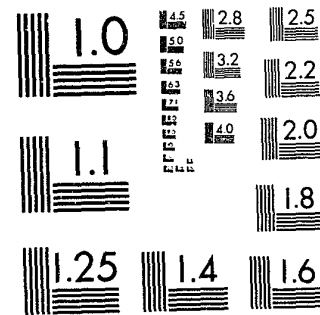


National Criminal Justice Reference Service

ncjrs

This microfiche was produced from documents received for inclusion in the NCJRS data base. Since NCJRS cannot exercise control over the physical condition of the documents submitted, the individual frame quality will vary. The resolution chart on this frame may be used to evaluate the document quality.



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Microfilming procedures used to create this fiche comply with the standards set forth in 41CFR 101-11.504.

Points of view or opinions stated in this document are those of the author(s) and do not represent the official position or policies of the U. S. Department of Justice.

National Institute of Justice
United States Department of Justice
Washington, D. C. 20531

3/30/84

HOW TO HANDLE SEASONALITY

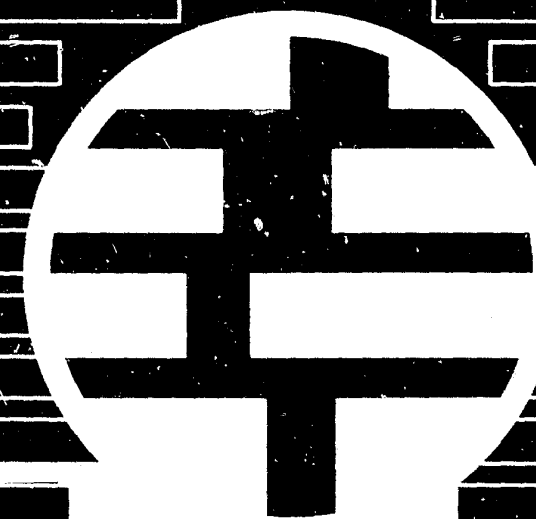
Introduction to the Detection
and Analysis of Seasonal Fluctuation
in Criminal Justice Time Series

May 1983

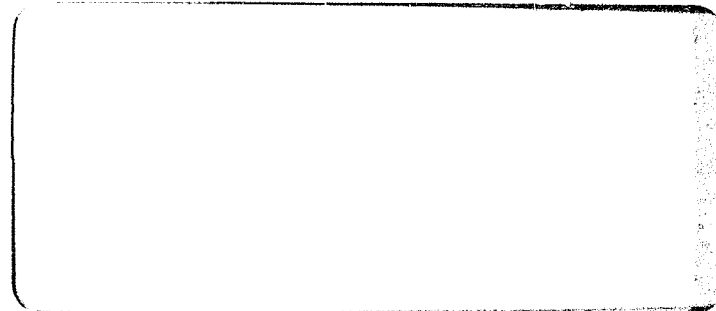


**ILLINOIS
CRIMINAL JUSTICE
INFORMATION AUTHORITY**

92928



92928



HOW TO HANDLE SEASONALITY

Introduction to the Detection
and Analysis of Seasonal Fluctuation
in Criminal Justice Time Series

May 1983

by Carolyn Rebecca Block
Statistical Analysis Center

ILLINOIS CRIMINAL JUSTICE INFORMATION AUTHORITY
William Gould, Chairman
J. David Coldren, Executive Director

U.S. Department of Justice
National Institute of Justice

92928

This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the National Institute of Justice.

Permission to reproduce this copyrighted material has been granted by
Illinois Criminal Justice
Information Authority
to the National Criminal Justice Reference Service (NCJRS).

Further reproduction outside of the NCJRS system requires permission of the copyright owner.

PROLOGUE

When most consumers of seasonally adjusted series -- and that includes nearly every economically literate person -- are confronted by the question of why they prefer such a series to the original, the most common and natural reaction is that the answer is obvious. Yet on further reflection the basis for such a preference becomes less clear, and those who give the matter extensive thought often finish by becoming hopelessly confused.

-- Grether and Nerlove (1970:685)

TABLE OF CONTENTS

PROLOGUE.....	i
TABLE OF CONTENTS.....	iii
EXECUTIVE SUMMARY.....	v
ACKNOWLEDGEMENTS.....	vii
INTRODUCTION.....	1
WHY DOES SEASONALITY MATTER?.....	3
WHAT IS SEASONALITY?.....	9
The Component Definition of Seasonality.....	9
The Stochastic Definition of Seasonality.....	14
Summary: Two Definitional Approaches.....	16
TOOLS FOR DETECTING AND ANALYZING SEASONALITY.....	19
Component Methods.....	19
Moving Average.....	20
Additive/Multiplicative Assumption.....	21
F of Stable Seasonality, and	
Relative Contribution of the Irregular.....	22
Average Duration of Run.....	24
Months for Cyclical Dominance.....	25
Pattern Consistency.....	27
Trading Day Option.....	28
Appropriate Applications.....	29
Extremes.....	30
Series Length.....	30
Discontinuities.....	30
Moving Seasonality.....	31
Stochastic Methods.....	33
Moving Average and Autoregressive Processes.....	33
Identifying the Process of a Series.....	34
Stationarity.....	36
Model Evaluation.....	43
Correlogram of Residuals.....	45
Cumulative Periodogram of Residuals.....	45
Appropriate Applications.....	48
Discontinuities.....	48
Extremes.....	50
Moving Seasonality.....	50
Summary.....	50
ANNOTATED BIBLIOGRAPHY.....	51

NCJRS

FEB 13 1984

ACQUISITIONS

EXECUTIVE SUMMARY

This report is an introduction to the fundamentals of seasonal analysis, with an emphasis on practical applications to criminal justice. Administrators, policy makers, researchers, and others who make decisions based on crime data now have time series data available that allow them to answer questions that could not be answered only a few years ago. But to answer these questions, it is necessary to use methods appropriate to the analysis of time series, including methods of detecting and analyzing seasonality. Many fields outside of criminology have long had a wealth of time series data available to them, and have developed methods to analyze seasonality in those data. This report guides the reader to the use of the most common of these methods.

In the analysis of time series data, as in the analysis of cross-sectional data, description must precede explanation. We must describe the past before we can forecast the future. We must become familiar with patterns of change over time in the original data before we can develop complex causal models. If we do not, we risk misspecifying the model, and forecasts and policy decisions based on that model may be erroneous.

An elementary part of the description of patterns over time in monthly or quarterly data is the description of seasonal fluctuation. Some monthly and quarterly series fluctuate with the seasons of the year; others do not. If we assume that a series is seasonal, when it is not, or that a series is not seasonal, when it is, we risk erroneous forecasts and explanatory models.

This report discusses the two major approaches to defining and detecting seasonality -- the component approach and the stochastic approach. Although the two approaches are mathematically similar, there are practical differences in emphasis. The component approach emphasizes a separate description of seasonal fluctuation, while the stochastic approach emphasizes forecasting the future with a model that incorporates seasonality. The component approach focuses on seasonality itself, while the stochastic approach focuses on seasonality as it affects the accuracy of a forecast.

No single method of analysis is appropriate in every situation. The method of choice depends upon the objectives of the analysis. For example, a decision to build a new prison will depend upon a forecast of the total number of inmates, with seasonal fluctuation included in the total. On the other hand, if there are wide seasonal fluctuations in the number of inmates, it might be necessary to open an additional wing during some months of the year. The decision to do this would depend on an analysis of the seasonal component.

Neither the component approach nor the stochastic approach offers a simple, objective, yes-or-no criterion for detecting the presence of seasonality in a time series. Both approaches depend heavily on the judgment of the analyst, although each approach gives the analyst a number of statistical tools upon which to base that judgment. This report discusses and compares these tools, and gives the analyst some basic rules of thumb for using them in various practical situations.

In addition, for those who need more detail than this report provides, it includes an annotated bibliography of 110 references to literature about seasonal analysis and to reports analyzing the seasonality of crime.

ACKNOWLEDGEMENTS

The writing of this report was supported, in part, by a grant from the Bureau of Justice Statistics, U.S. Department of Justice.

The staff of the Bureau of Labor Statistics, especially Robert J. McIntire and Kathryn Beale, helped greatly in our initial use of the Census X-11 program. Not only were their suggestions and practical advice useful, but their enthusiasm for this relatively new application of X-11 techniques was contagious. Similarly, Harry V. and June Roberts were of great help in our initial use of the IDA package for stochastic time series modeling. They were never too busy to offer hints and practical suggestions for using the package and for interpreting its results.

This report has been a "working draft" for several years. During that time, many people have commented on its various drafts. The staff of the Statistical Analysis Center, including James Coldren, Paul Fields and Jim Lucas, made helpful suggestions. Larry Dykstra and Keith Coopride used the manual to do seasonal analysis, and commented on its use in practical situations. In addition, the comments of Richard L. Block and Richard E. Barrett were especially helpful.

Louise S. Miller played a unique role in the writing of "How to Handle Seasonality." As an apprentice in the Statistical Analysis Center, she was one of the first persons to use the manual as a beginning text on seasonality, and made many comments and suggestions based on her experience. In addition, she has written a users' guide to the version of the X-11 available at the Authority, and has handled requests from users for seasonal analysis. She produced many of the graphs that the report uses as examples.

INTRODUCTION

Administrators, policy makers and researchers now have time series data available that allow them to answer questions that could not be answered only a few years ago. But to answer these questions, it is necessary to use methods appropriate to the analysis of time series, including methods of detecting and analyzing seasonality. Many fields outside of criminology have long had a wealth of time series data available to them, and have developed methods to analyze seasonality in those data. This report is an introduction to the most commonly used of these methods, with practical crime data examples.¹

The question of seasonality is a paradox. On one hand, the concept seems simple. Criminologists have traditionally believed (see Wolfgang, 1966:96-106 for a review) that more crimes occur during some months of the year than others. On the other hand, this simplicity is deceptive: a precise definition of seasonality is elusive, and the detection and measurement of seasonality are subjective.

The quote by Grether and Nerlove in the prologue exactly describes the Statistical Analysis Center staff's experience when we first confronted the question of seasonality. We naively thought that it would be a simple problem, that all we had to do would be to discover the standard "cookbook" seasonal adjustment method and apply it. However, we soon found that there is no standard cookbook approach to seasonality. Our routine search for a standard program soon became a lengthy investigation of the philosophical approaches and related mathematical methods for the detection, measurement and adjustment of seasonal fluctuation.

This report is a summary of the results of that investigation. It reviews the two most common approaches to detecting and measuring seasonality. It also discusses the qualitative and

¹A complete review of all seasonal analysis methods would fill at least one book. This report is limited to the two most commonly used methods, the seasonal component method and the stochastic modelling method. Readers who want to investigate alternative methods should see Kendall (1976), Zellner (1978), or Pierce (1980) for an overview; Lovell (1963) or Dutta (1975) for dummy regression; Shiskin (1957) for same-month-last-year; Land (1978, 1980) and Land and Felson (1976) for econometric and time-inhomogenous methods; Bliss (1958) or Warren, et al. (1981) for periodic regression analysis (PRA); Cleveland, et al. (1979) or Velleman and Hoaglin (1981) for resistant methods, and Rosenblatt (1965) for spectral analysis. For a technical guide to using the seasonality and other time series computer programs that are available at SAC, see the SAC report, "Technical Manual for Time Series Pattern Description," by Louise S. Miller.

quantitative choices that a user of any seasonal analysis method must make. As a simple introduction to seasonality, it includes statistics only when necessary, but it also includes a long, annotated bibliography of technical reports, for those who need more detail. In short, it is the report that I wish had existed when I first began to analyze the seasonality of time series.

WHY DOES SEASONALITY MATTER?

