINEAR EXPONENTIAL SMOOTHING		
		n				n				n				Φ		
Technique	C	3	4	5	6	3	4	5	6	3	4	5	6	.3	.5	.7
	.1	.94	.94	.90	.89	1.06	1.04	1.02	1.02	1.50	1.59	1.62	1.63	.96	1.03	1.30
DOUBLE	.2	.98	.96	.93	.93	1.09	1.10	1.06	1.05	1.39	1.47	1.53	1.58	1.01	1.04	1.15
SMOOTHED	.3	.99	.99	.93	.92	1.17	1.19	1.08	1.07	1.46	1.57	1.49	1.58	1.01	1.06	1.13
MEDIAN	.4	1.02	1.00	.95	.93	1.20	1.23	1.13	1.07	1.44	1.44	1.60	1.52	1.02	1.05	1.06
	.5	1.07	1.01	.96	.94	1.24	1.19	1.15	1.09	1.37	1.56	1.61	1.50	1.07	1.07	1.07
	.6	1.11	1.03	.96	.94	1.31	1.17	1.18	1.12	1.30	1.35	1.56	1.50			
ROBUST REGRESSION		1.34	1.17	1.08	1.03	1.23	1.15	1.18	1.14	1.12	1.07	1.20	1.32			
DOUBLE RUNNING MEDIAN		1.31	1.11	.99	.94	1.36	1.30	1.19	1.13	1.33	1.34	1.53	1.53			
DOUBLE MOVING AVERAGE		1.31	1.19	1.17	1.17	1.37	1.23	1.29	1.29	1.17	1.28	1.42	1.59			

APPENDIX D

SAMPLE EMPIRICAL PROGRAMS

These programs were written and run using the VMS Version of SAS, Release 5.16. Reference SAS Institute Inc. (1985).

PROGRAM TO READ THE DATA FROM THE M-COMPETITION DATA SET AND CHANGE IT INTO A SAS DATA SET FOR FURTHER PROCESSING. THE DATA WAS RECEIVED FROM PROFESSOR MAKRIDAKIS ON MS-DOS COMPATIBLE FILES ON FLOPPY DISKS, THEN READ INTO THE MAINFRAME USING A CARDWARE SYSTEM. THE FORMAT OF THE DATA WAS NOT USEFUL FOR SAS PROCESSING, SO IT HAD TO BE MODIFIED.

```
LIBNAME STAT '[RKANKEY.SIX2]';
DATA STAT.S111;
INFILE SER111;
INPUT SID / NFIT / SEASON / NOF / DUMMY;
DO I = 1 TO NFIT;
  INPUT X;                  Key:
  FITCODE = 'FIT';           SID  Series I.D.
  OBSNUM = I;                NFIT Number of obs used for fitting
  OUTPUT;                    SEASON identifies if nonseasonal,
END;                              quarterly or monthly seasonal
DO I = 1 TO NOF;             NOF number of hold out observations
  INPUT X;                   DUMMY a duplicate series ID
  FITCODE = 'NOF';           X  the series values
  OBSNUM = NFIT + I;         FITCODE is 'fit' for fitted values,
  OUTPUT;                       'nof' for hold out values, and
END;                            'adj' for the seasonal indices
DO I = 1 TO 12;              OBSNUM is the observation number
  INPUT X;
  FITCODE = 'ADJ';
  OBSNUM = .;
  OUTPUT;
END;
DROP DUMMY I;
```

MOVING AVERAGE PROGRAM. THIS PROGRAM READS DATA FROM THE SUBSET OF NO MONOTONIC TREND SERIES AND CALCULATES THE ONE PERIOD AHEAD FIT MEASURES FOR MOVING AVERAGES OF LENGTH ONE THROUGH MASIZE. EACH SERIES HAS NFIT VALUES USED FOR FITTING AND NOF HOLDOUT VALUES. WHEN TESTED FOR MONOTONIC TREND, THE NULL HYPOTHESIS OF NO MONOTONIC TREND COULD NOT BE REJECTED FOR NINETEEN. THESE NINETEEN SERIES ARE IN A SAS DATA SET NOTREND2.SSD.

```
LIBNAME STUFF '[RKANKEY.SIX2]';
DATA INTERIM;
SET STUFF.NOTREND2;
RETAIN;
IF OBSNUM GT NFIT THEN GOTO NEXTOBS;          Do not
ARRAY MA{10} MA1-MA10;                      calculate
ARRAY SUMS{10,6} SUM1-SUM60;         anything for the
IF OBSNUM = 1 THEN DO;                  hold out data.
  MASIZE = 10;
  IF NFIT/2 LT 10 THEN DO;          MASIZE is equal to
    MASIZE = CEIL(NFIT/2);              ten unless the
  END;                               series is short.
END;
LAGX = LAG(X);                            Renaming for
L1 = LAG(X);                      easier programming.
L2 = LAG2(X);
L3 = LAG3(X);
L4 = LAG4(X);
L5 = LAG5(X);
L6 = LAG6(X);
L7 = LAG7(X);
L8 = LAG8(X);
L9 = LAG9(X);
SUM = X;
MA1 = SUM;                             Calculating the
SUM = SUM+L1;                          moving averages.
MA2 = SUM/2;
SUM = SUM+L2;
MA3 = SUM/3;
SUM = SUM+L3;
MA4 = SUM/4;
SUM = SUM+L4;
MA5 = SUM/5;
SUM = SUM+L5;
MA6 = SUM/6;
SUM = SUM+L6;
MA7 = SUM/7;
SUM = SUM+L7;
MA8 = SUM/8;
SUM = SUM+L8;
MA9 = SUM/9;
SUM = SUM+L9;
MA10 = SUM/10;
```

```
DO J = 1 TO MASIZE;
  N = J;
  F = LAG(MA{J});                                  Forecasting.
  IF N GE I THEN GO TO SKIP;
  E = X-F;                Calculating the fit error.
  V8 = ABS(E);                 Manipulating the error.
  V9 = E**2;
  V10 = (V8/X)**2;
  V11 = ABS(E/X);
  V12 = (X-LAGX)**2;
  V13 = ABS(E/LAGX);
  V14 = ABS((X-LAGX)/LAGX);
  V16 = ABS(X-LAGX);
  V17 = (E/LAGX)**2;
  V18 = ((X-LAGX)/LAGX)**2;
  SUMS{J,1} + V8;                                     Summing.
  SUMS{J,2} + V9;
  SUMS{J,3} + V10;             Note that the use of
  SUMS{J,4} + V11;             eight error measures
  SUMS{J,5} + V17;            is discontinued.  Now
  SUMS{J,6} + V18;              only four are used.
  SKIP:;
  IF OBSNUM = NFIT THEN DO;
    MAD = SUMS{J,1}/(NFIT-J);    Calculating the
    MSE = SUMS{J,2}/(NFIT-J);     error measures
    MAPE = SUMS{J,4}/(NFIT-J);   for the series.
    TH3 = SQRT(SUMS{J,5}/SUMS{J,6});
    FILE MATEMP1;                                  Filing.
    PUT SID N MASIZE SEASON RSCORR NFIT
    MAD MSE MAPE TH3 MA{J};
  END;
END;
IF OBSNUM = NFIT THEN DO;       Resetting for the
  DO M = 1 TO 10;                          next series.
    DO N = 1 TO 6;                 Each series starts
      SUMS{M,N} = 0;          with OBSNUM = 1, but
    END;                      with NFIT equal to the
  END;               number of fitted values in the
END;                                        new series.
NEXTOBS:;
```

EXPONENTIAL AVERAGE PROGRAM. THIS PROGRAM CALCULATES THE ERROR MEASURES BY SERIES OVER THE NO MONOTONIC TREND SUBSET. THE SMOOTHING CONSTANT VARIES FROM ZERO TO ONE BY INCREMENTS OF .05. RESULTS ARE FILED SO THAT A LATER SEARCH PROGRAM CAN IDENTIFY THE BEST FITTING SMOOTHING CONSTANTS FOR EACH SERIES FOR EACH OF THE FOUR ERROR MEASURES (MAD, MSE, MAPE, THEIL'S U) PRIOR TO FORECASTING.

```
LIBNAME RDK '[RKANKEY.SIX2]';
DATA _NULL_;
SET RDK.NOTREND2;
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
RETAIN;
ARRAY EX{21} EX1-EX21;
ARRAY SUMS{21,6} S1-S126;
LAGX = LAG(X);
IF OBSNUM = 1 THEN DO;
  INFILE EASTART;            Used the average of the
  INPUT SID XBAR;             first six observations
END;                                of each series to
DO J = 1 TO 21;                           initialize.
  C = (J-1)*.05;            Smoothing constants from
  IF OBSNUM = 1 THEN DO;        zero to one by .05.
    EX0 = XBAR;
    EX{J} = C*X + (1-C)*EX0;
  END;
  IF OBSNUM GT 1 THEN DO;
    EX{J} = C*X + (1-C)*EX{J};
  END;
  F = LAG(EX{J});
  IF OBSNUM = 1 THEN GO TO SKIP;    There was no
  E = X - F;                     forecast before
  V8 = ABS(E);                    time period 2.
  V9 = E**2;
  V11 = ABS(E/X);             Note that many of the
  V17 = (E/LAGX)**2;     unnecessary V variables
  V18 = ((X-LAGX)/LAGX)**2;   have been dropped.
  SUMS{J,1} + V8;
  SUMS{J,2} + V9;                         Summing.
  SUMS{J,3} + V11;
  SUMS{J,4} + V17;
  SUMS{J,5} + V18;
  IF OBSNUM = NFIT THEN DO;
    MAD = SUMS{J,1}/(NFIT-1);       Calculating.
    MSE = SUMS{J,2}/(NFIT-1);
    MAPE = SUMS{J,3}/(NFIT-1);
    TH3 = SQRT(SUMS{J,4}/SUMS{J,5});
    FILE EAMEAS;                           Filing.
    PUT SID C MAD MSE MAPE TH3 EX{J};
  END;
SKIP:;
END;
```

```
IF OBSNUM = NFIT THEN DO;      Reset for the next
DO M = 1 TO 21;                            series.
  DO N = 1 TO 6;
    SUMS{M,N} = 0;
  END;
END;
END;
NEXTOBS:;
```