Time series containing time periods shorter than a year, such as monthly or quarterly series, may vary according to the season of the year. That is, a phenomenon may occur more frequently at certain times of the year, and less frequently at other times. On the other hand, not every monthly or quarterly time series is seasonal. For example, the number of aggravated assault offenses known to the police in Illinois (figure 1) is seasonal, but the number of homicide offenses known to the police in Illinois (figure 2) is not seasonal.²

If we ignore the question of seasonality, we may make the error of assuming that a series is not seasonal, when in fact it is. On the other hand, if we automatically adjust for seasonality without first analyzing the series to see if it is seasonal or not, we may make the error of adjusting for nonexistent seasonality. What difference would either sort of error make to common administrative or policy decisions?

If we make the first error, ignore the question of seasonality in a series that is seasonal, we ignore two kinds of information that may be useful in making decisions: a description of seasonal fluctuation, and a description of the variation in the series with the seasonal fluctuation removed. Such descriptions provide a necessary foundation for explanatory models, forecasts, and tests of intervention hypotheses. Without a prior description, models may be misspecified, forecasts inaccurate, and hypothesis tests erroneous.

Policy makers and administrators often need to know the amount of seasonal fluctuation in order to allocate resources. For example, if more rapes occur in the summer, a police chief may want to allocate more resources to a rape crisis center or to a rape investigation unit in the summer months. If more people are sentenced to prison in the fall, a prison administrator may want to arrange for more beds in the fall months. Knowledge of the pattern of seasonal fluctuation around the overall trend helps the administrator estimate the resources needed from month to month.

Ignoring seasonality may also lead to erroneous conclusions in comparing one month and another. Suppose that a crime prevention program were instituted in May, and that one of the goals of this program was to reduce larceny. If more larceny incidents

²The lines superimposed on the raw data in figures 1 and 2, and other figures in this report, are "line segment fits," which use linear spline regression to describe the general pattern of change over time in a variable. For more information, see the Statistical Analysis Center report, "Manual for the Pattern Description of Time Series."

Figure 1

ILLINOIS INDEX AGGRAVATED ASSAULT, 1972 TO 1981

RAW DATA SERIES = □
 MULTI-SEGMENT LINE = ○
 SOURCE: SAC EDITION ILLINOIS UNIFORM CRIME REPORTS
 OFFENSE DATA, 1981 PRELIMINARY
 INDEX AGG. ASSAULT = AGG. ASSAULT, AGG. BATTERY, ATT. MURDER

FIRST SLOPE = -5.11
 Y ZERO INTERCEPT = 2394.48
 SECOND SLOPE = -18.20
 Y ZERO INTERCEPT = 336.35
 THIRD SLOPE = -29.70
 Y ZERO INTERCEPT = 642.11

FIRST TURNING POINT 8:79.50
 SECOND TURNING POINT 8:105.10
 TOTAL SSR = 1791090.09

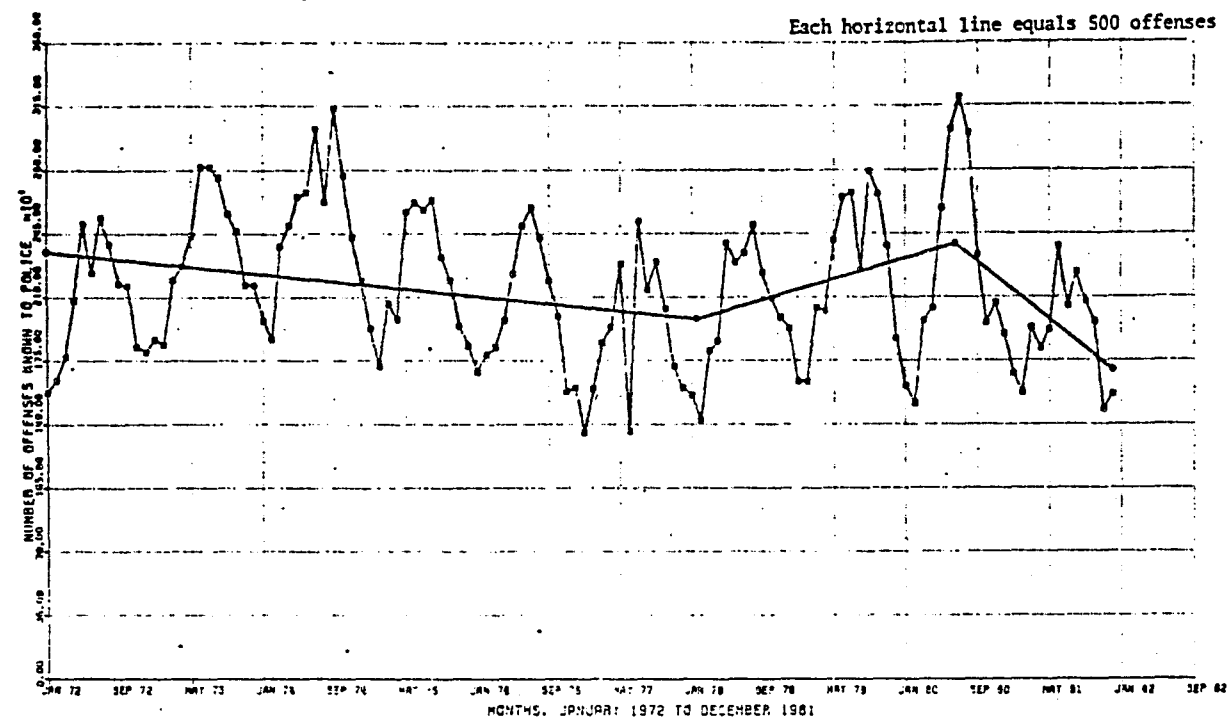


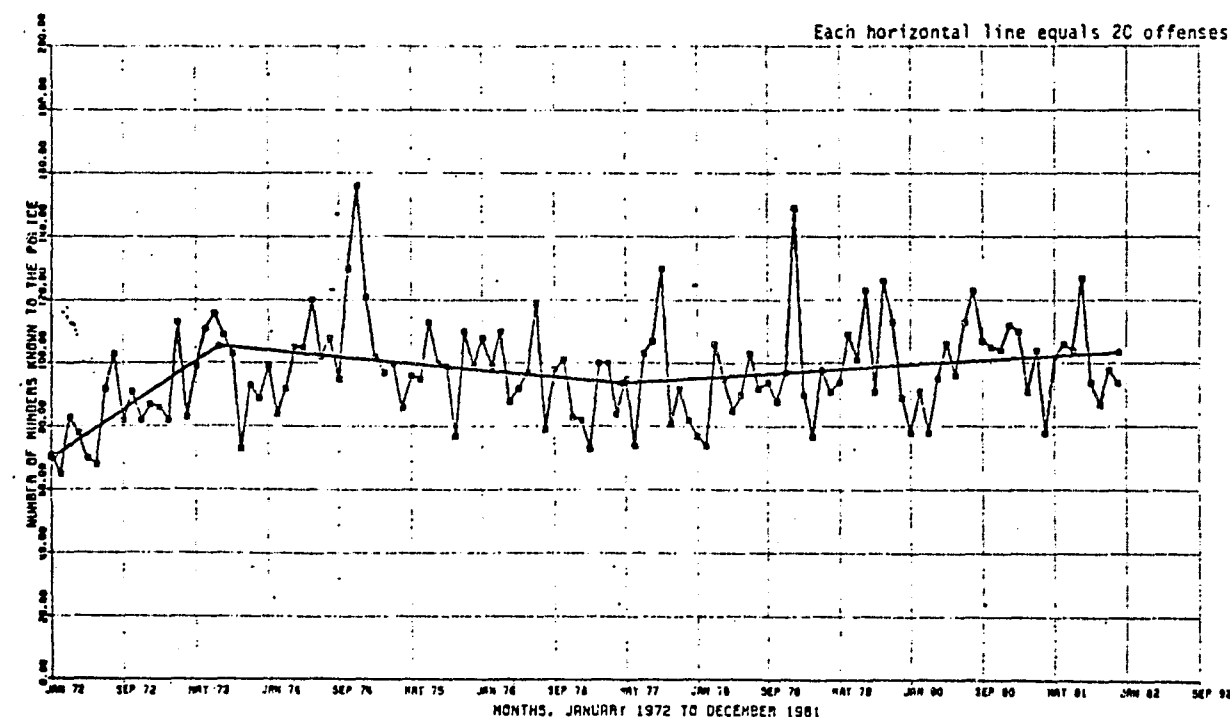
Figure 2

ILLINOIS INDEX MURDER, 1972 TO 1981

RAW DATA SERIES = □
 MULTI-SEGMENT LINE = ○
 SOURCE: SAC EDITION ILLINOIS UNIFORM CRIME REPORTS
 OFFENSE DATA, 1981 PRELIMINARY
 INDEX MURDER = MURDER AND VOLUNTARY MANSLAUGHTER

FIRST SLOPE = 1.91
 Y ZERO INTERCEPT = 82.19
 SECOND SLOPE = -0.26
 Y ZERO INTERCEPT = 118.37
 THIRD SLOPE = 0.18
 Y ZERO INTERCEPT = 82.17

FIRST TURNING POINT 8:18.50
 SECOND TURNING POINT 8:146.50
 TOTAL SSR = 23015.52



ordinarily occur in the summer than in the spring, the effect of the program might be obscured by seasonal variation. The number of larcenies occurring in June might be as high or even higher than the number of larcenies occurring in April, even if the program actually decreased larceny. In such a situation, the policy maker or administrator is not primarily interested in seasonal fluctuation, but is interested in the overall trend, with seasonal fluctuation removed. Once seasonality has been taken into account, were there fewer larcenies after the crime prevention program?

These two kinds of descriptions -- description of the pattern of seasonal fluctuation, and description of the pattern of the variable with the seasonal fluctuation removed -- can help in communication to policy makers, and other users of crime data (see Granger 1978:38-39). Seasonal fluctuation may be so great that it obscures any other pattern over time. Suppose that a reporter or a member of the City Council asks the Police Department's crime analysis unit whether larceny offenses are increasing or decreasing. The unit's answer will be more easily understood if it is accompanied by a graph of the seasonally adjusted data (figure 4), than if it is accompanied by a graph of the original data (figure 3). There is much less variation in the seasonally adjusted larceny series than in the original larceny series. The variation due to a known cause, seasonality, has been removed. With this seasonal fluctuation removed, the general pattern of larcenies over time appears much more clearly.³

The second kind of error, to assume that a series is seasonal when, in fact, it is not, may also lead to an inaccurate description of the pattern of the series. Failure to recognize a lack of seasonality may lead to model misspecification and inaccurate forecasts in the same way as failure to recognize the presence of seasonality (see Fromm, 1978:26). We will make the same descriptive mistakes discussed above, but for the opposite sort of error. An erroneous assumption that all Mays are higher than average, for example, might lead to a misallocation of May resources.

In addition, if we seasonally adjust a nonseasonal series, or build a complex model under the incorrect assumption that a series is seasonal, we will add error to the analysis.⁴ Such a misspecified model "overadjusts" for seasonality; it removes or

³Departures from the general pattern, such as the extremely low observation in May 1979, also appear more clearly in a graph of a seasonally adjusted series than in a graph of the raw data.

⁴For discussions of the problem of overadjustment, see Nettheim (1965), Rosenblatt (1965), Grether and Nerlove (1970: 682-683), or Kalleck (1978). For a discussion of other errors that may result from erroneous assumptions about seasonality in a regression model, see Wallis (1974).

Figure 3

ILLINOIS INDEX LARCENY-THEFT, 1972 TO 1981

RAW DATA SERIES = 0
MULTI-SEGMENT LINE = 0
SOURCE: SAC EDITION ILLINOIS UNIFORM CRIME REPORTS
OFFENSE DATA, 1981 PRELIMINARY
INDEX LARCENY-THEFT = THEFT, THEFT FROM MOTOR VEHICLE, ATTEMPTS

FIRST SLOPE = 247.00 FIRST TURNING POINT 5412.50
T ZERO INTERCEPT = 18617.05 T=1974.25
SECOND SLOPE = -1.20 TOTAL SSA = 220449461.50
T ZERO INTERCEPT = 27269.22

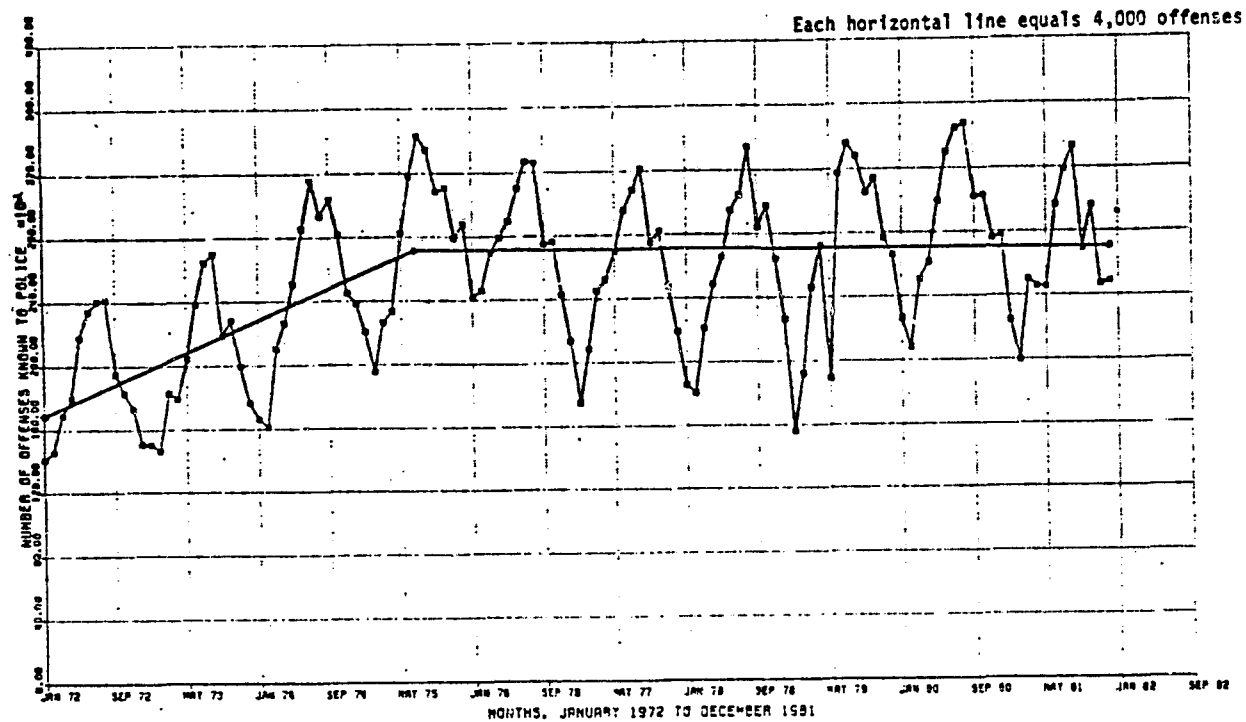
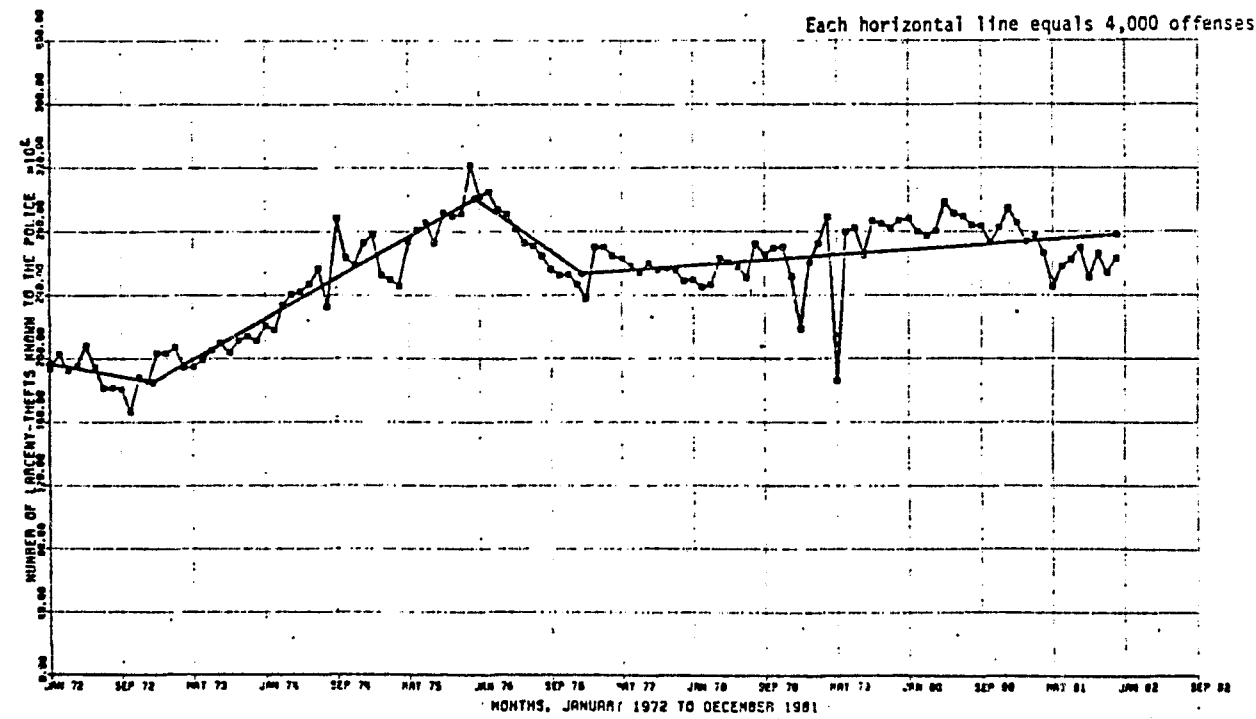


Figure 4

ILLINOIS INDEX LARCENY-THEFT, SEASONALLY ADJUSTED

RAW DATA SERIES = 0
MULTI-SEGMENT LINE = 0
SOURCE: SAC EDITION ILLINOIS UNIFORM CRIME REPORTS
OFFENSE DATA, 1981 PRELIMINARY
INDEX LARCENY-THEFT = THEFT, THEFT FROM MOTOR VEHICLE, ATTEMPTS

FIRST SLOPE = -181.52 FIRST TURNING POINT 5412.50
T ZERO INTERCEPT = 18782.10 T=1974.25
SECOND SLOPE = 531.52 SECOND TURNING POINT 5448.50
T ZERO INTERCEPT = 1984.22 T=2007.00
THIRD SLOPE = -394.85 THIRD TURNING POINT 5482.10
T ZERO INTERCEPT = 48297.45 T=2318.07
FOURTH SLOPE = 61.25 TOTAL SSA = 258210100.01
T ZERO INTERCEPT = 22854.08



otherwise controls for seasonal fluctuation that never existed. This produces a series that is negatively seasonal -- observations twelve months apart are negatively associated with each other. Not realizing that the negative seasonal pattern is the result of, not the reason for, statistical manipulation, the analyst may then correct the model for this imaginary seasonality. If the model becomes complex, it may be very difficult to detect this error.

Thus, if we knew a priori that some variable fluctuated with the seasons, it would be a good idea to take seasonality into consideration when we analyzed, or based any decision upon, the series.⁵ Conversely, if we had reliable evidence that a variable did not fluctuate with the seasons, we would know that a model of that variable would be misspecified if it incorporated a seasonal assumption. In practical situations, however, we usually do not know whether a series is seasonal or not. Therefore, in order to avoid both of these errors -- assuming a series is seasonal when it is not and assuming a series is not seasonal when it is -- an analysis of monthly or quarterly data should begin with the question: Is this series seasonal?

⁵Even when we know a series is seasonal, some decisions may require the actual raw data, not the seasonally adjusted data. As Fromm (1978:26) argues, "It does not help workers seeking jobs to tell them that seasonally adjusted they are employed." Consumers have to pay the actual price, not a seasonally adjusted price, for out-of-season fruits and vegetables. The prison administrator must find a bed for each new prisoner, without regard to whether the prisoner is part of a seasonal fluctuation or not.

WHAT IS SEASONALITY?

To answer the question, "Is this series seasonal?" we must first define seasonality. As Granger (1978:35) notes, "It is remarkable how many papers discuss [seasonality] without consideration of definition." It is not surprising that two investigators would come to conflicting conclusions about the presence of seasonal fluctuation in a series, if neither began the analysis with a definition of seasonality.

Such a definition needs to be more than a mathematical formula. The method used to calculate the presence of seasonality should have some basis in the analyst's concept of what seasonality is. For example, if we conceive of seasonal fluctuation as being relatively constant from year to year, consistency should be included in the measure of seasonality. By not explicitly stating our definition of seasonality, we risk using a measure that conflicts with that definition, and the analysis will yield confusing if not erroneous conclusions.

To avoid this, we need a clear conceptual definition of seasonality. There are two major empirical approaches to defining and detecting the presence of seasonality, the component approach and the stochastic approach. Although these two traditions are historically distinct, with adherents, literature and jargon that seldom overlap, there is a close mathematical similarity. Each approach can be expressed in terms of the other, and it is possible to combine the two to reap the benefits of both.⁶ However, because they are generally seen as separate approaches to seasonality, the following discussion treats them separately.

The Component Definition of Seasonality

Perhaps the most common conceptual approach to seasonality is the component approach, expressed by Kallek's (1978:15) simple and straightforward definition:

Seasonality refers to regular periodic fluctuations which recur every year with about the same timing and with the same intensity and which, most importantly, can be measured and removed from the time series under review.

A series with strong seasonal fluctuation, such as long gun registrations, (figure 5) easily qualifies as seasonal under Kallek's definition. However, the seasonality present in many

⁶This combination, an "X-11 ARIMA" method (Dagum, 1978, 1980), uses a stochastic time series model to improve the quality of X-11 forecasts. For definitions of these terms, see "The Stochastic Definition of Seasonality," below.