RUNNING MEDIAN PROGRAM. THIS PROGRAM DEVELOPS THE FITTING MEASURES FOR EACH OF THE NO MONOTONIC TREND SERIES WHEN THE TECHNIQUE OF RUNNING MEDIANS IS APPLIED. RUNNING MEDIANS OF LENGTH ONE THROUGH SIX ARE CONSIDERED. RESULTS ARE FILED FOR A LATER SEARCH PROGRAM PRIOR TO FORECASTING VALUES OF THE HOLD OUT DATA.

```
LIBNAME STUFF '[RKANKEY.SIX2]';
DATA _NULL_;
SET STUFF.NOTREND2;
ARRAY SUMS{6,5} S1-S30;
XI = X;
LAGXI = LAG(XI);
LAGX = LAG(X);
L1=LAGX;                                  Renaming.
L2=LAG2(X);
L3=LAG3(X);
L4=LAG4(X);
L5=LAG5(X);
IF OBSNUM GT NFIT THEN GOTO NEXTOBS;
ARRAY TEMP3{3} X L1 L2;
ARRAY TEMP4{4} X L1 L2 L3;
ARRAY TEMP5{5} X L1 L2 L3 L4;
ARRAY TEMP6{6} X L1 L2 L3 L4 L5;
RM1 = XI;                        Finding the running
RM2 = (XI+L1)/2;                            medians.
DO H=1 TO 3;
MORE=0;
DO K = 1 TO 2;
     IF TEMP3{K+1} LT TEMP3{K} THEN DO;
          TEMP = TEMP3{K};
          TEMP3{K} = TEMP3{K+1};
          TEMP3{K+1} = TEMP;
          MORE = 1;                  Using a bubble
          END;                     sort for those of
     END;                      length three or more.
     IF MORE = 0 THEN GO TO DONE3;
END;
DONE3:;
RM3 = TEMP3{2};
DO H = 1 TO 4;
MORE = 0;
DO I = 1 TO 3;
   IF TEMP4{K+1} LT TEMP4{K} THEN DO;
     TEMP = TEMP4{K};
     TEMP4{K} = TEMP4{K+1};
     TEMP4{K+1} = TEMP;
     MORE = 1;
     END;
   END;
   IF MORE = 0 THEN GO TO DONE4;
END;
```

```
DONE4:;
RM4 = (TEMP4{2} + TEMP4{3})/2;
DO H = 1 TO 5;
MORE = 0;
  DO K = 1 TO 4;
    IF TEMP5{K+1} LT TEMP5{K} THEN DO;
    TEMP = TEMP5{K};
    TEMP5{K} = TEMP5{K+1};
    TEMP5{K+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE5;
END;
DONE5:;
RM5 = TEMP5{3};
DO H = 1 TO 6;
MORE = 0;
  DO K = 1 TO 5;
    IF TEMP6{K+1} LT TEMP6{K} THEN DO;
    TEMP = TEMP6{K};
    TEMP6{K} = TEMP6{K+1};
    TEMP6{K+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE6;
END;
DONE6:;
RM6 = (TEMP6{3} + TEMP6{4})/2;
ARRAY RM{6} RM1 RM2 RM3 RM4 RM5 RM6;
DO J=1 TO 6;                  The forecast is the most
  F = LAG(RM{J});             recent running median of
  IF OBSNUM LT J+1 THEN GO TO CUT;      the right
  E = XI-F;                                 length.
  V8 = ABS(E);
  V9 = E**2;                          Manipulating.
  V11 = ABS(E/XI);
  V17 = (E/LAGXI)**2;
  V18 = ((XI-LAGXI)/LAGXI)**2;
  SUMS{J,1} + V8;
  SUMS{J,2} + V9;                        Summing.
  SUMS{J,3} + V11;
  SUMS{J,4} + V17;
  SUMS{J,5} + V18;
  IF OBSNUM EQ NFIT THEN DO;
    MAD = SUMS{J,1}/(NFIT-J);       Calculating
    MSE = SUMS{J,2}/(NFIT-J);    for the series.
    MAPE = SUMS{J,3}/(NFIT-J);
    TH3 = SQRT(SUMS{J,4}/SUMS{J,5});
    FILE RMRES;                          Filing.
    PUT SID J MAD MSE MAPE TH3 RM{J};
```

```
  END;
CUT:END;
IF OBSNUM = NFIT THEN DO;                Resetting for
  DO M = 1 TO 6;                     the next series.
    DO N = 1 TO 5;
      SUMS{M,N} = 0;
    END;
  END;
END;
NEXTOBS:;
```

SMOOTHED MEDIAN PROGRAM. THIS PROGRAM CONSIDERS MEDIANS FROM LENGTH ONE THROUGH SIX, AND SMOOTHING CONSTANTS OF .01, .05 TO .95 BY .05, AND .99. IT CAN BE VERY STABLE (LARGE LENGTH MEDIANS AND SMALL SMOOTHING CONSTANT) OR VERY RESPONSIVE (SHORT LENGTH MEDIANS AND A LARGE SMOOTHING CONSTANT). RESULTS ARE FILED AND USED BY A LATER SEARCH PROGRAM FOR THE IDENTIFICATION OF THE BEST FITTING COMBINATION OF LENGTH AND SMOOTHING CONSTANT FOR EACH SERIES FOR EACH ERROR MEASURE.

```
LIBNAME STUFF '[RKANKEY.SIX2]';
DATA _NULL_;
SET STUFF.NOTREND2;
RETAIN;
ARRAY SUMS{6,21,5} S1-S630;
XJ = X;
LAGX = LAG(X);
L1=LAGX;                                       Renaming.
L2=LAG2(X);
L3=LAG3(X);
L4=LAG4(X);
L5=LAG5(X);
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
ARRAY TEMP3{3} X L1 L2;
ARRAY TEMP4{4} X L1 L2 L3;
ARRAY TEMP5{5} X L1 L2 L3 L4;
ARRAY TEMP6{6} X L1 L2 L3 L4 L5;
RM1 = X;
RM2 = (X+LAGX)/2;                          Developing the
DO H=1 TO 3;                              running medians.
MORE=0;
DO J = 1 TO 2;
     IF TEMP3{J+1} LT TEMP3{J} THEN DO;
          TEMP = TEMP3{J};
          TEMP3{J} = TEMP3{J+1};
          TEMP3{J+1} = TEMP;
          MORE = 1;                        Several bubble
          END;                                      sorts.
     END;
     IF MORE = 0 THEN GO TO DONE3;
END;
DONE3:;
RM3 = TEMP3{2};
DO H = 1 TO 4;
MORE = 0;
DO J = 1 TO 3;
   IF TEMP4{J+1} LT TEMP4{J} THEN DO;
     TEMP = TEMP4{J};
     TEMP4{J} = TEMP4{J+1};
     TEMP4{J+1} = TEMP;
     MORE = 1;
     END;
```

```
      END;
      IF MORE = 0 THEN GO TO DONE4;
    END;
    DONE4:;
    RM4 = (TEMP4{2} + TEMP4{3})/2;
    DO H = 1 TO 5;
    MORE = 0;
      DO J = 1 TO 4;
        IF TEMP5{J+1} LT TEMP5{J} THEN DO;
        TEMP = TEMP5{J};
        TEMP5{J} = TEMP5{J+1};
        TEMP5{J+1} = TEMP;
        MORE = 1;
        END;
      END;
      IF MORE = 0 THEN GO TO DONE5;
    END;
    DONE5:;
    RM5 = TEMP5{3};
    DO H = 1 TO 6;
    MORE = 0;
      DO J = 1 TO 5;
        IF TEMP6{J+1} LT TEMP6{J} THEN DO;
        TEMP = TEMP6{J};
        TEMP6{J} = TEMP6{J+1};
        TEMP6{J+1} = TEMP;
        MORE = 1;
        END;
      END;
      IF MORE = 0 THEN GO TO DONE6;
    END;                              Running medians are all
    DONE6:;                         calculated, place them in
    RM6 = (TEMP6{3} + TEMP6{4})/2;                  an array.
    ARRAY RM{6} RM1 RM2 RM3 RM4 RM5 RM6;
    LAGXJ = LAG(XJ);
    ARRAY SM{6,21} SM1-SM126;               Six lengths of
    DO J = 1 TO 6;                  medians are considered
      DO K = 1 TO 21;     and 21 smoothing constants
        C = (K-1)/20;                     so 126 models are
        IF C = 0 THEN C = .01;                  considered.
        IF C = 1 THEN C = .99;
        IF OBSNUM LT J THEN SM{J,K} = XI;   Initial-
        IF OBSNUM EQ J THEN SM{J,K} = RM{J};  izing.
        IF OBSNUM GT J THEN SM{J,K} = C*RM{J} +
                                        (1-C)*SM{J,K};
        F = LAG(SM{J,K});
        IF OBSNUM LE J THEN GO TO CUT;
        E = XJ - F;                          Manipulating.
        V8 = ABS(E);
        V9 = E**2;
        V11 = ABS(E/XJ);
        V17 = (E/LAGXJ)**2;
```

```
        V18 = ((XJ-LAGXJ)/LAGXJ)**2;
        SUMS{J,K,1} + V8;                          Summing.
        SUMS{J,K,2} + V9;
        SUMS{J,K,3} + V11;
        SUMS{J,K,4} + V17;
        SUMS{J,K,5} + V18;
        IF OBSNUM = NFIT THEN DO;
          MAD = SUMS{J,K,1}/(NFIT-J);   Calculating.
          MSE = SUMS{J,K,2}/(NFIT-J);
          MAPE = SUMS{J,K,3}/(NFIT-J);
          TH3 = SQRT(SUMS{J,K,4}/SUMS{J,K,5});
          FILE SMMEAS;                          Filing.
          PUT SID C J MAD MSE MAPE TH3 SM{J,K};
        END;
      CUT:END;
    END;
    IF OBSNUM = NFIT THEN DO;                Resetting.
      DO J = 1 TO 6;
        DO K = 1 TO 21;
          DO N = 1 TO 5;
            SUMS{J,K,N} = 0;
          END;
        END;
      END;
    END;
    NEXTOBS:;
```

WALSH AVERAGE PROGRAM - FIRST HALF. THIS PART OF THE PROGRAM CALCULATES THE WALSH AVERAGES AND FILES THEM FOR LATER USE.

```
LIBNAME KANK '[RKANKEY.SIX2]';
DATA _NULL_;                        Uses the no monotonic
SET KANK.NOTREND2;                             trend data.
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
W1 = X;
W2 = LAG(X);                                     Renaming.
W3 = LAG2(X);
W4 = LAG3(X);
W5 = LAG4(X);
W6 = LAG5(X);
WA1 = 0; WA2 = 0; WA3 = 0; WA4 = 0; WA5 = 0;
  WA6 = 0;
ARRAY VALUES{6} W1-W6;
ARRAY Z{6,6} Z1-Z36;
ARRAY ORDERING{21} UNORD1-UNORD21;
ARRAY WAL{6} WA1-WA6;
DO J = 1 TO 6;
  DO K = J TO 6;
    Z{J,K} = (VALUES{J} + VALUES{K});
  END;
END;                             This program uses Walsh
WAL{1} = W1;                      averages of length one
WAL{2} = (W1+W2)/2;                         through six.
IF OBSNUM EQ 1 THEN WAL{2} = 0;
DO N = 3 TO 6 BY 1;
  OLDK=1;
  DO J = 1 TO N;
    DO K = J TO N;
      ORDERING{OLDK} = Z{J,K};
      OLDK+1;
    END;
  END;
  IF OBSNUM LT N THEN GO TO PLUSDAT;
  IF N EQ 3 THEN MAX = 6;
  IF N EQ 4 THEN MAX = 10;
  IF N EQ 5 THEN MAX = 15;
  IF N EQ 6 THEN MAX = 21;
  DO L = 1 TO MAX;
    MORE = 0;                                 Bubble sort.
    DO J = 1 TO (MAX-1);
      IF ORDERING{J+1} LT ORDERING{J} THEN DO;
      TEMP = ORDERING{J};
      ORDERING{J} = ORDERING{J+1};
      ORDERING {J+1} = TEMP;
      MORE = 1;
    END;
  END;
  IF MORE EQ 0 THEN GO TO NEXTN;
```

```
    END;
    NEXTN: IF MOD(MAX,2) EQ 1 THEN WAL{N} =
           ORDERING{(OLDK+1)/2}/2;
    ELSE WAL{N} = (ORDERING{OLDK/2} +
           ORDERING{OLDK/2+1})/4;
  END;
  PLUSDAT:;
  FILE WAVGSNF;                                  Filing.
  PUT SID X WA1-WA6;
  NEXTOBS:;
```