Figure 5

CHICAGO LONG GUN REGISTRATIONS, JANUARY 1969-JULY 1980

SOURCE: GUN REGISTRATION SECTION OF THE,
CHICAGO COMPTROLLER'S OFFICE

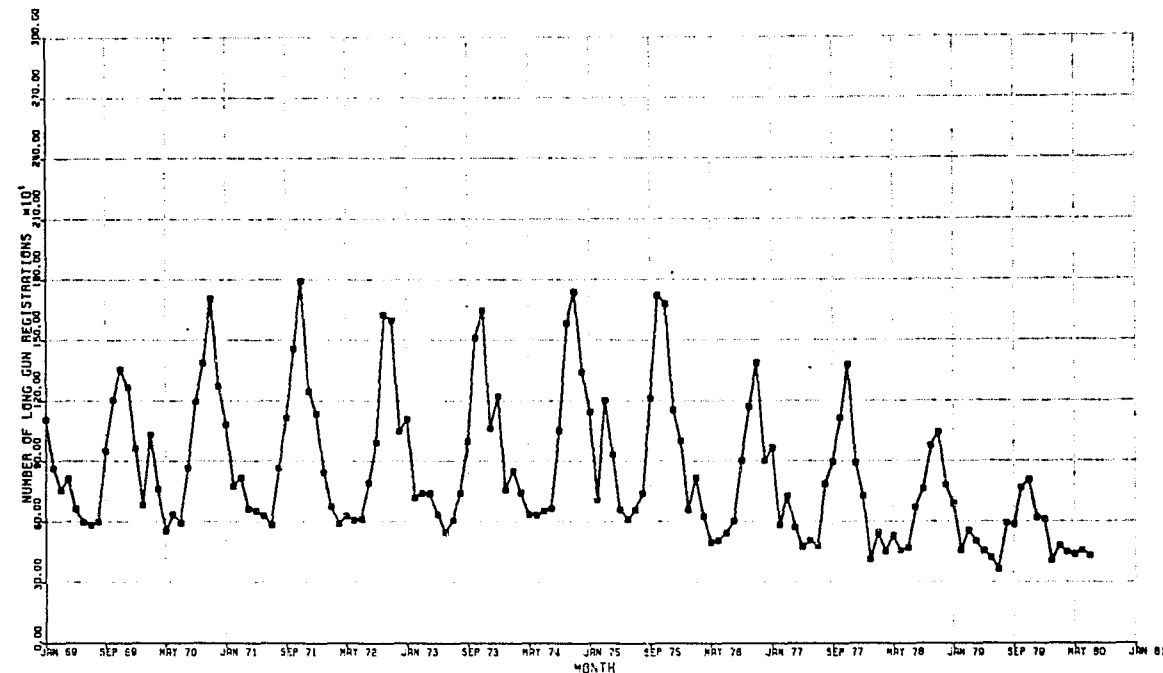
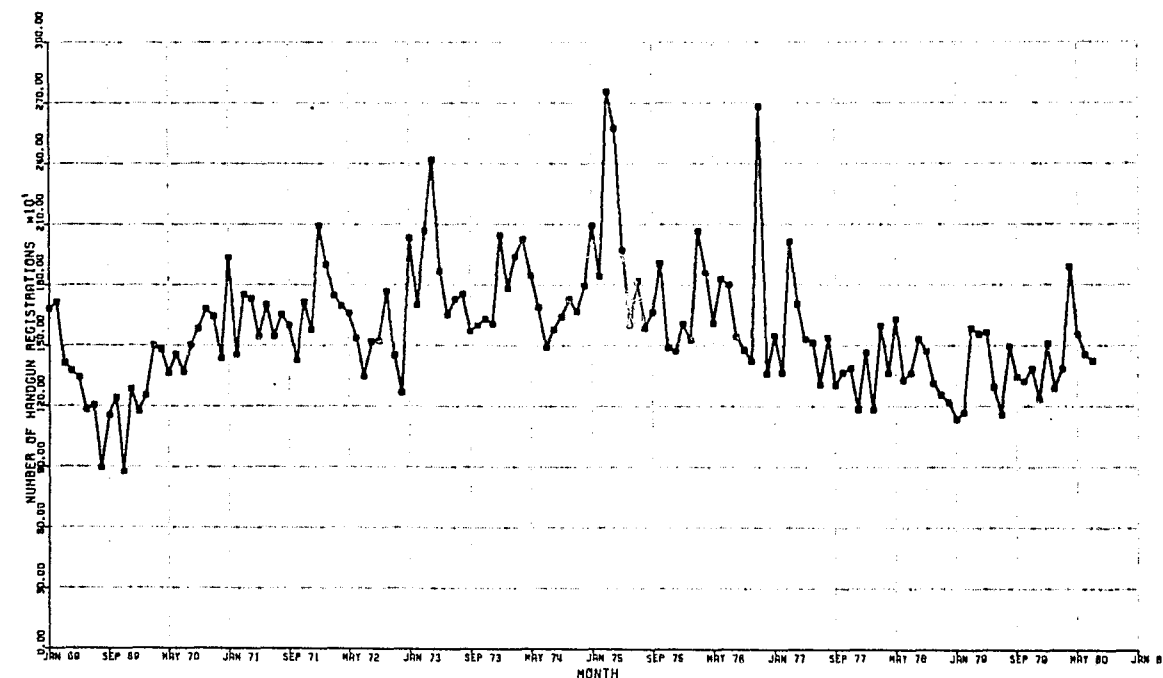


Figure 6

CHICAGO HANDGUN REGISTRATIONS, JANUARY 1969-JULY 1980

SOURCE: GUN REGISTRATION SECTION OF THE,
CHICAGO COMPTROLLER'S OFFICE



crime series, such as handgun registrations (figure 6) is less obvious, and categorizing the series as "seasonal" or "not seasonal" becomes a subjective question. To reduce the subjectivity, or at least make it explicit, we need measures for aspects of the conceptual definition, such as "regular periodic fluctuations," "same timing," and "same intensity." For example, what if all summers were high except one, and that summer were abnormally low? What if the degree to which the summer months were high were less than the degree to which the summer months varied among themselves? In such cases, which in crime series are very common, we need objective criteria to measure, or "operationalize" Kallek's definition. The component approach operationalizes seasonality by separating seasonal fluctuation from the rest of the series.

The final clause of Kallek's definition, that seasonal fluctuation "can be measured and removed from the time series under review," is the foundation of the component approach. The analyst imagines that each seasonal series has three components. The trend/cycle component consists of long-term trend and any nonseasonal but regular fluctuations. The seasonal component is, "the intrayear pattern of variation which is repeated constantly or in an evolving fashion from year to year," (Shiskin et al. 1967:1). The irregular component consists of everything else, the "residual variation." Thus, the total number of occurrences in a given month equals the number due to the trend/cycle, the number due to seasonality, and the number due to irregular fluctuation.⁷ A "seasonally adjusted" series is a series from which the seasonal component has been removed. It has all the characteristics of the original, except seasonal fluctuation.

The component approach is commonly referred to as "seasonal adjustment," or as "Census X-11 adjustment." Since 1954, when the U.S. Bureau of the Census introduced an early version of the X-11 seasonal adjustment program, it has become one of the standards against which seasonal adjustment methods are measured.⁸ It is widely used by both governmental agencies and academic scholars in the United States and elsewhere. When you see economic data labeled "seasonally adjusted," with no other qualifying statement, you can usually assume that the data were seasonally adjusted by the X-11 program, or some version of it.⁹

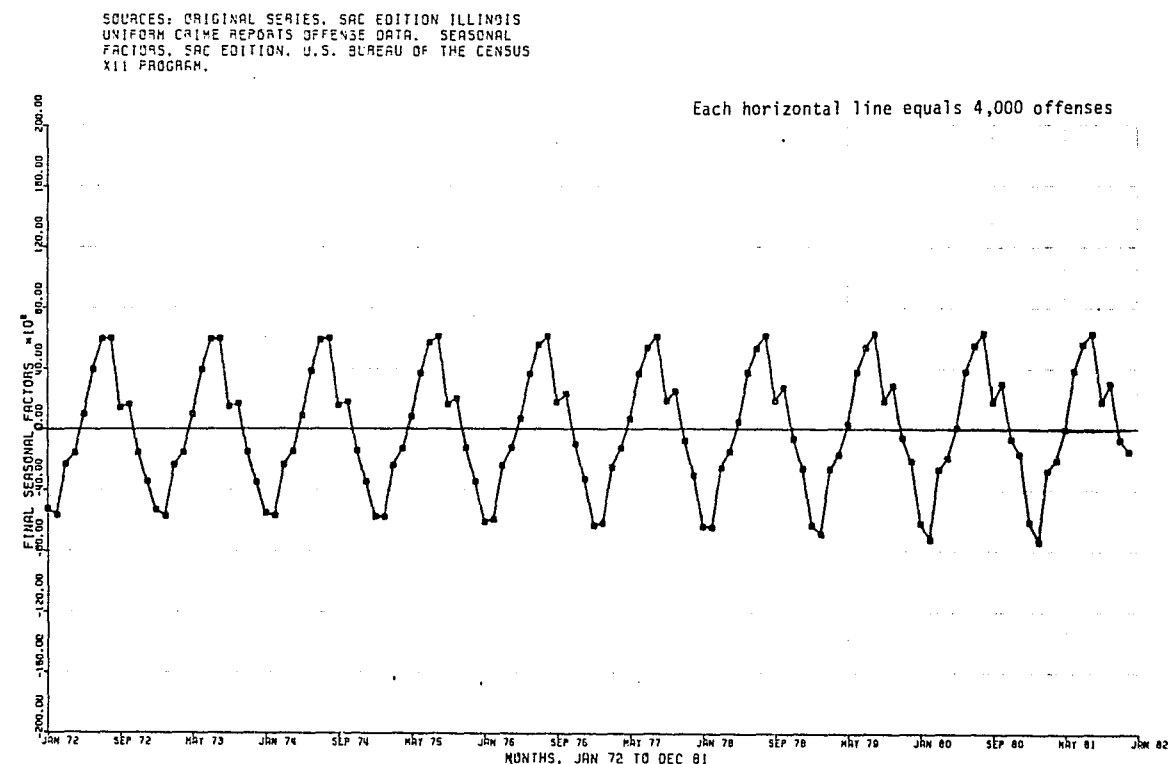
⁷The relation between components may be additive or multiplicative. See "Component Methods," below.

⁸For more information on the Census X-11 and other seasonal component methods, see Shiskin (1967), Plewes (1977), Grether and Nerlove (1970), Hannon (1960,1963), Lovell (1963), Willson (1973), Nettheim (1965) and Rosenblatt (1965). In addition, more than twenty papers on aspects of seasonal adjustment and analysis are contained in the Census Bureau publication, Seasonal Analysis of Economic Time Series (Zellner, 1978).

⁹Bell Laboratories has recently developed a program, called SABL, that is similar in concept to the X-11, but contains several improvements. See Cleveland et al. (1978).

Thus, the problem of detecting seasonality becomes a problem of dividing a series into its three components. The usual method for doing this is to smooth the series by some variation of a moving average, isolate the seasonal component, and then remove it. (For details, see "Component Methods," below.) Once the seasonal fluctuation has been separated from the rest of the series, the component method uses a variety of statistical tests, which compare the removed seasonal component to the trend/cycle and irregular components, as criteria for the presence of seasonality. If the seasonal component is large enough relative to the irregular component, then the component approach decides the series is seasonal. For example, the three components of the seasonal larceny/theft series are graphed in figures 7, 8, and 9. Figure 7 shows the seasonal fluctuation, figure 8 shows the irregular, and figure 9 shows the trend/cycle.¹⁰

Figure 7
ILLINOIS INDEX LARCENY-THEFT, FINAL SEASONAL FACTORS



¹⁰These components were calculated by the X-11 program under the additive assumption. The F value for the amount of variation in the seasonal relative to the variation in the irregular is 96. For details, definitions, and other seasonal component analysis examples, see the section, "Component Methods," below.