WALSH AVERAGE PROGRAM - SECOND HALF. THIS HALF OF THE PROGRAM READS IN THE WALSH AVERAGES, CALCULATES THE FITTING ERROR MEASURES BY SERIES, AND FILES THE RESULTS.

```
LIBNAME KANK '[RKANKEY.SIX2]';
DATA _NULL_;
SET KANK.NOTREND2;
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
INFILE WAVGSNF;
INPUT SID2 X2 WA1 WA2 WA3 WA4 WA5 WA6;
XI = X;
ARRAY WAL{6} WA1 WA2 WA3 WA4 WA5 WA6;
ARRAY SUMS{6,5} S1-S30;
IF OBSNUM EQ 1 THEN DO;                      Setting SUMS
  DO Q = 1 TO 6;                             equal to zero.
    DO R = 1 TO 5;
      SUMS{Q,R} = 0;
    END;
  END;
END;
LAGXI = LAG(XI);
DO J=1 TO 6;
  F = LAG(WAL{J});                              Forecasts.
  N = J;
  E = XI-F;
  IF OBSNUM LT J+1 THEN GO TO CUT;
  V8 = ABS(E);                               Manipulating.
  V9 = E**2;
  V11 = ABS(E/XI);
  V17 = (E/LAGXI)**2;
  V18 = ((XI-LAGXI)/LAGXI)**2;
  SUMS{J,1} + V8;
  SUMS{J,2} + V9;                                Summing.
  SUMS{J,3} + V11;
  SUMS{J,4} + V17;
  SUMS{J,5} + V18;
  IF OBSNUM EQ NFIT THEN DO;                 Calculating.
    MAD = SUMS{J,1}/(NFIT-J);
    MSE = SUMS{J,2}/(NFIT-J);
    MAPE = SUMS{J,3}/(NFIT-J);
    TH3 = SQRT(SUMS{J,4}/SUMS{J,5});
    FILE TEMPWA;                                  Filing.
    PUT SID J MAD MSE MAPE TH3 WAL{J};
  END;
CUT:END;
NEXTOBS:;
```

FITTING PROGRAM TO SELECT BEST PARAMETER FOR EXPONENTIAL SMOOTHING. THIS PROGRAM TAKES THE COMPLETE RESULTS FROM THE EXPONENTIAL SMOOTHING PROGRAM AND FINDS THE SMOOTHING CONSTANTS THAT YIELDED THE BEST FIT FOR MAD, MSE, MAPE, AND THEIL'S U.

```
LIBNAME STUFF '[RKANKEY.SIX2]';
DATA ONE;
INFILE MEAS111D;
INPUT J SID C NFIT MAD MSE MAPE MSPE TH3;
RETAIN;
IF J = 1 THEN DO;
    MSEMIN = MSE;  MSEC = C;
    MADMIN = MAD;  MADC = C;
    MAPEMIN = MAPE;  MAPEC = C;
    MSPEMIN = MSPE;  MSPEC = C;
    TH3C = C;  TH3MIN = TH3;
    END;
ELSE DO;
    IF MSE LT MSEMIN THEN DO;
        MSEMIN = MSE;  MSEC = C;
    END;
    IF MAD LT MADMIN THEN DO;
        MADMIN = MAD;   MADC = C;
    END;
    IF MAPE LT MAPEMIN THEN DO;
        MAPEMIN = MAPE;  MAPEC = C;
    END;
    IF MSPE LT MSPEMIN THEN DO;
        MSPEMIN = MSPE;  MSPEC = C;
    END;
    IF TH3 LT TH3MIN THEN DO;
        TH3MIN = TH3;  TH3C = C;
    END;
END;
IF J = 99 THEN DO;
    FILE PARAMSD;          Filing best constants.
    PUT SID MADC MSEC MAPEC MSPEC TH3C;
END;
```

FITTING PROGRAM TO SELECT THE BEST PARAMETER FOR RUNNING MEDIANS. THIS PROGRAM FINDS THE BEST LENGTHS FOR EACH SERIES CONSIDERED AND FOR EACH OF THE FOUR ERROR MEASURES.

```
LIBNAME STUFF '[RKANKEY.SIX2]';
DATA ONE;
INFILE RMRES;
INPUT SID N MAD MSE MAPE TH3 FORECAST;
RETAIN;
IF N = 1 THEN DO;
    MSEMIN = MSE;  MSEN = N;
    MADMIN = MAD;  MADN = N;
    MAPEMIN = MAPE;  MAPEN = N;
    TH3N = N;  TH3MIN = TH3;
    FMAD=FORECAST;  FMSE=FORECAST;
    FMAPE=FORECAST; FTH3=FORECAST;
    END;
ELSE DO;
    IF MSE LT MSEMIN THEN DO;
        MSEMIN = MSE;  MSEN = N;
        FMSE = FORECAST;
    END;
    IF MAD LT MADMIN THEN DO;
        MADMIN = MAD;   MADN = N;
        FMAD = FORECAST;
    END;
    IF MAPE LT MAPEMIN THEN DO;
        MAPEMIN = MAPE;  MAPEN = N;
        FMAPE = FORECAST;
    END;
    IF TH3 LT TH3MIN THEN DO;
        TH3MIN = TH3;  TH3N = N;
        FTH3 = FORECAST;
    END;
END;
IF N = 6 THEN DO;
    FILE RMTEMP1;
    PUT SID MADN MSEN MAPEN TH3N FMAD FMSE
        FMAPE FTH3 MADMIN MSEMIN MAPEMIN
        TH3MIN;
END;
```

This program files the best lengths and the best fit measure values.

PROGRAM TO CALCULATE FORECASTING ERRORS OVER HORIZONS ONE THROUGH SIX USING THE BEST MAPE MODEL FOR EACH TECHNIQUE OVER EACH SERIES, AND TO ACCUMULATE AND REPORT THE FORECASTING MAPE

```
LIBNAME KANK '[RKANKEY.SIX2]';
DATA ONE;
SET KANK.BYSID8;
ARRAY DEVS{5,6} DEVM1-DEVM6 DEVE1-DEVE6 DEVR1-
      DEVR6 DEVS1-DEVS6 DEVW1-DEVW6;
ARRAY PD{5,6} PD1-PD30;
ARRAY MAPEFIT{5} MAPEMFIT MAPEEFIT MAPERFIT
      MAPESFIT MAPEWFIT;
ARRAY XVALS{7} XNFIT XNOF1-XNOF6;
ARRAY SUMPDEV{5,6} SUM1-SUM30;
ARRAY SMAPEFIT{5} SMAPE1-SMAPE5;
ARRAY AMAPEF{5,6} AMAPEF1-AMAPEF30;
ARRAY AMAPEFT{5} AMAPEFT1-AMAPEFT5;
DEVM1 = XNOF1 - FMMAPE;
DEVM2 = XNOF2 - FMMAPE;      Deviations from use
DEVM3 = XNOF3 - FMMAPE;    of moving average for
DEVM4 = XNOF4 - FMMAPE;    horizons 1 through 6.
DEVM5 = XNOF5 - FMMAPE;
DEVM6 = XNOF6 - FMMAPE;
DEVE1 = XNOF1 - FEMAPE;
DEVE2 = XNOF2 - FEMAPE;      Deviations from use
DEVE3 = XNOF3 - FEMAPE;           of exponential
DEVE4 = XNOF4 - FEMAPE;   smoothing for horizons
DEVE5 = XNOF5 - FEMAPE;             1 through 6.
DEVE6 = XNOF6 - FEMAPE;
DEVR1 = XNOF1 - FRMAPE;
DEVR2 = XNOF2 - FRMAPE;          Deviations using
DEVR3 = XNOF3 - FRMAPE;          running medians.
DEVR4 = XNOF4 - FRMAPE;
DEVR5 = XNOF5 - FRMAPE;
DEVR6 = XNOF6 - FRMAPE;
DEVS1 = XNOF1 - FSMAPE;
DEVS2 = XNOF2 - FSMAPE;          Deviations using
DEVS3 = XNOF3 - FSMAPE;         smoothed medians.
DEVS4 = XNOF4 - FSMAPE;
DEVS5 = XNOF5 - FSMAPE;
DEVS6 = XNOF6 - FSMAPE;
DEVW1 = XNOF1 - FWMAPE;
DEVW2 = XNOF2 - FWMAPE;          Deviations using
DEVW3 = XNOF3 - FWMAPE;           Walsh averages.
DEVW4 = XNOF4 - FWMAPE;
DEVW5 = XNOF5 - FWMAPE;
DEVW6 = XNOF6 - FWMAPE;
DO J = 1 TO 5;
  SMAPEFIT{J} + MAPEFIT{J};  Summing the fitting
  DO K = 1 TO 6;                           MAPEs.
    PD{J,K} = ABS((DEVS{J,K})/(XVALS{K+1}));
```

```
    SUMPDEV{J,K} + PD{J,K};      Calculating and
  END;                         summing the forecasting
END;                                             MAPEs.
IF SID LT 994 THEN GO TO NEXTOBS;
DO J = 1 TO 5;
  AMAPEFT{J} = SMAPEFIT{J}/19; Avg fitting MAPE.
  DO K = 1 TO 6;
    AMAPEF{J,K} = SUMPDEV{J,K}/19;        Average
  END;                                 forecasting
END;                                         MAPEs.
PUT AMAPEFT1 AMAPEF1-AMAPEF6;             Average
      fitting MAPE, and average forecasting MAPE
      for horizons one through six using moving
      averages.  PUT into the log.
PUT AMAPEFT2 AMAPEF7-AMAPEF12;    Ex. Smoothing.
PUT AMAPEFT3 AMAPEF13-AMAPEF18; Running Medians.
PUT AMAPEFT4 AMAPEF19-AMAPEF24;    Smoothed Mds.
PUT AMAPEFT5 AMAPEF25-AMAPEF30;      Walsh Avgs.
NEXTOBS:;
```

PROGRAM TO CALCULATE THE MEDIANS OF THE THEIL'S U VALUES. SINCE THEIL'S U VALUES WERE NOT SYMMETRIC, RESULTS BY THEIL'S U WERE FELT TO BE BEST DISPLAYED THROUGH USE OF THE MEDIAN THEIL'S U VALUES RATHER THAN THE AVERAGES.

```
LIBNAME KANK '[RKANKEY.SIX2]';
DATA ONE;
INFILE TH3CK;
INPUT VAR1-VAR30;        Six horizon values times
PROC UNIVARIATE NOPRINT;          five techniques.
VAR VAR1-VAR30;
OUTPUT OUT=KANK.Z5MEDRES Q1=LOWER1-LOWER30
       MEDIAN=MED1-MED30 Q3=UPPER1-UPPER30
       MAX=MAX1-MAX30;
```

DOUBLE MOVING AVERAGE PROGRAM - FIRST HALF. THIS PROGRAM IS APPLIED TO THE 92 SERIES THAT DID TEST TO DEMONSTRATE MONOTONIC TREND OVER THE FIT DATA. THESE SERIES ARE ALL STORED IN THE FILE TREND2.SSD. THIS HALF OF THE PROGRAM CALCULATES THE SINGLE AND DOUBLE MOVING AVERAGES FOR EACH APPROPRIATE OBSERVATION OF EACH SERIES AND FILES THEM.

```
LIBNAME STAT '[RKANKEY.SIX2]';
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA _NULL_;
SET STAT.TREND2;
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
ARRAY MA{10} MA1-MA10;              Moving averages.
ARRAY DMA{10} DMA1-DMA10;        Double moving avgs.
ARRAY FC{10} FC1-FC10;                    Forecasts.
ARRAY SUMS{10,5} S1-S50;
LAGX = LAG(X);
L1 = LAGX;
L2 = LAG2(X);                              Renaming.
L3 = LAG3(X);
L4 = LAG4(X);
L5 = LAG5(X);
L6 = LAG6(X);
L7 = LAG7(X);
L8 = LAG8(X);
L9 = LAG9(X);
SUM = X;
MA1 = SUM;
SUM = SUM+L1;                       Calculating the
MA2 = SUM/2;                         single moving
SUM = SUM+L2;                            averages.
MA3 = SUM/3;
SUM = SUM+L3;
MA4 = SUM/4;
SUM = SUM+L4;
MA5 = SUM/5;
SUM = SUM+L5;
MA6 = SUM/6;
SUM = SUM+L6;
MA7 = SUM/7;
SUM = SUM+L7;
MA8 = SUM/8;
SUM = SUM+L8;
MA9 = SUM/9;
SUM = SUM+L9;
MA10 = SUM/10;
SUM = 0;
DMA1 = MA{1};      Calculating the double moving
DMA2 = (MA{2} + LAG(MA{2}))/2;           averages.
DMA3 = (MA{3} + LAG(MA{3}) + LAG2(MA{3}))/3;
DMA4 = (MA{4}+LAG(MA{4})+ LAG2(MA{4}) +
```

```
                LAG3(MA{4}))/4;
        DMA5 = (MA{5} + LAG(MA{5}) + LAG2(MA{5}) +
                LAG3(MA{5}) + LAG4(MA{5}))/5;
        DMA6 = (MA{6} + LAG(MA{6}) + LAG2(MA{6}) +
                LAG3(MA{6}) + LAG4(MA{6}) +
                LAG5(MA{6}))/6;
        DMA7 = (MA{7} + LAG(MA{7}) + LAG2(MA{7}) +
                LAG3(MA{7})  + LAG4(MA{7}) +
                LAG5(MA{7}) + LAG6(MA{7}))/7;
        DMA8 = (MA{8} + LAG(MA{8}) + LAG2(MA{8}) +
                LAG3(MA{8}) + LAG4(MA{8}) + LAG5(MA{8})
                + LAG6(MA{8}) + LAG7(MA{8}))/8;
        DMA9 = (MA{9} + LAG(MA{9}) + LAG2(MA{9}) +
                LAG3(MA{9}) +  LAG4(MA{9}) + LAG5(MA{9})
                + LAG6(MA{9}) + LAG7(MA{9})+
                LAG8(MA{9}))/9;
        DMA1Ø = (MA{1Ø} + LAG(MA{1Ø}) + LAG2(MA{1Ø}) +
                 LAG3(MA{1Ø}) + LAG4(MA{1Ø}) +
                 LAG5(MA{1Ø}) + LAG6(MA{1Ø}) +
                 LAG7(MA{1Ø}) + LAG8(MA{1Ø}) +
                 LAG9(MA{1Ø})) / 1Ø;
        DO K = 1 TO 1Ø;
          IF OBSNUM LT K THEN DO;        Cleaning up some
            MA{K} = Ø;                 missing values for
          END;                          the next program.
          IF OBSNUM LT 2*K-1 THEN DO;
            DMA{K} = Ø;
          END;
        END;
        FILE MOVAVGS;                             Filing.
        PUT SID OBSNUM X MA1-MA1Ø DMA1-DMA1Ø;
        NEXTOBS:;
```

DOUBLE MOVING AVERAGE PROGRAM - SECOND HALF. THIS PART OF THE PROGRAM READS IN THE AVERAGES, CALCULATES THE APPROPRIATE MAXIMUM LENGTH AVERAGE TO USE, THEN FINDS THE FITTING ERROR MEASURE RESULTS FOR MOVING AVERAGES OF LENGTH ONE THROUGH TEN.

```
LIBNAME STAT '[RKANKEY.SIX2]';
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA _NULL_;
SET STAT.TREND2;                    Reading the raw data.
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
INFILE MOVAVGS;                     Reading the averages.
INPUT SID2 OBSNUM2 X2 MA1-MA10 DMA1-DMA10;
RETAIN;
ARRAY MA{10} MA1-MA10;                     Setting up the
ARRAY DMA{10} DMA1-DMA10;                          arrays.
ARRAY FC{10} FC1-FC10;
ARRAY SUMS{10,5} S1-S50;
IF OBSNUM = 1 THEN DO;                    Determining the
  DMASIZE = 10;                    desired maximum length
  IF NFIT/2 LT 20 THEN DO;                averages to be
  LENGTH = CEIL(NFIT/2);                      considered.
  DMASIZE = CEIL(LENGTH/2);
  END;
END;
LAGX = LAG(X);
DO J = 1 TO DMASIZE;
  N = J;
  IF OBSNUM LT 2*N-1 THEN DO;   Getting started.
    AT = 0;
    BT = 0;
    GO TO SKIP;
  END;
  IF N EQ 1 THEN DO;
    AT = X;
    BT = X-LAGX;
    GO TO SKIP;
  END;
  AT = 2*MA{J} - DMA{J};                     Initializing.
  BT = (2/(N-1))*(MA{J}-DMA{J});
  SKIP:;
  FC{J} = AT + BT;
  F = LAG(FC{J});                             Forecasting.
  IF N EQ 1 THEN DO;
    IF OBSNUM LE 2 THEN GO TO CUT;
    END;
  ELSE IF OBSNUM LT 2*J THEN GO TO CUT;
  E = X-F;                                Manipulating the
  V8 = ABS(E);                                 error term.
  V9 = E**2;
  V11 = ABS(E/X);
  V17 = (E/LAGX)**2;
```

```
    V18 = ((X-LAGX)/LAGX)**2;
    SUMS{J,1} + V8;                    Summing over the
    SUMS{J,2} + V9;                             series.
    SUMS{J,3} + V11;
    SUMS{J,4} + V17;
    SUMS{J,5} + V18;
    IF OBSNUM EQ NFIT THEN DO;
      IF N = 1 THEN DEN = NFIT-2*N;
      ELSE DEN = NFIT-2*N+1;              Calculating
      MAD = SUMS{J,1}/DEN;               series error
      MSE = SUMS{J,2}/DEN;                  measures.
      MAPE = SUMS{J,3}/DEN;
      TH3 = SQRT(SUMS{J,4}/SUMS{J,5});
      FILE RESDMA;                            Filing.
      PUT SID N NFIT DMASIZE MAD MSE MAPE TH3
          AT BT;
    END;
  CUT:END;

  IF OBSNUM EQ NFIT THEN DO;          Resetting for
    DO M = 1 TO 10;            for the next series.
      DO N = 1 TO 5;
        SUMS{M,N} = 0;
      END;
    END;
  END;
  NEXTOBS:;
```

PROGRAM TO GET THE LEAST-SQUARES, BEST-FIT INTERCEPT AND SLOPE OVER THE FIRST SIX OBSERVATIONS OF EACH OF THE 111 SERIES, USED TO START THE LINEAR EXPONENTIAL SMOOTHING TECHNIQUE.

```
LIBNAME STUFF '[RKANKEY.SIX2]';
DATA ONE;
SET STUFF.DS111;
IF OBSNUM GT 6 THEN GO TO RDK;
RETAIN;
Y = X;
SUMY + Y;
SUMIY + (Y*I);
IF OBSNUM = 6 THEN DO;
  YBAR = SUMY/6;
  LSBFB = (SUMIY-21*YBAR)/17.5;
  LSBFA = YBAR -(LSBFB*3.5);
  FILE LSBFAB;
  PUT SID LSBFA LSBFB;
  SUMY = 0;
  SUMIY = 0;
END;
RDK:;
```

Just an implementation of the least-squares, best fit formulas where the independent variable values are known and constant. For comparison with other references, the dependent variable is renamed from X to Y.

LINEAR EXPONENTIAL SMOOTHING PROGRAM. THIS PROGRAM CALCULATES THE VARIOUS ERROR MEASURES ACHIEVED BY LINEAR EXPONENTIAL SMOOTHING MODELS WITH SMOOTHING CONSTANTS OF .01, .05 TO .95 BY .05, AND .99 OVER THE SET OF NINETY TWO MONOTONIC TREND SERIES.

```
LIBNAME STAT '[RKANKEY.SIX2]';
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA _NULL_;
SET STAT.TREND2;
RETAIN;
IF OBSNUM = 1 THEN DO;
  ONEMORE:;
  INFILE LSBFAB;          Reading in the LSBF values
  INPUT SID2 LSBFA LSBFB;     for initialization.
  IF SID GT SID2 THEN GO TO ONEMORE;      Making
END;                    sure the proper 92 are read.
IF OBSNUM GT NFIT THEN GO TO RDK;
ARRAY EX{21} EX1-EX21;                Setting up the
ARRAY E2X{21} E2X1-E2X21;                    arrays.
ARRAY FA{21} F1-F21;
ARRAY SUMS{21,5} S1-S105;
LAGX = LAG(X);

DO J = 1 TO 21;
  C = (J-1)*.05;
  IF C = 0 THEN C = .01;                Initializing
  IF C = 1 THEN C = .99;                    section.
  IF OBSNUM = 1 THEN DO;
    EX0 = LSBFA - ((LSBFB*(1-C))/C);
    E2X0 = LSBFA -((2*LSBFB*(1-C))/C);
    EX{J} = (C*X)+(1-C)*EX0;
    E2X{J} = C*EX{J} + (1-C)*E2X0;
  END;
  IF OBSNUM GT 1 THEN DO;
    EX{J} = C*X + (1-C)*EX{J};
    E2X{J} = C*EX{J} + (1-C)*E2X{J};
  END;


  AT = 2*EX{J} - E2X{J};
  BT = (C/(1-C))*(EX{J} - E2X{J});
  FA{J} = AT + BT;
  F = LAG(FA{J});                       Forecasting.
  IF OBSNUM = 1 THEN GO TO SKIP;
  E = X-F;


  V8 = ABS(E);                       Manipulating the
  V9 = E**2;                                error term.
  V11 = ABS(E/X)*100;
  V17 = (E/LAGX)**2;
```

```
    V18 = ((X-LAGX)/LAGX)**2;
    SUMS{J,1} + V8;                           Summing over
    SUMS{J,2} + V9;                            the series.
    SUMS{J,3} + V11;
    SUMS{J,4} + V17;
    SUMS{J,5} + V18;

    IF OBSNUM = NFIT THEN DO;                 Calculating.
      MAD = SUMS{J,1}/(NFIT-1);
      MSE = SUMS{J,2}/(NFIT-1);
      MAPE = SUMS{J,3}/(NFIT-1);
      TH3 = SQRT(SUMS{J,4}/SUMS{J,5});
      FILE DEARES;                                 Filing.
      PUT J SID C RSCORR NFIT MAD MSE MAPE TH3
          AT BT;
    END;

    SKIP:;

  END;
  IF OBSNUM = NFIT THEN DO;                 Resetting for
      DO M = 1 TO 21;                     the next series.
        DO N = 1 TO 5;
          SUMS{M,N} = 0;
        END;
      END;
  END;
  RDK:;
```