Figure 8
ILLINOIS INDEX LARCENY-THEFT, FINAL IRREGULAR SERIES

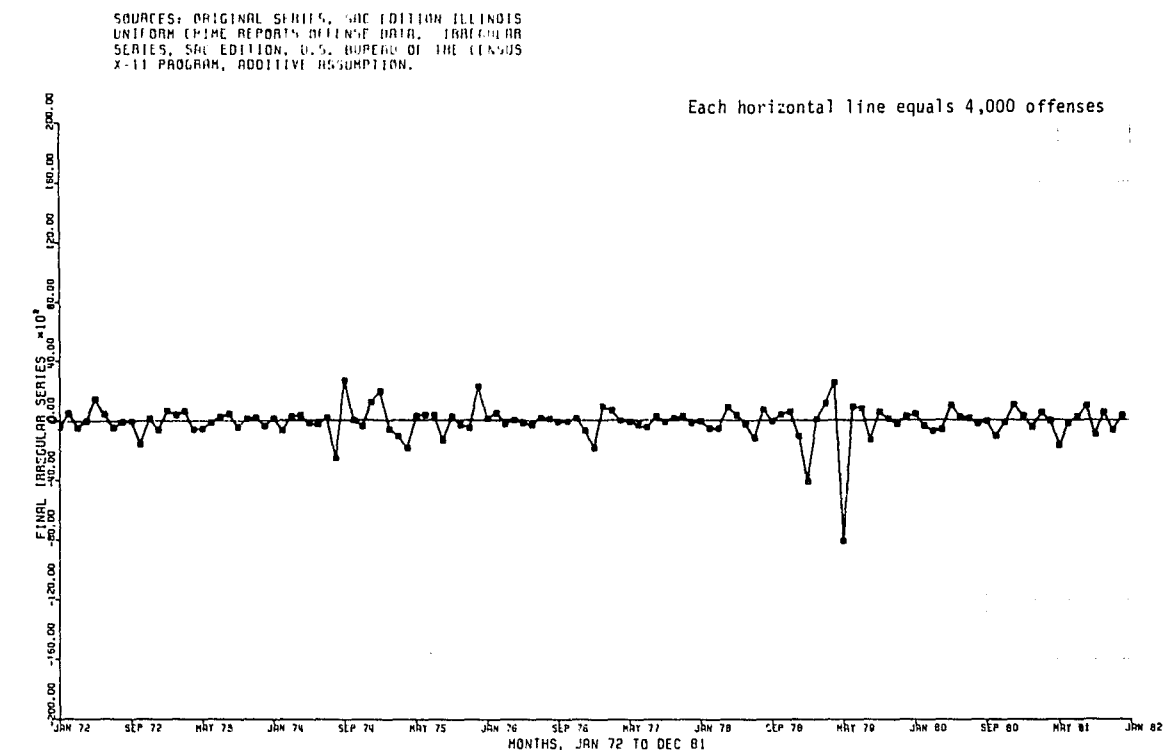
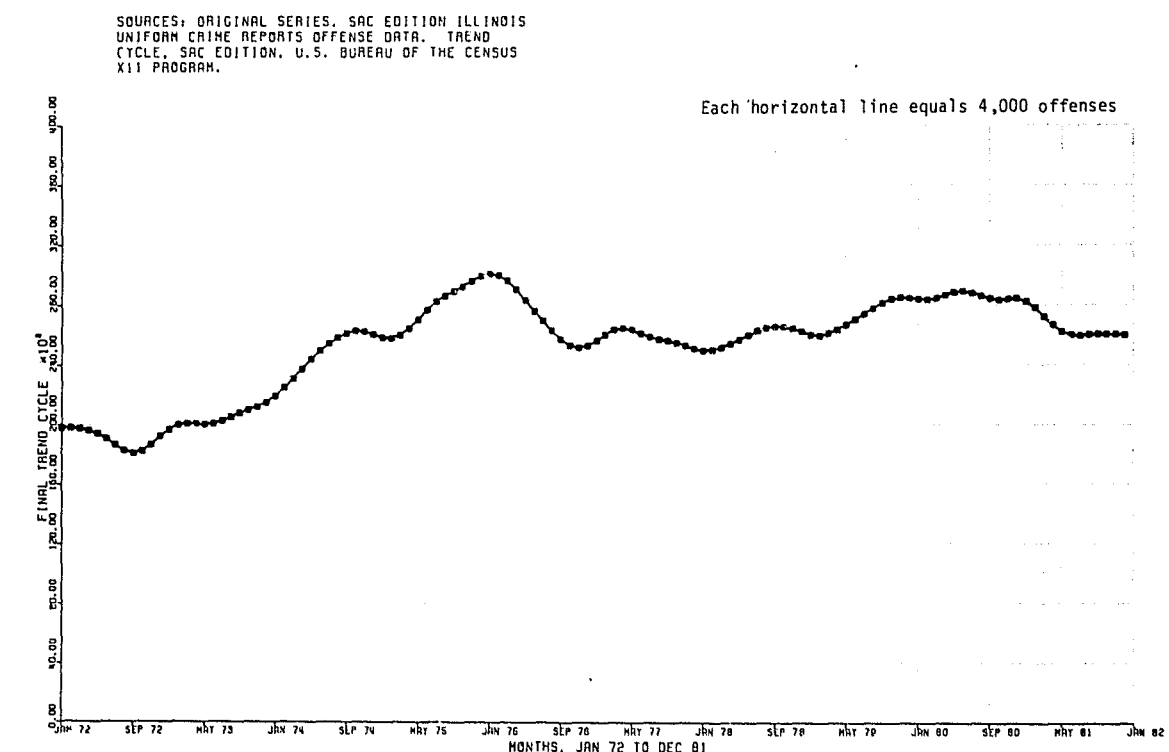


Figure 9
ILLINOIS INDEX LARCENY-THEFT, FINAL TREND CYCLE



The Stochastic Definition of Seasonality

The foundation of the component approach to seasonality is its conception of seasonal fluctuation as separate from the rest of the series. In contrast, the "stochastic time series analysis" or "autoprojection" method incorporates seasonal fluctuation into a single descriptive model of the series, a description of the stochastic process.¹¹ In a stochastic process, one observation follows the next with a certain probability. In a series with seasonal fluctuation, observations twelve months apart are correlated, which means that they follow each other with a certain probability.¹² Thus, seasonality may be part of a stochastic process.

In the stochastic literature, as in the component literature, it is unusual to find an explicit conceptual definition of seasonality. The closest thing to such a definition in Box and Jenkins (1976:301) is the following:

In general, we say that a series exhibits periodic behavior with period s , when similarities in the series occur after s basic time intervals.

Nelson (1973:168) paraphrases this in less mathematical language:

Seasonality means a tendency to repeat a pattern of behavior over a seasonal period, generally one year.

Therefore, like the component definition, the stochastic definition of seasonality emphasizes the existence of regular periodic fluctuation. However, unlike the component definition, the stochastic definition does not emphasize separating this fluctuation from the rest of the series.

An additional difference between the two approaches is that the stochastic approach is not so much concerned about describing the past as it is about forecasting the future. Box and Jenkins (1976:301) emphasize that a reasonable description of each of the three components of a series may not produce a good forecast. Of course, description of the past must precede a forecast of the

¹¹Stochastic time series analysis is also called Box-Jenkins or ARIMA, in reference to the authors and to their mnemonic for the method. There are many approaches to the problem of stochastic time series modeling, and an extensive literature on the subject, some of which requires a knowledge of advanced statistics. This report is only a brief guide to these methods, especially as they pertain to seasonal analysis. For a more complete, but still elementary introduction, see Nelson (1973) or McCleary and Hay (1981).

¹²However, the opposite is not always true. If observations twelve months apart are correlated, the series is not necessarily seasonal. See "Stochastic Methods," below.

future. However, the stochastic approach does not describe the series with a regression or harmonic function, or describe the separate components. Instead, it describes the stochastic process of the series. If we can describe the probability, or set of probabilities, under which observations followed one another in the past (the stochastic process), and if the same process continues unchanged, then we can accurately forecast the future.

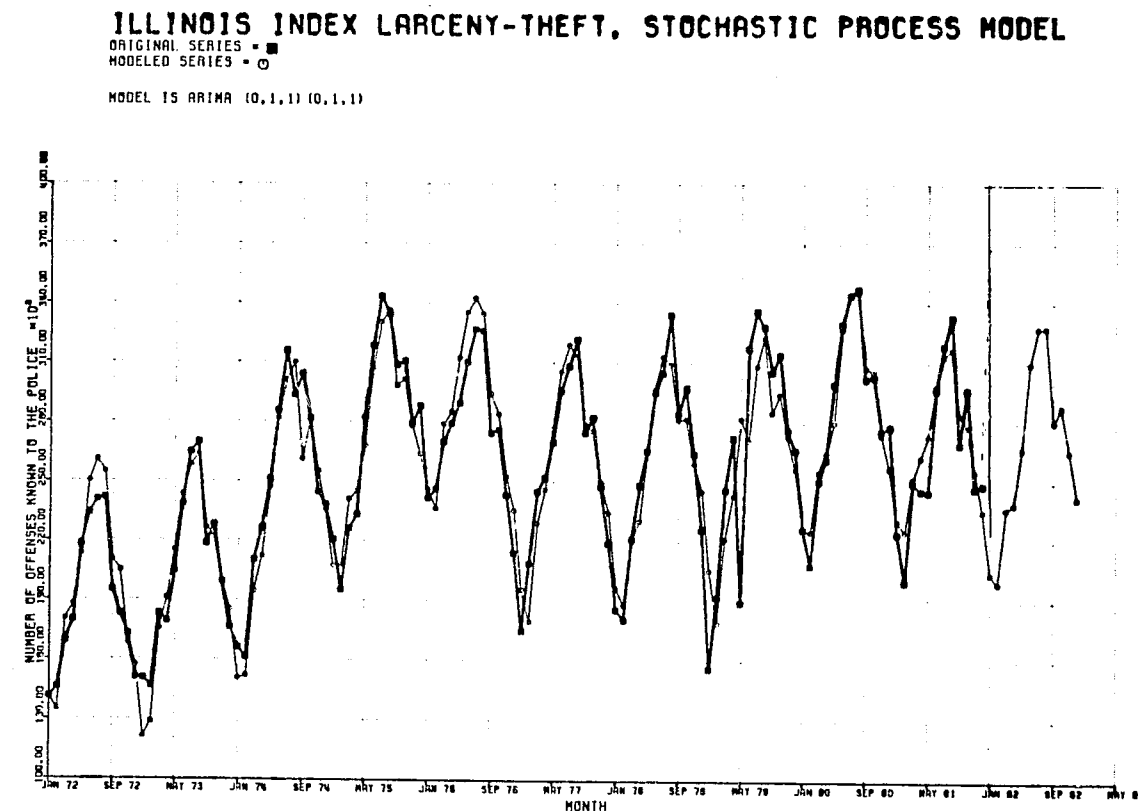
Stochastic time series analysis assumes that a times series has followed some unknown stochastic process in the past. The problem is to identify or "model" that process.¹³ The stochastic method uses trial and error, "iterative decisions," to arrive at the best model. These iterative decisions begin with an initial diagnosis of the series. With respect to seasonality, the analyst uses descriptions of the relationship between observations at one time period and another to discover any systematic seasonal movement. If this diagnosis suggests seasonality, the analyst then considers a number of alternative seasonal processes that may describe the pattern of the series. Each set of alternative processes becomes a tentative model, which the analyst then evaluates. By fitting a tentative model to the series, calculating the difference between the model and the actual data, and analyzing these "residuals," the analyst determines whether or not the chosen seasonal process successfully describes the series. Eventually, the analyst reaches a stochastic model that appears to describe the series better than alternative models. If this final model includes a seasonal term, the stochastic approach decides that the series is seasonal.

For example, stochastic diagnosis, estimation and evaluation of the Illinois larceny/theft series suggests that the series follows a seasonal process.¹⁴ The current observation is related to the error of the observation twelve months ago. In addition, the current observation is related to the error of the preceding observation. (For details, see "Stochastic Methods," below.) Figure 10 shows the original larceny/theft series (dark line) and the modelled series (light line). Each of the modelled values was calculated from the immediately preceding observation and the observation one year ago. For the years 1972 through 1981, we used the actual numbers of larceny/theft to calculate the modelled value. The January, 1982 modelled value was calculated from the actual December, 1982 and January, 1981 values. To forecast for February, 1982, we used the actual February, 1981 observation and the modelled value for January, 1982. By continuing this process, we calculated the 1982 forecasted values in figure 10.

¹³A model is "a set of assumptions concerning the origin or generating mechanism of a series," (Pierce, 1980:125).

¹⁴The model realized in figure 10 is a (0,1,1)(0,1,1) ARIMA process, or a first-order seasonal and serial moving average process with seasonal and serial differencing. For definitions and details, see "Stochastic Methods," below.

Figure 10



Summary: Two Definitional Approaches

This section discussed two major approaches to defining and detecting seasonality, the component approach and the stochastic approach. Although the two approaches are mathematically similar, there are practical differences in emphasis. The component approach emphasizes a separate description of seasonal fluctuation, while the stochastic approach emphasizes forecasting the future with a model that incorporates seasonality. The component approach is more interested in seasonality itself, while the stochastic approach is more interested in seasonality as it affects the accuracy of a forecast.