DOUBLE RUNNING MEDIAN PROGRAM - FIRST PART. THIS PART OF THE PROGRAM CALCULATES THE SINGLE RUNNING MEDIANS OF LENGTHS ONE THROUGH SIX AND FILES THEM.

```
LIBNAME STAT '[RKANKEY.SIX2]';
DATA _NULL_;
SET STAT.TREND2;
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
XI = X;
LAGX = LAG(X);                                 Renaming.
L1=LAGX;
L2=LAG2(X);
L3=LAG3(X);
L4=LAG4(X);
L5=LAG5(X);
ARRAY TEMP3{3} X L1 L2;                   Setting up the
ARRAY TEMP4{4} X L1 L2 L3;                        arrays.
ARRAY TEMP5{5} X L1 L2 L3 L4;
ARRAY TEMP6{6} X L1 L2 L3 L4 L5;
RM1 = X;                                        Defining/
IF OBSNUM LT 2 THEN GO TO OBSOUT;             calculating
RM2 = (X+LAGX)/2;                             the running
IF OBSNUM LT 3 THEN GO TO OBSOUT;                medians.
DO H=1 TO 3;
  MORE=0;
  DO I = 1 TO 2;
    IF TEMP3{I+1} LT TEMP3{I} THEN DO;
      TEMP = TEMP3{I};
      TEMP3{I} = TEMP3{I+1};     Bubble sorts for
      TEMP3{I+1} = TEMP;       medians of lengths
      MORE = 1;                three through six.
    END;
  END;
  IF MORE = 0 THEN GO TO DONE3;
END;
DONE3:;
RM3 = TEMP3{2};
IF OBSNUM LT 4 THEN GO TO OBSOUT;
DO H = 1 TO 4;
  MORE = 0;
  DO I = 1 TO 3;
    IF TEMP4{I+1} LT TEMP4{I} THEN DO;
      TEMP = TEMP4{I};
      TEMP4{I} = TEMP4{I+1};
      TEMP4{I+1} = TEMP;
      MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE4;
END;
DONE4:;
RM4 = (TEMP4{2} + TEMP4{3})/2;
```

```
IF OBSNUM LT 5 THEN GO TO OBSOUT;
DO H = 1 TO 5;
  MORE = 0;
  DO I = 1 TO 4;
    IF TEMP5{I+1} LT TEMP5{I} THEN DO;
    TEMP = TEMP5{I};
    TEMP5{I} = TEMP5{I+1};
    TEMP5{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE5;
END;
DONE5:;
RM5 = TEMP5{3};
IF OBSNUM LT 6 THEN GO TO OBSOUT;
DO H = 1 TO 6;
  MORE = 0;
  DO I = 1 TO 5;
    IF TEMP6{I+1} LT TEMP6{I} THEN DO;
    TEMP = TEMP6{I};
    TEMP6{I} = TEMP6{I+1};
    TEMP6{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE6;
END;
DONE6:;
RM6 = (TEMP6{3} + TEMP6{4})/2;
OBSOUT:;
FILE MEDIANS;                  Filing the medians.
PUT SID OBSNUM XI RM1-RM6;
NEXTOBS:;
```

DOUBLE RUNNING MEDIANS PROGRAM - SECOND PART. THIS PART OF THE PROGRAM READS THE SINGLE RUNNING MEDIANS, CALCULATES THE DOUBLE RUNNING MEDIANS, AND FILES BOTH FOR LATER USE.

```
DATA _NULL_;
INFILE MEDIANS;
INPUT SID OBSNUM XI RMI1-RMI6;
RM1 = RMI1;                                    Renaming.
RM2 = RMI2;
RM3 = RMI3;
RM4 = RMI4;
RM5 = RMI5;
RM6 = RMI6;
ARRAY RM{6} RMI1 RMI2 RMI3 RMI4 RMI5 RMI6;
ARRAY FC{6} FC1-FC6;
LRM2 = LAG(RM2);                               Renaming.
LT1 = LAG(RM3);
LT2 = LAG2(RM3);
LF1 = LAG(RM4);
LF2 = LAG2(RM4);
LF3 = LAG3(RM4);
LFI1 = LAG(RM5);
LFI2 = LAG2(RM5);
LFI3 = LAG3(RM5);
LFI4 = LAG4(RM5);
LS1 = LAG(RM6);
LS2 = LAG2(RM6);
LS3 = LAG3(RM6);
LS4 = LAG4(RM6);
LS5 = LAG5(RM6);
ARRAY SC3{3} RM3 LT1 LT2;         Setting up arrays
ARRAY SC4{4} RM4 LF1-LF3;       of running medians.
ARRAY SC5{5} RM5 LFI1-LFI4;
ARRAY SC6{6} RM6 LS1-LS5;
DRM1 = RM1;
IF OBSNUM LT 3 THEN GO TO SKIP;
DRM2 = (RM2+LRM2)/2;
IF OBSNUM LT 5 THEN GO TO SKIP;
DO H = 1 TO 3;                  Sorting the running
MORE = 0;                       medians to find the
DO I = 1 TO 2;              double running medians.
  IF SC3{I+1} LT SC3{I} THEN DO;
    TEMP = SC3{I};
    SC3{I} = SC3{I+1};
    SC3{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE3;
END;
DONE3:;                        Double running median of
DRM3 = SC3{2};                             length three.
```

```
IF OBSNUM LT 7 THEN GO TO SKIP;
DO H = 1 TO 4;
MORE = 0;
DO I = 1 TO 3;
  IF SC4{I+1} LT SC4{I} THEN DO;
    TEMP = SC4{I};
    SC4{I} = SC4{I+1};
    SC4{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE4;
END;
DONE4:;
DRM4 = (SC4{2} + SC4{3})/2;  DRM of length four.
IF OBSNUM LT 9 THEN GO TO SKIP;
DO H = 1 TO 5;
MORE = 0;
DO I = 1 TO 4;
  IF SC5{I+1} LT SC5{I} THEN DO;
    TEMP = SC5{I};
    SC5{I} = SC5{I+1};
    SC5{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE5;
END;
DONE5:;
DRM5 = SC5{3};               DRM of length five.
IF OBSNUM LT 11 THEN GO TO SKIP;
DO H = 1 TO 6;
MORE = 0;
DO I = 1 TO 5;
  IF SC6{I+1} LT SC6{I} THEN DO;
    TEMP = SC6{I};
    SC6{I} = SC6{I+1};
    SC6{I+1} = TEMP;
    MORE = 1;
    END;
  END;
  IF MORE = 0 THEN GO TO DONE6;
END;
DONE6:;                            DRM of length
DRM6 = (SC6{3} + SC6{4})/2;                 six.
SKIP:;
ARRAY DRM{6} DRM1-DRM6;
DO K = 1 TO 6;
  IF OBSNUM LT 2*K-1 THEN DRM{K} = 0;
END;
FILE MEDSALL;                            Filing.
PUT SID OBSNUM XI RMI1-RMI6 DRM1-DRM6;
```

DOUBLE RUNNING MEDIANS PROGRAM - THIRD PART. THIS PART OF THE PROGRAM READS THE SINGLE AND DOUBLE RUNNING MEDIANS, MAKES THE FORECASTS, AND RECORDS THE FITTING PERFORMANCE FOR EACH LENGTH MEDIAN BY SERIES.

```
LIBNAME STAT '[RKANKEY.SIX2]';
DATA _NULL_;
SET STAT.TREND2;              Reading the raw data.
RETAIN;
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
INFILE MEDSALL;                 Reading the medians.
INPUT SID2 OBSNUM2 X2 RM1-RM6 DRM1-DRM6;
IF SID NE SID2 THEN STOP;
ARRAY RM{6} RM1-RM6;                       Setting up
ARRAY DRM{6} DRM1-DRM6;                   the arrays.
ARRAY FC{6} FC1-FC6;
ARRAY SUMS{6,5} S1-S30;
LAGX = LAG(X);
IF OBSNUM = 1 THEN DO;              Calculating the
  DRMSIZE = 6;                maximum length medians
  IF NFIT/2 LT 12 THEN DO;       to be considered.
    LENGTH = CEIL(NFIT/2);
    DRMSIZE = CEIL(LENGTH/2);
    END;
END;
DO J=1 TO DRMSIZE;
  N = J;
  IF OBSNUM LT 2*N-1 THEN DO;
    AT = 0;
    BT = 0;
    GO TO SKIP;
  END;
  IF N EQ 1 THEN DO;
    AT = X;
    BT = X-LAGX;                            Calculating
    GO TO SKIP;                            the line for
  END;                                     forecasting,
  AT = 2*RM{J} - DRM{J};              i.e. AT and BT.
  BT = (2/(N-1))*(RM{J} -DRM{J});
  SKIP:;
  FC{J} = AT +BT;
  F = LAG(FC{J});                          Forecasting.
  E = X-F;
  IF N EQ 1 THEN DO;
    IF OBSNUM LE 2 THEN GO TO CUT;
    END;
  ELSE IF OBSNUM LT 2*N THEN GO TO CUT;
  V8 = ABS(E);
  V9 = E**2;                               Manipulating
  V11 = ABS(E/X);                       the error term.
```

```
        V17 = (E/LAGX)**2;
        V18 = ((X-LAGX)/LAGX)**2;
        SUMS{J,1} + V8;
        SUMS{J,2} + V9;                       Summing over
        SUMS{J,3} + V11;                       the series.
        SUMS{J,4} + V17;
        SUMS{J,5} + V18;
        IF OBSNUM = NFIT THEN DO;
          IF N = 1 THEN DEN = NFIT-2*N;
          ELSE DEN = NFIT-2*N+1;         Calculating the
          MAD = SUMS{J,1}/DEN;            error measures
          MSE = SUMS{J,2}/DEN;           for the series.
          MAPE = SUMS{J,3}/DEN;
          TH3 = SQRT(SUMS{J,4}/SUMS{J,5});
          FILE RESDRM;                           Filing.
          PUT SID N DRMSIZE MAD MSE MAPE TH3 AT BT;
        END;
      CUT:END;

      IF OBSNUM = NFIT THEN DO;             Resetting for
        DO M = 1 TO 6;                    the next series.
          DO N = 1 TO 5;
            SUMS{M,N} = 0;
          END;
        END;
      END;
      NEXTOBS:;
```

DOUBLE SMOOTHED MEDIAN PROGRAM. THIS PROGRAM TAKES ADVANTAGE OF PREVIOUS WORK BY CALLING IN THE LSBF VALUES FROM A PREVIOUSLY DEVELOPED FILE, AND BY CALLING IN THE RUNNING MEDIANS THAT WERE ALSO PREVIOUSLY CALCULATED. SMOOTHING CONSTANTS OF .05 TO .95 BY .05 ARE CONSIDERED.

```
LIBNAME STAT '[RKANKEY.SIX2]';
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA _NULL_;
SET STAT.TREND2;
RETAIN;
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
IF OBSNUM = 1 THEN DO;
  ONEMORE:;
  INFILE LSBFAB;
  INPUT SID3 LSBFA LSBFB;
  IF SID GT SID3 THEN GO TO ONEMORE;     Assures
  DROP SID3;                 proper observations are
  LENGTH = 6;                      used from LSBFAB.
  IF NFIT/2 LT 6 THEN DO;
    LENGTH = CEIL(NFIT/2);   Here DMASIZE refers
    DMASIZE = CEIL(LENGTH/2);      to the length
  END;              running medians considered.  A
END;                very short series should not be
INFILE MEDSALL;      forcast with a long median.
INPUT SID2 OBSNUM2 X2 RM1-RM6 DRM1-DRM6;
IF SID NE SID2 THEN STOP;
ARRAY SUMS{6,19,5} S1-S570;
LAGX = LAG(X);
ARRAY RM{6} RM1-RM6;                     Setting up
ARRAY SM{6,19} SM1-SM114;               the arrays.
ARRAY DSM{6,19} DSM1-DSM114;
ARRAY FI{6,19} F1-F114;
DO K=1 to LENGTH;
  N = K;
  DO J = 1 TO 19;
    C = .05*J;
    IF OBSNUM = 1 THEN DO;         Initializing.
      IF N = 1 THEN DO;
        SM0 = LSBFA-LSBFB*(1-C)/C;
        DSM0 = LSBFA - ((2*LSBFB*(1-C))/C);
        SM{K,J} = C*RM{K} + (1-C)*SM0;
        DSM{K,J} = C*SM{K,J} + (1-C)*DSM0;
      END;
    IF N EQ 2 THEN DO;
      DSM{K,J} = LSBFA+K*LSBFB - 2*(LSBFB*
                   (1-C)/C);
      SM{K,J} = DSM{K,J} + LSBFA*(1-C)/C;
    END;
    IF N GT 2 THEN DO;
      SM{K,J} = 0;
      DSM{K,J} = 0;
```

```
      END;
      GO TO ENDNEQ1;
    END;
    IF 1 LT OBSNUM LT N-1 THEN DO;
      SM{K,J} = 0;
      DSM{K,J} = 0;
    END;
    IF OBSNUM = N-1 THEN DO;
      DSM{K,J} = LSBFA+K*LSBFB - 2*(LSBFB*
                 (1-C)/C);
      SM{K,J} = DSM{K,J} + LSBFA*(1-C)/C;
    END;
    IF OBSNUM GE N THEN DO;   Recursive equations.
      SM{K,J} = C*RM{K} + (1-C)*SM{K,J};
      DSM{K,J} = C*SM{K,J} + (1-C)*DSM{K,J};
    END;
    ENDNEQ1:;
    AT = 2*SM{K,J} - DSM{K,J};
    BT = (C/(1-C))*(SM{K,J} - DSM{K,J});
    FI{K,J} = AT + ((N+1)/2)*BT;
    F = LAG(FI{K,J});                    Forecasting.
    IF OBSNUM LT N+1 THEN GO TO CUT;
    E = X-F;
    V8 = ABS(E);                    Manipulating the
    V9 = E**2;                          error terms.
    V11 = ABS(E/X)*100;
    V17 = (E/LAGX)**2;
    V18 = ((X-LAGX)/LAGX)**2;
    SUMS{K,J,1} + V8;
    SUMS{K,J,2} + V9;                   Summing over
    SUMS{K,J,3} + V11;                   the series.
    SUMS{K,J,4} + V17;
    SUMS{K,J,5} + V18;
    IF OBSNUM = NFIT THEN DO;
      DEN = NFIT - N;          Calculating fitting
      MAD = SUMS{K,J,1}/DEN;       errors for the
      MSE = SUMS{K,J,2}/DEN;              series.
      MAPE = SUMS{K,J,3}/DEN;
      TH3 = SQRT(SUMS{K,J,4}/SUMS{K,J,5});
      FILE RESDSM1;
      PUT SID N C MAD MSE MAPE TH3 AT BT;
    END;
  CUT:END;
  END;
```

```
IF OBSNUM = NFIT THEN DO;            Resetting for
  DO M = 1 TO 6;                   the next series.
   DO P = 1 TO 19;
    DO N = 1 TO 5;
      SUMS{M,P,N} = 0;
    END;
   END;
  END;
END;
NEXTOBS:;
```

ROBUST REGRESSION PROGRAM - FIRST HALF. THIS HALF OF THE PROGRAM CALCULATES STATISTICS FOR ROBUST REGRESSION LINES FOR LENGTHS OF THREE THROUGH SIX, CALCULATES THE FORECASTS AT EACH OBSERVATION, AND STORES RESULTS IN FILES.

```
LIBNAME STAT '[RKANKEY.SIX2]';
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA _NULL_;
SET STAT.TREND2;
IF OBSNUM GT NFIT THEN GO TO NEXTOBS;
L1 = LAG(X);                                  Renaming.
L2 = LAG2(X);
L3 = LAG3(X);
L4 = LAG4(X);
L5 = LAG5(X);
XI = X;
IF OBSNUM LT 3 THEN GO TO NEXTOBS;
ARRAY LAGS{6} L5 L4 L3 L2 L1 X;
ARRAY SLOPE{15} S1-S15;
L = 0;                                        Finding the
DO I = 1 TO 5;                                    slopes.
DO J = I+1 TO 6;
  L+1;
  SLOPE{L} = (LAGS{J} - LAGS{I})/(J-I);
END;
END;
ARRAY SL3{3} S13 S14 S15;
DO I = 1 TO 3;
REDO = 0;
  DO J = 1 TO 2;
    IF SL3{J+1} LT SL3{J} THEN DO;
      TEMP = SL3{J};
      SL3{J} = SL3{J+1};
      SL3{J+1} = TEMP;
      REDO = 1;
    END;
  END;
  IF REDO = 0 THEN GO TO DONESL3;
END;
DONESL3:;
SLOPE3 = SL3{2};
IF OBSNUM LT 4 THEN GO TO INT3;
ARRAY SL4{6} S10-S15;
DO I = 1 TO 6;
REDO = 0;
  DO J = 1 TO 5;
    IF SL4{J+1} LT SL4{J} THEN DO;
      TEMP = SL4{J};
      SL4{J} = SL4{J+1};
      SL4{J+1} = TEMP;
      REDO = 1;
    END;
```

```
    END;
    IF REDO = 0 THEN GO TO DONESL4;
  END;
  DONESL4:;
  SLOPE4 = (SL4{3} + SL4{4})/2;
  IF OBSNUM LT 5 THEN GO TO INT3;
  ARRAY SL5{10} S6-S15;
  DO I = 1 TO 10;
  REDO = 0;
    DO J = 1 TO 9;
      IF SL5{J+1} LT SL5{J} THEN DO;
        TEMP = SL5{J};
        SL5{J} = SL5{J+1};
        SL5{J+1} = TEMP;
        REDO = 1;
      END;
    END;
    IF REDO = 0 THEN GO TO DONESL5;
  END;
  DONESL5:;
  SLOPE5 = (SL5{5} + SL5{6})/2;
  IF OBSNUM LT 6 THEN GO TO INT3;
  DO I = 1 TO 15;
    DO J = 1 TO 14;
      IF SLOPE{J+1} LT SLOPE{J} THEN DO;
        TEMP = SLOPE{J};
        SLOPE{J} = SLOPE{J+1};
        SLOPE{J+1} = TEMP;
        REDO = 1;
      END;
    END;
    IF REDO = 0 THEN GO TO DONESL6;
  END;
  DONESL6:;
  SLOPE6 = (SLOPE{8});         Array B has slopes for
  ARRAY B{4} SLOPE3-SLOPE6;       robust regression
  INT3:;                            of lengths three
  ARRAY TEMP3{3} L2 L1 X;                through six.
  DO I = 1 TO 3;
    REDO = 0;              Finding the "intercepts".
    DO J = 1 TO 2;
      IF TEMP3{J+1} LT TEMP3{J} THEN DO;
        TEMP = TEMP3{J};
        TEMP3{J} = TEMP3{J+1};
        TEMP3{J+1} = TEMP;
        REDO = 1;
      END;
    END;
    IF REDO = 0 THEN GO TO DONE3;
  END;
  DONE3:;
  A3 = TEMP3{2};
```

```
IF OBSNUM LT 4 THEN GO TO KEEP;
ARRAY TEMP4{4} L3 L2 L1 X;
DO I = 1 TO 4;
  REDO = 0;
  DO J = 1 TO 3;
    IF TEMP4{J+1} LT TEMP4{J} THEN DO;
      TEMP = TEMP4{J};
      TEMP4{J} = TEMP4{J+1};
      TEMP4{J+1} = TEMP;
      REDO = 1;
    END;
  END;
  IF REDO = 0 THEN GO TO DONE4;
END;
DONE4:;
A4 = (TEMP4{2} + TEMP4{3})/2;
IF OBSNUM LT 5 THEN GO TO KEEP;
ARRAY TEMP5{5} L4 L3 L2 L1 X;
DO I = 1 TO 5;
  REDO = 0;
  DO J = 1 TO 4;
    IF TEMP5{J+1} LT TEMP5{J} THEN DO;
      TEMP = TEMP5{J};
      TEMP5{J} = TEMP5{J+1};
      TEMP5{J+1} = TEMP;
      REDO = 1;
    END;
    IF REDO = 0 THEN GO TO DONE5;
  END;
  IF REDO = 0 THEN GO TO DONE5;
END;
DONE5:;
A5 = TEMP5{3};
IF OBSNUM LT 6 THEN GO TO KEEP;
ARRAY TEMP6{6} L5 L4 L3 L2 L1 X;
DO I = 1 TO 6;
  REDO = 0;
  DO J = 1 TO 5;
    IF TEMP6{J+1} LT TEMP6{J} THEN DO;
      TEMP = TEMP6{J};
      TEMP6{J} = TEMP6{J+1};
      TEMP6{J+1} = TEMP;
      REDO = 1;
    END;
    IF REDO = 0 THEN GO TO DONE6;
  END;
  IF REDO = 0 THEN GO TO DONE6;
END;
DONE6:;
A6 = (TEMP6{3} + TEMP6{4})/2;
KEEP:;
ARRAY A{4} A3-A6;   Array with the "intercepts".
```

```
ARRAY FI{4} FI1-FI4;
DO I = 1 TO 4;
  H = I+3;
  FI{I} = A{I} + (H/2)*B{I};      Forecasts over
END;                                the fit data.
FILE ROBUST1;                 Filing the forecasts.
PUT SID NFIT OBSNUM XI FI1-FI4;
IF OBSNUM = NFIT THEN DO;
  FILE RRAB;    Filing the slope and "intercept"
  PUT SID OBSNUM A3-A6 SLOPE3-SLOPE6;    for the
END; last           fitting observation, used to
NEXTOBS:;        forecast over the holdout data.
```