Both approaches model seasonal fluctuation. However, in the component model, seasonality is separated from the rest of the series, while in the stochastic model, it is not. There are two schools of thought concerning the separation of seasonal fluctuation from the rest of the series. One school (see Kendall, 1976: 66) argues that, since seasonality is variation due to a known cause, it should be removed prior to building an explanatory model, forecasting, or any other complex analysis. The other school (see Plosser, 1978) holds that it is more logical to include seasonal fluctuation as an integral part of the final analysis. The first school of thought would use the component method, but the second would not.

In reality, models of separate components are necessary to answer some questions, and one model incorporating seasonality is necessary to answer other questions. For example, a decision to build a new prison will depend upon a forecast of the total number of inmates, with seasonal fluctuation included in the total. On the other hand, if there are wide seasonal fluctuations in the number of inmates, it might be necessary to open an additional wing during some months of the year. The decision to do this would depend upon an analysis of the seasonal component.

There have been several experimental comparisons of various approaches to detecting the presence of seasonality (Kuiper, 1978; Granger, 1978; Grether and Nerlove, 1970; Buchin, 1982). However, Kendall and Stuart (1966) probably give the best advice: "Try several methods and choose the one which appears to give the best results." No single method of analysis is appropriate in every situation. The method of choice depends upon the objectives of the analysis. In the following sections of this report, we discuss and compare the tools for detecting and analyzing seasonality that are offered by the two approaches, and give the analyst some basic rules of thumb for using these tools in various practical situations.

TOOLS FOR DETECTING AND ANALYZING SEASONALITY

Neither the component nor the stochastic approach to seasonality offers a simple, objective, yes-or-no criterion for detecting the presence of seasonality in a time series. Both approaches depend heavily on the judgment of the analyst, although each approach gives the analyst a number of statistical tools upon which to base that judgment. In the following sections, we introduce the reader to some of these tools for detecting, measuring and adjusting for seasonality.

Component Methods

The X-11 program, developed by the U.S. Bureau of the Census, gives the user a vast amount of information that can be used to answer the question, "Is this series seasonal?"¹⁵ If the answer is "yes," the X-11 program allows the user to describe both the seasonal fluctuation and the pattern of the series with the seasonal fluctuation removed.

The X-11 program partitions a series into three components (seasonal, trend-cycle, and irregular) by fitting a moving average to it (for details, see below). This smooths the series, and allows seasonal fluctuation, if present, to be isolated and described.

The standard output of the X-11 program is voluminous, and its interpretation is an art as much as it is a science. The user must weigh the results of various diagnostic tests against each other, and make a number of subjective judgments. The final decision as to whether or not a given series fluctuates with the seasons is a function of the analyst's interpretation of these diagnostics. Two analysts may disagree. Thus, published results should mention the diagnostic tests the analyst used to arrive at the decision, and the results of those tests.

Pierce (1980:130) argues that, "seasonal adjustment models are never more than approximations." However, the objectivity of

¹⁵The Bureau of Labor Statistics (BLS) of the U. S. Department of Labor provided the X-11 program we use at SAC. BLS staff also provided documentation, and were very helpful in answering questions about interpretation. The X-11 is available in the SAS/ETS (Econometric and Time Series) package. In addition, an alternative component program, SABL, is available from Bell Labs (Cleveland et al., 1978). We use an abbreviated component program as a screener. This "Bell-Canada" package was developed by John Higginson of Statistics Canada. For instructions on using both of these programs on the SAC system, see Miller (1982).

these approximations can be improved if analysts use the same diagnostic tests, interpret these tests using general guidelines or "rules of thumb," and explicitly state any deviations from the use of these guidelines. At the Statistical Analysis Center, we have found the following guidelines to be helpful.

Moving Average

A moving average calculation takes successive averages from the beginning to the end of the series.¹⁶ For example, it might calculate the average of observations 1 through 5, then observations 2 through 6, observations 3 through 7, and so on to the end of the series. These means are, then, a transformed series in which random variation and periodic fluctuation (within a five-month span) are "averaged out." A moving average is smoother than the original series, and, depending on the number of observations within each average, does not contain periodic fluctuation.¹⁷ It is also shorter than the original series. This "end effect" is often important, because we may be most interested in the most recent past of the series.¹⁸

The goal of a moving average is to produce a smoothed series that does not contain random variation or periodicity, but still contains the other patterns in the series. However, there are many kinds of moving average, and not every kind will meet this goal for every series.¹⁹ According to Kendall (1976:53),

Trend-fitting and trend-estimation are very far from being a purely mechanical process which can be handed over regardless to an electronic computer. In the choice of the extent of the average, the nature of the weights, and the order of the polynomial on which these weights are based, there is great scope -- even a necessity -- for personal judgment. To a scientist

¹⁶The concept and calculation of moving average, in this context, is very different from the moving average process in stochastic time series analysis. The moving average (MA) process received its name because it is similar to a conventional moving average in one way: it assumes that each observation is affected by a finite number of other observations. For more detail, see Nelson (1973:33).

¹⁷In the context of spectral analysis, a moving average is called a "filter." For example, a "low-pass filter" removes high frequency periodicity. See Kendall (1976:44) and the section, "Cumulative Periodogram of Residuals," above.

¹⁸There are various statistical techniques to handle the end effect in a moving average. See Kendall (1976) for a review. The X-11 ARIMA method uses stochastic time series analysis to estimate end values. See Dagum (1978,1980).

¹⁹For a clear discussion of the effects that various moving averages have on a series, see Kendall (1976:29-54). For ratio-to-moving average, see Hickman and Hilton (1971).

it is always felt as a departure from correctness to incorporate subjective elements into his work. The student of time-series cannot be a purist in that sense.

Therefore, the X-11 program utilizes iterative approximations of the best moving average, and offers the user a choice of moving average options (see Shiskin et al., 1967 for details).

Additive/Multiplicative Assumption

The seasonal, trend/cycle, and irregular components have two possible relationships to each other, dependent or independent.²⁰ If we consider them to be independent of each other, then we add them together to equal the total number of occurrences. If we consider them to be dependent on each other, then we multiply them together to equal the total number of occurrences. For example, if the relationship for larceny were additive, then the number of larcenies due to seasonal fluctuation would remain the same no matter if the total number of larcenies were 50 or 500. If the relationship were multiplicative, the number of larcenies due to seasonal fluctuation would be greater if the total number of larcenies were higher. The additive/multiplicative assumption is the analyst's choice. Most economic series are assumed to be multiplicative. However, we know of no theoretical argument for assuming the components of a crime series to be either dependent or independent. In our experience, the additive assumption has produced the better adjustment in the majority of crime series analyzed.

Our general procedure at SAC is to make no prior judgment about whether seasonal fluctuation is additive or multiplicative, but to adjust the series under both assumptions, and choose the best adjustment of the two according to the diagnostic tests discussed below.²¹ The two assumptions usually produce very similar results, but, when they do not, we assume that the better adjustment, additive or multiplicative, reflects the true relationship among the components.²²

²⁰For a discussion of additive versus multiplicative relationships in stochastic process models, see Box and Jenkins (1976:322-324). The relationship in stochastic models, as in component models, is usually assumed to be multiplicative.

²¹Of course, it is not possible to adjust under a multiplicative assumption if the series contains an observation that is zero.

²²For an important series in which the multiplicative or additive relationship is not clear, it may be necessary to use more complex analytical methods than this report can cover. The literature on seasonal adjustment contains many discussions of the problem. For an introduction, see several of the papers in Zellner (1978). For some practical hints, see Plewes (1977).

F of Stable Seasonality, and
Relative Contribution of the Irregular

The F of stable seasonality is a ratio between the seasonal component and the irregular component.²³ The F value's significance is based on the assumption that the irregular is normally distributed, homoscedastic, and varies randomly over time (Shiskin et al., 1967:59). With time series data, the assumption of independence of successive observations may be violated (Anderson, 1950). Therefore, there is some question as to the proper interpretation of this F value.²⁴

Seasonal series typically have very high F values. The stable seasonality F is 96, for example, for the Illinois larceny/theft series, and it is not unusual to find an F value of 100 or more. In light of this, how should we interpret an F that is much smaller, but not small enough to be statistically insignificant, providing we could assume independence? If we cannot apply the usual significance tables, what does an F value of 5 or 10 indicate about the presence of seasonality?

As a guide to interpreting such X-11 results, Plewes (1977) prepared a set of "rules of thumb" for the staff of the Bureau of Labor Statistics. We have found these guidelines to be very helpful, and describe some of them here. Plewes suggests that interpretation of the stable seasonality F value should be guided by information about the irregular. This makes sense, when we realize that the assumptions upon which the F is based have to do with the behavior of the irregular.

Another diagnostic computed by the X-11 program, the "relative contribution of the irregular" varies from 0% to 100%, and indicates the contribution of the irregular component to total month-to-month variation, relative to the contributions of the seasonal and the trend/cycle components. It indicates the absolute importance of each component to the variation in the total series.²⁵ Plewes (1977:4) suggests that the F value should be interpreted in light of the relative contribution of the irregular, according to the following rule of thumb:

²³Stable seasonality assumes that seasonal fluctuation is constant from year to year. For "moving seasonality," which does not assume consistency, See the "Appropriate Applications" section, below.

²⁴If we could assume independence of observations and use the F table, a value of 2.41 would be significant. This is the 1% level for a 10-year series. Differences in significance levels for series of other lengths are negligible (Shiskin et al., 1967:59).

²⁵In X-11 printed results, we also find the relative contributions of each of the three components over a two-month span, three-month span, and so on, up to a twelve month span.

<u>F Value</u>	<u>% Cont. of I</u>	<u>Decision</u>
0.00-2.40	> 0%	no stable seasonality
2.41-15.00	> 14%	no stable seasonality
15.01-50.00	> 25%	no stable seasonality
50.01 and up	> 30%	no stable seasonality

To this rule of thumb, we would add a qualification.²⁶ The percent contribution of the irregular reflects the relative contributions of both the seasonal and the trend/cycle. In crime series, in contrast to many economic series, the contribution of the trend/cycle may be very low. As a result, both the irregular and the seasonal relative contributions may be high. Therefore, with a stable seasonality F value over 15 and a percent contribution of the irregular about 30, before rejecting the stable seasonality hypothesis, check the percent contribution of the seasonal. According to Plewes (1977:7) "a seasonal component with a [relative contribution] value of less than 50.0 percent in a one-month span signals a weak seasonal." If the seasonal contribution is 50 percent or more, use additional diagnostics (see below) to make the final decision.

Therefore, even if it cannot be interpreted as an exact statistic, the F of stable seasonality can be used in an exploratory way as one indicator of the amount of seasonality in a series. For example, as we mentioned above, the stable seasonality F value for Illinois larceny/theft is 95.82. The contribution of the irregular over a one-month span is 18 percent. According to Plewes's rule of thumb, we should not reject the hypothesis of stable seasonality. In contrast, for Illinois Index homicide (figure 2) the stable seasonality F value is 2.78 and the contribution of the irregular is 70 percent. This indicates that the series does not contain stable seasonality. On the other hand, for Index aggravated assault (figure 1) the stable seasonality F value is 45.70, and the contribution of the irregular is 38 percent. According to Plewes's rule of thumb, we should reject the hypothesis of stable seasonality. However, the contribution of the seasonal component over a one-month span is 60 percent. Therefore, other diagnostics should be consulted before making the final decision.²⁷

²⁶Kathryn Beale of the Bureau of Labor Statistics, who was very helpful in explaining X-11 interpretation, pointed this out to us.

²⁷The statistics given here are for the additive or multiplicative adjustment, whichever has the highest stable seasonality F. Statistics for the alternative adjustment for these series are very similar.

Average Duration of Run

The average duration of run (ADR) is a simple test of the smoothness of variation over time. By definition, the irregular component varies randomly over time. If it does not, then the calculation of the seasonal adjustment should be suspect.