ROBUST REGRESSION PROGRAM - SECOND HALF. THIS HALF OF THE PROGRAM READS IN THE FITTING FORECASTS, CALCULATES THE VARIOUS ERROR MEASURES FOR EACH LENGTH ROBUST REGRESSION FOR EACH SERIES, AND FILES THE RESULTS.

```
DATA _NULL_;
INFILE ROBUST1;                              File with the
INPUT SID NFIT OBSNUM XI FI1-FI4;               forecasts.
ARRAY FI{4} FI1-FI4;
ARRAY SUMS{4,5} S1-S20;
LAGXI = LAG(XI);
DO I = 1 TO 4;
  N = I+2;
  F = LAG(FI{I});                       Finding the fitting
  E = XI-F;                                         error.
  IF OBSNUM LT N+1 THEN GO TO CUT;
  V8 = ABS(E);
  V9 = E**2;                               Manipulating the
  V11 = ABS(E/XI);                                  error.
  V17 = (E/LAGXI)**2;
  V18 = ((XI-LAGXI)/LAGXI)**2;
  SUMS{I,1} + V8;
  SUMS{I,2} + V9;                          Summing over the
  SUMS{I,3} + V11;                                 series.
  SUMS{I,4} + V17;
  SUMS{I,5} + V18;
  IF OBSNUM = NFIT THEN DO;
    DEN = NFIT - N;                         Calculating the
    MAD = SUMS{I,1}/DEN;                    error measures.
    MSE = SUMS{I,2}/DEN;
    MAPE = (SUMS{I,3}/DEN)*100;
    TH3 = SQRT(SUMS{I,4}/SUMS{I,5});
FILE RRRES;                                      Filing the
PUT SID N MAD MSE MAPE TH3;                        results.
  END;
CUT:END;
IF OBSNUM = NFIT THEN DO;                     Resetting for
  DO M = 1 TO 4;                           the next series.
    DO N = 1 TO 5;
      SUMS{M,N} = 0;
    END;
  END;
END;
```

PROGRAM TO FIND THE BEST PERIOD LENGTH FOR THE ROBUST REGRESSION TECHNIQUE. THIS PROGRAM SEARCHES THROUGH THE FILE CREATED BY THE PREVIOUS PROGRAM AND IDENTIFIES THE BEST LENGTHS IN TERMS OF FITTED MAD, MSE, MAPE, AND THEIL'S U.

```
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA ONE;
INFILE RRRES;
INPUT SID N MAD MSE MAPE TH3;
RETAIN;
IF N = 3 THEN DO;
  MADMIN = MAD;  MADN = N;
  MSEMIN = MSE;  MSEN = N;
  MAPEMIN = MAPE;  MAPEN = N;
  TH3MIN = TH3;  TH3N = N;
  END;
ELSE DO;
  IF MAD LT MADMIN THEN DO;
    MADMIN = MAD; MADN = N;
  END;
  IF MSE LT MSEMIN THEN DO;
    MSEMIN = MSE;  MSEN = N;
  END;
  IF MAPE LT MAPEMIN THEN DO;
    MAPEMIN = MAPE; MAPEN = N;
  END;
  IF TH3 LT TH3MIN THEN DO;
    TH3MIN = TH3; TH3N = N;
  END;
  IF N = 6 THEN DO;
    FILE RRBEST1;
    PUT SID MADMIN MSEMIN MAPEMIN TH3MIN MADN
    MSEN MAPEN TH3N;
  END;
END;
```

THIS PROGRAM TAKES THE BEST FITTING RESULTS FROM THE PREVIOUS PROGRAM, USES THIS INFORMATION TO PULL THE FORECASTING SLOPE AND "INTERCEPT" FROM THE PREVIOUSLY CREATED FILE RRAB.DAT, AND REFILE THE COMPLETE ROBUST REGRESSION FITTING RESULTS.

```
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA _NULL_;
INFILE RRBEST1;
INPUT SID MADRFIT MSERFIT MAPERFIT TH3RFIT MADRN
  MSERN MAPERN TH3RN;
INFILE RRAB;
INPUT SID NFIT A3-A6 B3-B6;
ARRAY A{4} A3-A6;
ARRAY B{4} B3-B6;
MADRA = A{MADRN-2};
MADRB = B{MADRN-2};
MSERA = A{MSERN-2};
MSERB = B{MSERN-2};
MAPERA = A{MAPERN-2};
MAPERB = B{MAPERN-2};
TH3RA = A{TH3RN-2};
TH3RB = B{TH3RN-2};
FILE RRBEST1A;
PUT SID MADRFIT MSERFIT MAPERFIT TH3RFIT MADRN
  MSERN MAPERN TH3RN MADRA MADRB MSERA MSERB
  MAPERA MAPERB TH3RA TH3RB;
```

The variable MADRN is the best length robust regression over the fitting data, based upon MAD.

PROGRAM TO CALCULATE THE FORECAST ERRORS FOR HORIZONS ONE THROUGH SIX USING THE BEST MODEL FOR EACH TECHNIQUE AND EACH SERIES, AND TO CALCULATE THE MAPE AVERAGE VALUES OVER THE NINETY TWO MONOTONIC TREND SERIES. THIS PROGRAM OPERATES ON THE SAS DATA SET TBYSID11.SSD WHERE ALL BEST FIT INFORMATION AND FORECASTING EQUATIONS WERE STORED.

```
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA ONE;
SET TREND.TBYSID11;
ARRAY DEVS{5,6} DEVM1-DEVM6 DEVE1-DEVE6
      DEVR1-DEVR6 DEVS1-DEVS6 DEVW1-DEVW6;
ARRAY PAD{5,6} PAD1-PAD30;
ARRAY MAPEFIT{5} MAPEMFIT MAPEEFIT MAPERFIT
      MAPESFIT MAPEWFIT;
ARRAY XVALS{7} XNFIT XNOF1-XNOF6;
ARRAY SUMAPDEV{5,6} SUM1-SUM30;
ARRAY SMAPEFIT{5} SMAPE1-SMAPE5;
ARRAY AMAPFO{5,6} AMAPFO1-AMAPFO30;
ARRAY AMAPEFT{5} AMAPEFT1-AMAPEFT5;
DO J = 1 TO 6;
  DEVS{1,J} = XVALS{J+1} - (MAPEMA + J*MAPEMB);
  DEVS{2,J} = XVALS{J+1} - (MAPEEA + J*MAPEEB);
  DEVS{3,J} = XVALS{J+1} - (MAPERA + J*MAPERB);
  DEVS{4,J} = XVALS{J+1} - (MAPESA + ((J+1)/2)*
              MAPESB);
  DEVS{5,J} = XVALS{J+1} - (MAPEWA + ((J+1)/2) *
              MAPEWB);
END;
DO J = 1 TO 5;
  SMAPEFIT{J} + MAPEFIT{J};
  DO K = 1 TO 6;
    PAD{J,K} = ABS(DEVS{J,K}/XVALS{K+1});
    SUMAPDEV{J,K} + PAD{J,K};
  END;              The series with SID=976 is the
END;            last series in the trend data set.
IF SID LT 976 THEN GO TO NEXTOBS;
DO J = 1 TO 5;
  AMAPEFT{J} = SMAPEFIT{J}/92;
  DO K = 1 TO 6;
    AMAPFO{J,K} = SUMAPDEV{J,K}/92;
  END;
END; -fitting- forecasting -   MAPE results for:
PUT AMAPEFT1 AMAPFO1-AMAPFO6;      D. Moving Avg
PUT AMAPEFT2 AMAPFO7-AMAPFO12;     L. Exp. Smooth
PUT AMAPEFT3 AMAPFO13-AMAPFO18;    D. Running Med
PUT AMAPEFT4 AMAPFO19-AMAPFO24;    D. Smoothed M.
PUT AMAPEFT5 AMAPFO25-AMAPFO30;    Robust Regress
NEXTOBS:;
```

PROGRAM TO CALCULATE THE THEIL'S U VALUES OVER THE NINETY ONE SERIES. THE SERIES WITH SID=967 HAD ONE HORIZON VALUE WHERE THEIL'S U WAS NOT DEFINED DUE TO NO CHANGE IN THE SERIES BETWEEN TWO POINTS IN TIME, SO THE SERIES IS DROPPED.

```
LIBNAME TREND '[RKANKEY.SIX2T]';
DATA ONE;
SET TREND.TBYSID11;
IF SID = 967 THEN GO TO NEXTOBS;
ARRAY DEVS{5,6} DEVM1-DEVM6 DEVE1-DEVE6
      DEVR1-DEVR6 DEVS1-DEVS6
      DEVW1-DEVW6;
ARRAY TH{5,6} TH1-TH30;
ARRAY TH3FIT{5} TH3MFIT TH3EFIT TH3RFIT TH3SFIT
      TH3WFIT;
ARRAY XVALS{7} XNFIT XNOF1-XNOF6;
ARRAY SUMTH3{5,6} SUM1-SUM30;
ARRAY STH3FIT{5} STH31-STH35;
ARRAY ATH3F{5,6} ATH3F1-ATH3F30;
ARRAY ATH3FT{5} ATH3FT1-ATH3FT5;
ARRAY THNUM{5,6} THN1-THN30;
ARRAY THDEN{5,6} THD1-THD30;
ARRAY STHNUM{5} STHN1-STHN5;
ARRAY STHDEN{5} STHD1-STHD5;
ARRAY THALL{5} THALL1-THALL5;
DO J = 1 TO 6;
  DEVS{1,J} = XVALS{J+1} - (TH3MA+J*TH3MB);
  DEVS{2,J} = XVALS{J+1} - (TH3EA+J*TH3EB);
  DEVS{3,J} = XVALS{J+1} - (TH3RA+J*TH3RB);
  DEVS{4,J} = XVALS{J+1} - (TH3SA+((J+1)/2)*
              TH3SB);
  DEVS{5,J} = XVALS{J+1} - (TH3WA+((J+1)/2)*
              TH3WB);
END;
DO J = 1 TO 5;
STH3FIT{J} + TH3FIT{J};
DO K = 1 TO 6;
  THNUM{J,K} = ((DEVS{J,K}/XVALS{1})**2);
  THDEN{J,K} = ((XVALS{K+1} - XVALS{1})
                /XVALS{1})**2;
  STHNUM{J} + THNUM{J,K};
  STHDEN{J} + THDEN{J,K};
  TH{J,K}= SQRT(((DEVS{J,K}/XVALS{1})**2) /
           ((XVALS{K+1} - XVALS{1}) /
            XVALS{1})**2);
  SUMTH3{J,K} + TH{J,K};
END;
  THALL{J} = SQRT(STHNUM{J}/STHDEN{J});
END;


IF SID LT 976 THEN GO TO NEXTOBS;
```

```
DO J = 1 TO 5;
ATH3FT{J} = STH3FIT{J}/91;
DO K = 1 TO 6;
  ATH3F{J,K} = SUMTH3{J,K}/91;
END;
END; Fitting - forecasting   Theil's U results
PUT ATH3FT1 ATH3F1-ATH3F6;    D. Moving Avg.
PUT ATH3FT2 ATH3F7-ATH3F12;   D. Exponential Sm
PUT ATH3FT3 ATH3F13-ATH3F18;  D. Running Medians
PUT ATH3FT4 ATH3F19-ATH3F24;  D. Smoothed Meds.
PUT ATH3FT5 ATH3F25-ATH3F30;  Robust Regression
NEXTOBS:;
IF SID EQ 976 THEN DO;
PUT THALL1-THALL5;           Theil's U by technique
END;                         over all six horizons.
```

BIBLIOGRAPHY

Armstrong, J. S. (1978), Long Range Forecasting: From Crystal Ball to Computer, J. Wiley and Sons.

Armstrong, J. S. (1983), "Relative Accuracy of Judgemental and Extrapolative Methods in Forecasting Annual Earnings", Journal of Forecasting, Vol 2, pp 437-447.

Armstrong, J. S. (1985), Long Range Forecasting: From Crystal Ball to Computer, 2nd Ed, J. Wiley and Sons.

Ballou, R. H. (1985), Business Logistics Management: Planning and Control, Prentice-Hall.

Box, G. E. P. and G. M. Jenkins (1976), Time Series Analysis, Forecasting and Control, Holden-Day.

Brown, R. G. (1963), Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice-Hall.

Bustos, O. H. and V. J. Yohai (1986), "Robust Estimates for ARMA Models", Journal of the American Statistical Association, Vol 81, pp155-168.

Carbone, R. and J. S. Armstrong (1982), "Evaluation of Extrapolative Forecasting Methods: Results of a Survey of Academicians and Practitioners", Journal of Forecasting, Vol 1, pp 215-217.

Carbone, R. and W. L. Gorr (1985), "Accuracy of Judgemental Forecasting of Time Series", Decision Sciences, Vol 16, pp 153-160.

Carbone, R. and S. Makridakis (1986), "Forecasting When Pattern Changes Occur Beyond the Historical Data", Management Science, Vol 32, pp 257-271.

Chambers, J. C., S. K. Mullick, and D. D. Smith (1971), "How to Choose the Right Forecasting Technique", Harvard Business Review, Jul-Aug, pp 45-74.

_____, _____, and _____ (1974), An Executives Guide to Forecasting, J. Wiley and Sons.

Chatfield, C. (1979), "Discussion of the Paper by Professor Makridakis and Dr. Hibon", Journal of the Royal Statistical Society, Series A, Vol 142, pp 129-130.

Chatfield, C. (1980), The Analysis of Time Series, An Introduction, Chapman and Hall.

Conover, W. J. (1980), Practical Nonparametric Statistics, 2nd Ed, J. Wiley and Sons.

Dalrymple, D. J. and B. E. King (1981), "Selecting Parameters for Short-Term Forecasting Techniques", Decision Sciences, Vol 12, pp 661-669.

Draper, N. R. and H. Smith (1981), Applied Regression Analysis, 2nd Ed, J. Wiley and Sons.

Durbin, J. (1979), "Discussion of the Paper by Professor Makridakis and Dr. Hibon", J. of the Royal Statistical Society, Series A, Vol 142, pp 133-134.

Gardner, E. S. (1983), "The Trade-offs in Choosing a Time Series Method", J. of Forecasting, Vol 2, pp 263-267.

Gardner, E. S. and D. G. Dannenbring (1980), "Forecasting With Exponential Smoothing: Some Guidelines for Model Selection", Decision Sciences, Vol 11, pp 370-383.

Gardner, E. S. and E. McKenzie (1985), "Forecasting Trends in Time Series", Management Science, Vol 31, pp 1237-1246.

Gorr, W. L. and C. Hsu (1985), "An Adaptive Filtering Procedure for Estimating Regression Quantiles", Management Science, Vol 31, pp 1019-1029.

Granger, C. W. J. and R. Ramanthan (1984), "Improved Methods of Combining Forecasts", Journal of Forecasting, Vol 3, pp 197-204.

Graver, C. A. (1981), "Selecting Estimating Techniques Using Historical Simulation", Concepts, Vol 4 (Spring), pp 145-163.

Guerts, M. D. (1983), "Evaluating a Forecasting Competition with Emphasis on the Combination of Forecasts", comment on the M-competition, Journal of Forecasting, Vol 2, No. 3, pp 267-269.

Hardle, W. and T. Gasser (1984), "Robust Non-Parametric Function Fitting", J. Royal Statistical Society, Series B, Vol 46, pp 42-51.

Hibon, M. (1984), "Naive, Moving Average, Exponential Smoothing and Regression Methods", Chapter 8 of The Forecasting Accuracy of Major Time Series Methods, Makridakis, et. al., J. Wiley and Sons.

Hogarth, R. M. and S. Makridakis (1981), "Forecasting and Planning: An Evaluation", Management Science, Vol 27, pp 115-138.

Hogg, R. V. (1979), "Statistical Robustness: One View of Its Use in Applications Today", The American Statistician, Vol 33, No. 3, pp 108-115.

Hollander, M. and D. A. Wolfe (1973), Nonparametric Statistical Methods, J. Wiley and Sons.

Huss, W. R. (1985), "Selecting the Best Load Forecasting Techniques for Electric Utilities", Doctoral Dissertation, The Ohio State University (College of Industrial and Systems Engineering).

Hussain, S. S. and P. Sprent (1983), "Non-parametric Regression", J. Royal Statistical Society, Series A, Vol 146, pp 182-191.

Johnson, N. L. and S. Kotz (1970), Continuous Univariate Distributions - 1, Houghton Mifflin.

Kankey, R. D. and P. A. Thompson (1986), "Loss Functions and Forecast Accuracy: Some Relationships Between Measures Used to Compare Forecasting Techniques", College of Administrative Science Working Paper Series, The Ohio State University, WPS-86-91.

Kucukemiroglu, O. and K. Ord (1985), "The Impact of Extreme Observations on Simple Forecasting Methods", Decision Sciences, Vol 16, pp 299-308.

LaLonde, B. J. (1984), Lecture Notes from BMKTG 880, Physical Distribution Management, The Ohio State University, Spring Quarter.

Lawrence, M. J. (1983), "An Exploration of Some Practical Issues in the Use of Quantitative Forecasting Models", J. of Forecasting, Vol 2, pp 169-179.

Lawrence, M. J.; R. H. Edmundson; and M. J. O'Connor (1986), "The Accuracy of Combining Judgemental and Statistical Forecasts", Management Science, Vol 32, No 12, pp 1521-1532.

Levenbach, H. and J. P. Cleary (1984), *The Modern Forecaster*, Lifetime Learning Publications.

Mahmoud, E. (1984), "Accuracy in Forecasting: A Survey", *J. of Forecasting*, Vol 3, pp 139-159.

Makridakis, S. (1983), "Empirical Evidence versus Personal Experience", *J. of Forecasting*, Vol 2, pp 295-306.

Makridakis, S. (1984), "Forecasting: State of the Art", Chapter 1 of *The Forecasting Accuracy of Major Time Series Methods*, Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler, J. Wiley and Sons.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler (1982), "The Accuracy of Extrapolative (Time Series) Methods; Results of a Forecasting Competition", *J. of Forecasting*, Vol 1, pp 111-153.

___, ___, ___, ___, ___, ___, ___, ___ and ___ (1984), *The Forecasting Accuracy of Major Time Series Methods*, J. Wiley and Sons.

Makridakis, S. and M. Hibon (1979), "Accuracy of Forecasting: An Empirical Investigation", *J. of the Royal Statistical Society, Series A*, Vol 142, pp 97-145.

Makridakis, S. and S. C. Wheelwright (1978), *Interactive Forecasting*, 2nd Ed, Holden-Day.

Makridakis, S., S. C. Wheelwright, and V. E. McGee (1983), *Forecasting: Methods and Applications*, J. Wiley and Sons.

Makridakis, S., and R. Winkler (1983), "Averages of Forecasts: Some Empirical Results", *Management Science*, Vol 29, pp 987-996.

McLaughlin, R. L. (1972), "Measuring the Accuracy of Forecasts", *Business Economics*, Vol VII, No. 3, pp 27-35.

McLaughlin, R. L. (1975), "The Real Record of the Economic Forecasters", *Business Economics*, Vol X, No. 3, pp 28-36.

Mentzer, J. T. and J. E. Cox (1984), "Familiarity, Application, and Performance of Sales Forecasting Techniques", *J. of Forecasting*, Vol 3, pp 27-36.

Montgomery, D. C. and L. A. Johnson, *Forecasting and Time Series Analysis*, McGraw-Hill, 1976.

Newbold, P. (1983), "The Competition to End all Competitions", comment on the M-competition, Journal of Forecasting, Vol 2, No. 3, pp 276-279.

Newbold, P. and C. W. T. Granger (1974), "Experience With Forecasting Univariate Time Series and the Combination of Forecasts", J. of the Royal Statistical Society, Series A, Vol 137, pp 131-165.

Noether, G. E. (1984), "Nonparametrics: The Early Years - Impressions and Recollections", The American Statistician, Vol 38, No. 3, pp 173-178.

Pack, D. J. (1983), "What Do These Numbers Tell Us", Comment on the M-Competition, J. of Forecasting, Vol 2, No 3, pp 279-285.

Parzen, E. (1960), Modern Probability Theory and Its Applications, J. Wiley and Sons.

Sanders, N. R. and L. P. Ritzman (1987), "Some Empirical Findings on Short-Term Forecasting: Technique Complexity and Combinations", College of Business Working Paper Series, The Ohio State University, WPS 87-91.

SAS Institute Inc. (1982), SAS User's Guide: Basics, 1982 Edition. Cary, NC.

SAS Institute Inc. (1985), SAS User's Guide: Basics, Version 5 Edition. Cary, NC.

SAS Institute Inc. (1985), SAS/GRAPH User's Guide: Version 5 Edition. Cary, NC.

Stigler, S. M. (1974), "Comments on Newbold and Granger", J. of the Royal Statistical Society, Series A, Vol 137, pp 157-158.

Theil, H. (1966), "Measuring the Accuracy of Point Predictions", Chapter 2 in Applied Economic Forecasting, North-Holland.

Torfin, G. P. and T. R. Hoffman (1968), "Simulation Tests of Some Forecasting Techniques", Production and Inventory Management, 2nd Qtr, pp 71-78.

Tukey, J. W. (1977), Exploratory Data Analysis, Addison-Wesley.

Weiss, A. A. and A. P. Andersen (1984), "Estimating Time Series Models Using the Relevant Forecast Evaluation Criterion", J. Royal Statistical Society, Series A, Vol 147, pp 484-487.

Wheelwright, S. C. and D. G. Clarke (1976), "Corporate Forecasting: Promise and Reality", Harvard Business Review, Nov-Dec, p 40.

Wheelwright, S. C. and S. Makridakis (1985), Forecasting Methods for Management, 2nd Ed, J. Wiley and Sons.

Winkler, R. L. (1983), "The Effects of Combining Forecasts and the Improvement of the Overall Forecasting Process", J. of Forecasting, Vol 2, pp 293-294.

Winkler, R. L. and S. Makridakis (1983), "The Combination of Forecasts", J. of the Royal Statistical Society, Series A, Vol 146, pp 150-157.

Winters, P. R. (1960), "Forecasting Sales by Exponentially Weighted Moving Averages", Management Science, Vol 6, pp 324-342.

Wright, D. J. (1986), "Forecasting Data Published at Irregular Time Intervals Using an Extension of Holt's Method", Management Science, Vol 32, pp 499-510.