The ADR is the mean length of runs of values consecutively higher (or lower) than the preceding value. The higher the ADR, the fewer the total number of runs in the series. If the irregular ADR is lower than would be expected in a random series, the adjustment may have assigned some seasonal or trend/cycle variation to the irregular component. If the irregular ADR is higher than would be expected in a random series, the adjustment may have assigned variation that should be considered irregular to the seasonal or trend/cycle component. An ADR from 1.36 to 1.75 is considered random.

Again, Plewes (1977:8) provides a rule of thumb to interpret the irregular ADR. It is the following:

The ADR of the irregular (I) should fall between 1.36 and 1.75. When values fall outside of this range, the F-statistic and relative contribution of the irregular should be consulted. If both meet their tests, the series may still be accepted.

For example, for Illinois Index larceny/theft, the ADR of the irregular is 1.59. For Index homicide, it is 1.45, and for Index aggravated assault it is 1.51. These ADRs are all within the "random" range, which indicates that the adjustments can be trusted. The irregular components seem to vary randomly over time, as they should. The ADRs indicate that the irregular components do not contain seasonal fluctuation, nor do the other components contain irregular fluctuation.

In our experience, using the X-11 with hundreds of crime and crime-related series, we have found only four series in which the ADR indicated a non-random irregular. Three series with an ADR below the random range are very short (four to six years). One series with a high ADR, Chicago Index assault 1967-1978, is a moving-average transformation of an original series that was collected in units of thirteen police periods per year. This moving average probably has less irregular variation than the original series, resulting in an overly smooth irregular.

Thus, in practice, you may find very few series with an irregular ADR outside the random limits. If you do find one, consider it as a warning that something may be amiss. Look carefully at the series itself for an explanation. In the above examples, the low and high ADRs were apparently related to short series or to unusually smooth series. In any case, do not accept the adjustment unless other indicators, especially the F of stable seasonality and the percent contribution of the irregular, are unequivocal.

Months for Cyclical Dominance

Months for cyclical dominance (MCD) compares the relative contribution of the trend/cycle to the relative contribution of the irregular. As discussed above (see note 25), the standard output of the X-11 program includes a table giving the relative contributions of each of the three components over a one-month span, a two-month span, and so on.

From one month to the next, the irregular usually provides the most visible movement in a series. Thus, the relative contribution of the irregular over a one-month span is usually high. In contrast, the contribution of the trend/cycle to month-to-month variation is usually low. However, the trend/cycle contribution usually builds over time; its contribution over a two-month span is greater than over a one-month span, its contribution over a three-month span is still greater, and so on. Thus, in most series, the relative effect of the trend/cycle gradually increases, until it exceeds the contribution of the irregular. The span at which this occurs is the MCD.

An MCD of 1 means that the percent contribution of the trend/cycle over a one-month span exceeds the irregular contribution. An MCD of 2 means that the trend/cycle exceeds the irregular over a two-month span. In many economic series, the MCD is low. The relative contribution of the trend/cycle is substantial over a one-month span, and increases rapidly, until it exceeds the irregular contribution at the three- or four-month span. However, the contribution of the trend/cycle in many crime series is less than this. Consequently, we have found few crime series that meet Plewes's following rule of thumb:

Series with MCD values of 1, 2, or 3 usually exhibit sufficient smoothness to be acceptable; series with MCD's of 4 or 5 are borderline, and the impact of the irregular should be carefully analyzed; when an MCD of 6 appears, the particular month in which the I/C ratio becomes less than one should be identified (the X-11 program prints no value larger than 6). The decision to publish the series should be made on other grounds, since a long MCD is usually reflective of other problems in the series.

For example, table 1 shows the relative contributions of each component to the total variation in the larceny/theft series (additive adjustment) from a one-month to a twelve-month span. Because the trend/cycle contribution exceeds the irregular contribution for the first time at a five-month span, the MCD for larceny/theft is 5. For comparison, the MCD of the Index homicide series is over twelve months (the trend/cycle contribution never exceeds the irregular contribution). The MCD of the Index aggravated assault series is 6. Notice that the contribution of the seasonal component drops close to zero over a twelve-month span. This makes sense, because, by definition, seasonal values

Table 1
Relative Contributions of Components to Variance
Illinois Larceny/Theft, Additive Adjustment

Span in Months	Irregular	Trend/ Cycle	Seasonal
1	18.16%	1.16%	80.69%
2	7.29	1.49	91.22
3	3.95	1.97	94.07
4	2.65	2.53	94.82
5	2.40	2.97	94.63
6	2.11	3.60	94.29
7	2.15	4.82	93.02
9	3.67	9.70	86.64
11	10.83	39.28	49.89
12	21.01	78.75	0.24

Table 2
Relative Contributions of Components to Variance
Illinois Aggravated Assault, Additive Adjustment

Span in Months	Irregular	Trend/ Cycle	Seasonal
1	38.46%	1.92%	59.62%
2	16.40	2.94	80.66
3	10.50	3.53	85.97
4	8.07	4.07	87.86
5	6.30	4.44	89.26
6	4.79	4.84	90.36
7	5.15	6.36	88.49
9	8.41	12.59	79.00
11	23.55	43.26	33.19
12	36.53	63.20	0.27

twelve months apart are similar to each other. The seasonal differencing technique (see "Stationarity," below) takes advantage of this fact.

A high MCD is a warning that the series may contain so much irregular variation that the presence and degree of seasonal fluctuation cannot be reliably determined. In practice, we have found only a few crime series with an MCD of 3 or 4, and none with an MCD of 1 or 2 (although we commonly find a low MCD in non-crime series). Because the contributions of the irregular, the trend/cycle, and the seasonal add to 100%, a high MCD does

not always indicate that the adjustment should be rejected. If the MCD is high, look at the percent contribution of the seasonal over a one or two-month period. In a series containing little or no overall trend, both the irregular and the seasonal components may contribute more than the trend/cycle component. For example, in the Illinois aggravated assault series, the contribution of the trend/cycle does not exceed the contribution of the irregular until a six-month span (table 2). However, the contribution of the seasonal is 60% over a one-month span. In such a case, consider the possibility that the series may contain relatively weak, but consistent, seasonal fluctuation. Look at other diagnostics, in particular the final seasonal factors (see below).

Pattern Consistency

Consistency in the seasonal pattern is another important consideration in determining whether or not a series is seasonal. Both the component and the stochastic approaches include consistency, or regularly evolving fluctuation, in their conceptual definition of seasonality.²⁸ While a gradual change from year to year may indicate moving seasonality (see "Appropriate Applications," below), abrupt change and change in sign from one year to the next argue against the hypothesis that the series is seasonal, by this definition.

There are two kinds of seasonal consistency: year-to-year, and within-season. For example, if April observations are very high in four scattered years of a ten-year series, and very low in the other years, then April is not consistently high; the series does not have a consistent pattern of seasonal fluctuation from year to year. Similarly, we should conclude that a certain season tends to be high only if each month of that season tends to be high. For example, if June is always slightly high over a ten-year period, and July and August are very high, then we might say that summers are generally high. On the other hand, if June is always high, July is low, and August is high, then all we can say is that the patterns of the summer months vary.

The "seasonal factor" table, produced by the X-11 program, allows us to examine year-to-year and within-season consistency.²⁹ The seasonal factors indicate, for each month of the series, the amount by which it is high or low due to seasonal fluctuation. There are 144 seasonal factors in the seasonal factor table of a 12-year monthly series.

²⁸See Warren et al. (1981) for an example of an analysis of seasonality that does not include year-to-year consistency in the definition.

²⁹For a similar diagnostic check for consistency, but using stochastic methods, see Thompson and Tiao (1971:540-541).

In a multiplicative adjustment, the seasonal factors show the relative seasonal weight of each month. In an additive adjustment, the seasonal factors show the absolute amount by which the month is high or low. Thus, in a multiplicative adjustment, the seasonal factors range from .00 to 1.99, with 1.00 indicating an average month with no seasonal fluctuation. In an additive adjustment, the seasonal factors range above and below zero, and the scale depends upon the particular data. For example, in a homicide series, a seasonal factor of +20 for a certain month indicates that that month was seasonally high by about 20 homicides. With the standard deviation, which the table includes, you can decide whether a month was high, low, or average.

The seasonal factors (multiplicative adjustment) of homicides of male victims in Chicago from 1965 through 1978 (table 3) show that, while some months may be decernably high and others low in the number of male homicide victims, there is no consistent pattern from year to year.³⁰ January changes over time from an average month to a low month. March begins as an average month, but becomes high in later years, while April begins average but becomes low. July, August, September, October, and November, all change their seasonal factors over the time period. Only one month, February, is consistently high or low, although some argument could be made for July being high. If we consider all the evidence, including the lack of seasonal consistency, the low stable seasonality F (4.00), the high relative contribution of the irregular (63%), an irregular ADR of 1.52 (indicating that the irregular does not contain any seasonal fluctuation), and an MDR higher than twelve months, it becomes difficult to argue that murders of males occur seasonally.

Trading Day Option

The X-11 package provides a "trading-day adjustment" that gives the user an idea of the importance of each day of the week. The adjustment counts the number of Mondays, Tuesdays, and so on, in each month of the series, and determines whether months with three Mondays (for example) differ from months with five Mondays. The program then calculates weights for each day of the week, and computes standard tests of significance for each day. Thus, X-11 trading-day statistics are not a result of direct analysis of the effect of each day of the week. Rather, they are estimated from aggregate data.

³⁰We chose this homicide series as an example because it evidences the most seasonality of any homicide series we have analyzed (Block and Block, 1980; Block et al. 1982). However, it is commonly assumed that homicide is seasonal (Wolfgang, 1966; President's Commission, 1967; Warren et al. 1981). These differing conclusions are probably due to differing definitions and measures of both homicide and seasonality.

Table 3

Final Seasonal Factors, Multiplicative Adjustment
Male Homicide Victims, Chicago: 1965-1978

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1965	1.04	0.76	0.98	1.06	0.95	1.03	1.05	1.14	1.01	1.13	0.83	1.03
1966	1.01	0.77	0.97	1.06	0.96	1.02	1.04	1.15	1.02	1.13	0.85	1.04
1967	0.98	0.79	0.94	1.05	0.99	1.01	1.04	1.15	1.04	1.13	0.89	1.03
1968	0.94	0.82	0.91	1.04	1.01	1.01	1.04	1.13	1.05	1.12	0.92	1.02
1969	0.90	0.85	0.88	1.02	1.03	1.00	1.07	1.13	1.05	1.11	0.95	1.01
1970	0.87	0.88	0.88	1.01	1.03	1.01	1.11	1.12	1.02	1.09	0.95	1.01
1971	0.87	0.89	0.90	0.98	1.03	1.02	1.16	1.12	1.00	1.07	0.94	1.01
1972	0.87	0.89	0.95	0.94	1.03	1.03	1.19	1.09	0.99	1.07	0.92	1.02
1973	0.88	0.87	1.02	0.88	1.03	1.04	1.21	1.06	1.00	1.09	0.91	1.02
1974	0.87	0.85	1.08	0.85	1.03	1.04	1.20	1.02	1.04	1.09	0.91	1.03
1975	0.86	0.82	1.14	0.83	1.01	1.04	1.18	1.00	1.08	1.08	0.92	1.05
1976	0.84	0.80	1.17	0.83	0.98	1.04	1.16	0.98	1.13	1.07	0.94	1.07
1977	0.83	0.79	1.20	0.83	0.96	1.03	1.15	0.99	1.16	1.06	0.94	1.08
1978	0.82	0.78	1.21	0.83	0.94	1.03	1.15	0.99	1.17	1.05	0.94	1.08

Therefore, analysts who are primarily interested in diurnal periodicity might want to analyze daily data, if available, in preference to estimates from monthly data. On the other hand, use of the trading-day adjustment is quicker and less expensive than conducting an extensive analysis of daily data. It may uncover effects that might be overlooked by other methods. To utilize the advantages of both approaches, use them sequentially. The X-11 program allows the user to set a priori weights for days of the week. A direct analysis of daily data may provide the information with which to set these daily weights.

However, there are limits to the use of the trading-day option. It will not provide accurate estimates when the contribution of the irregular over a one-month span is eight percent or more (Shiskin et al., 1967). Because most crime series are more irregular than this, the trading-day option can seldom be used with crime data.

Appropriate Applications

At the Statistical Analysis Center, we have found component seasonality methods to be very useful in the initial description of a series. Since the X-11 program is relatively simple to use, it is especially appropriate when the patterns in a large number of series, for example the 714 series of seven Index crimes in Illinois' 102 counties, must be described and compared to each other. It is also appropriate when the decision at hand requires a separate description of the pattern of seasonal fluctuation, or the pattern of the series adjusted for seasonality.

The X-11 program is not appropriate for highly irregular series, short series (six or fewer years), or for series containing an abrupt change or discontinuity (Plewes 1977:2; Shiskin et al. 1965:5-6). For an overview of potential problems for X-11 users, see Fromm (1978).

Extremes

Although the X-11 program is not appropriate for highly irregular series, it is good to use when the series contains extreme values. It is "resistant" to the effect of extremes (Pierce, 1980:131), because it contains a graduated weighting system.³¹ Values exceeding 2.5 standard deviations are weighted zero, and values from 1.5 to 2.5 standard deviations are graduated linearly from full to zero weight. This is the default option, which the user is allowed to modify.

Series Length

The reason for the limit on series length becomes obvious if you consider that the X-11 algorithm searches for similarities among months, and that there is only one January, one February, and so on, per year. To analyze a six-year series, for example, is to look at the similarities among six Januaries, six Februaries, and so on. Thus, the number of observations is really only six.

Discontinuities

If there is an abrupt change or discontinuity in the series, no continuous method, component or stochastic, will work. The moving average is a linear smoothing technique, and all smoothing techniques are analytically continuous, defined in the same way throughout the series.³² Like other methods of continuous smoothing, the X-11 cannot accurately describe discontinuities or abrupt changes in the direction of a series. When discontinuities are suspected, Shiskin et al. (1967:5) suggests that they be "ascertained by inspection," and that the series then be broken into segments for analysis. The user should investigate the data source to determine whether there was a change in definition or data collection practices.

³¹For other resistant time series analysis methods, see Tukey (1977) or Velleman and Hoaglin (1981). For resistant seasonal adjustment with SABL, see Cleveland et al. (1978).

³²Macaulay (1931:21), in his classic text on time series smoothing, argues that, even though freehand smoothing with a French curve is generally unsatisfactory, "if the underlying ideal curve is itself not smooth," then a freehand method is better than mathematical curve-fitting. If there are discontinuities or sharp changes of direction (cusps) in the underlying series, then any overall, continuous smoothing method will obscure them, rather than describe them accurately. An analyst might then be misled into thinking that there was no abrupt change in the series. For a fuller discussion of this issue, see the Statistical Analysis Center publication, "Manual for the Pattern Description of Time Series," pages 7-8 and 58-59.

Moving Seasonality

If, instead of an abrupt change or discontinuity, the seasonal fluctuation gradually changes over the years, stochastic methods may be more appropriate than component methods. The X-11 program assumes that any seasonal fluctuation follows a consistent pattern from year to year (see "Pattern Consistency," above). When, over a period of years, the seasonal fluctuation gradually increases in strength in certain months and decreases in strength in other months, the series contains "moving seasonality." One of the X-11 diagnostics, an F of moving seasonality, will alert you to its presence. In contrast to the F of stable seasonality, which is the ratio of the between-month variance of the seasonal to the irregular, the F of moving seasonality is the between-year ratio. It tests the null hypothesis that the years all have the same seasonal pattern.

When X-11 results indicate a significant F of moving seasonality, we suggest the following procedure:

1. Inspect the series for abrupt changes or discontinuities. Is there an abrupt change in level? Does the series suddenly develop (or lose) seasonal fluctuation after a certain date? If so, no continuous method, whether component or seasonal, is appropriate. Check the definition and validity of the data set. If the definition of the series changed at some point, partition the series into two parts at that point, and analyze the parts separately.
2. If there is no discontinuity, compare the additive to the multiplicative adjustment. Do both contain moving seasonality? If not, assume that the adjustment that does not reflects the true nature of the series.
3. If both additive and multiplicative adjustments indicate moving seasonality, determine the particular month(s) that vary in seasonal fluctuation. Using options available in the X-11 program, change the moving average for these months. (For more detail, see Plewes 1977:5-6.)
4. In any case, do not accept the results of an adjustment in which the moving seasonality F value is significant.

The distinction between gradual change (moving seasonality) and abrupt change (discontinuity) requires subjective interpretation and an intimate knowledge of the data source. As an example, figure 11 shows a series containing an apparent discontinuity. In this series, the number of people in Illinois receiving the first 26 weeks of unemployment insurance, moving seasonality F values are significant in both the additive adjustment and the multiplicative adjustment.³³ Between October, 1974 and January,

³³The F values are 6.16 and 7.09, respectively. A value of 2.41 or higher should be considered significant. That is, the possibility that the series contains moving seasonality should not be ruled out. We cannot confidently assume that the seasonal pattern is the same in every year of the series.

Figure 11

ILLINOIS UNEMPLOYMENT INSURANCE RECIPIENTS, 1965-1978

FIRST 26 WEEKS
SOURCE: BUREAU OF LABOR STATISTICS

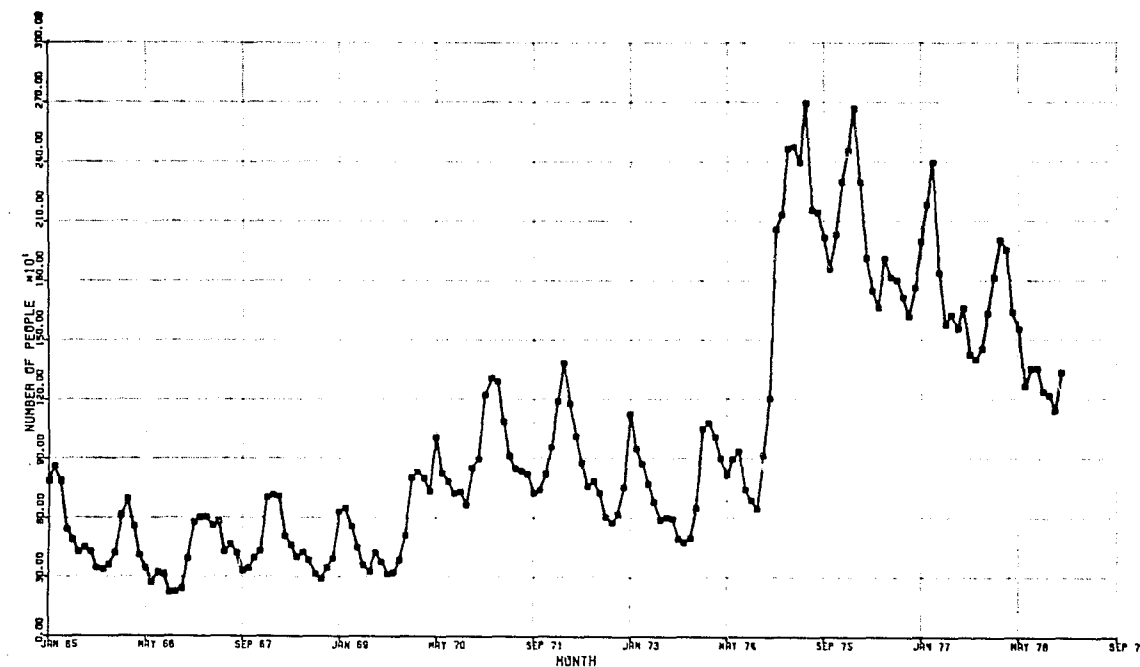


Table 4

Seasonal Factors, Additive Adjustment Illinois Unemployment Insurance Recipients

Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
1965	181	259	174	48	-17	-76	-39	-65	-150	-150	-116	-52
1966	184	256	176	49	-17	-75	-37	-67	-150	-153	-117	-51
1967	189	249	177	52	-15	-70	-35	-71	-153	-164	-117	-52
1968	199	247	179	63	-15	-65	-40	-79	-161	-177	-117	-50
1969	210	248	183	74	-14	-61	-47	-88	-172	-196	-123	-47
1970	230	255	197	88	-10	-62	-54	-96	-187	-213	-139	-45
1971	251	265	208	100	-3	-54	-52	-105	-204	-243	-165	-54
1972	274	277	229	121	8	-45	-46	-118	-227	-274	-193	-58
1973	291	292	260	128	11	-40	-31	-131	-238	-299	-221	-65
1974	300	310	292	132	2	-42	-26	-130	-243	-309	-238	-61
1975	290	332	318	120	-17	-47	-28	-123	-235	-303	-246	-60
1976	271	352	341	112	-34	-58	-46	-112	-228	-287	-241	-44
1977	250	366	352	94	-51	-75	-63	-101	-215	-261	-237	-35
1978	240	370	353	83	-63	-88	-75	-91	-205	-243	-235	-32

1975, the number of unemployed people in Illinois tripled. (For a discussion of this phenomenon, see Block et al. 1981.) The seasonal factors (table 4) reflect this drastic change. The seasonal factor of many months changes about 1974. April, for example, is usually a little high, but is very high in 1974. August follows the opposite pattern. October is always low, but 1974 is extremely low. These results should not be taken at face value, but should be considered to be an indication of problems in the adjustment. Since there was an abrupt change in 1974, the series should be split into two segments, and each segment should be analyzed separately.

Stochastic Methods

Stochastic time series models are a sophisticated way of using past observations of a series to forecast its future observations. Box and Jenkins (1970) suggested that most time series encountered in practice follow either (or a combination of) two types of stochastic process: moving average and autoregressive. By determining what process a series followed in the past, and assuming that that process will continue, we can forecast the future.

Seasonal fluctuation is one possible aspect of a stochastic process. To identify a model for a series, the analyst must decide whether or not a seasonal process should be part of the model. Just as the component method did not lend itself to one simple, objective interpretation of X-11 results to decide whether or not a series fluctuates with the seasons, the stochastic method also relies on the subjective interpretation of a number of diagnostic tests. In this section, we explain the stochastic method, and discuss the most important of these diagnostics.

Moving Average and Autoregressive Processes

In a moving average (MA) process, the current observation is a function of a past error. Error is a random disturbance, sometimes called "noise" or "shock." By definition, the error of one observation is independent of the error of other observations. However, errors can be correlated with the observations themselves. This happens in a moving average process.

An MA(1) moving average process means that the current observation is affected by the error of the previous observation.³⁴ An MA(2) moving average process means that the current observation is affected by the error of the second previous observation.

³⁴The word "affected" here simply means correlated. Although the current observation may be predicted from past data in the series, the past data do not directly "cause" the current observation.

A seasonal moving average process, MA(12), means that the current observation is affected by the error of the observation one year ago. In general, in a series following a moving average process, the current observation is correlated with past error(s).

In an autoregressive (AR) process, the current observation is a function of a past observation (not a past error). An AR(1) autoregressive process means that the current observation is affected by the previous observation. An AR(2) autoregressive process means that the current observation is affected by the second previous observation. A seasonal autoregressive process, AR(12), means that the current observation is affected by the observation one year ago.

The stochastic process of most series can be described as either MA(1), MA(2), AR(1), AR(2), or a combination of MA and AR processes. Some series are a combination of a serial MA or AR process (or both), and a seasonal MA or AR process (or both). How can we identify the process, or combination of processes, that best describes the series at hand?

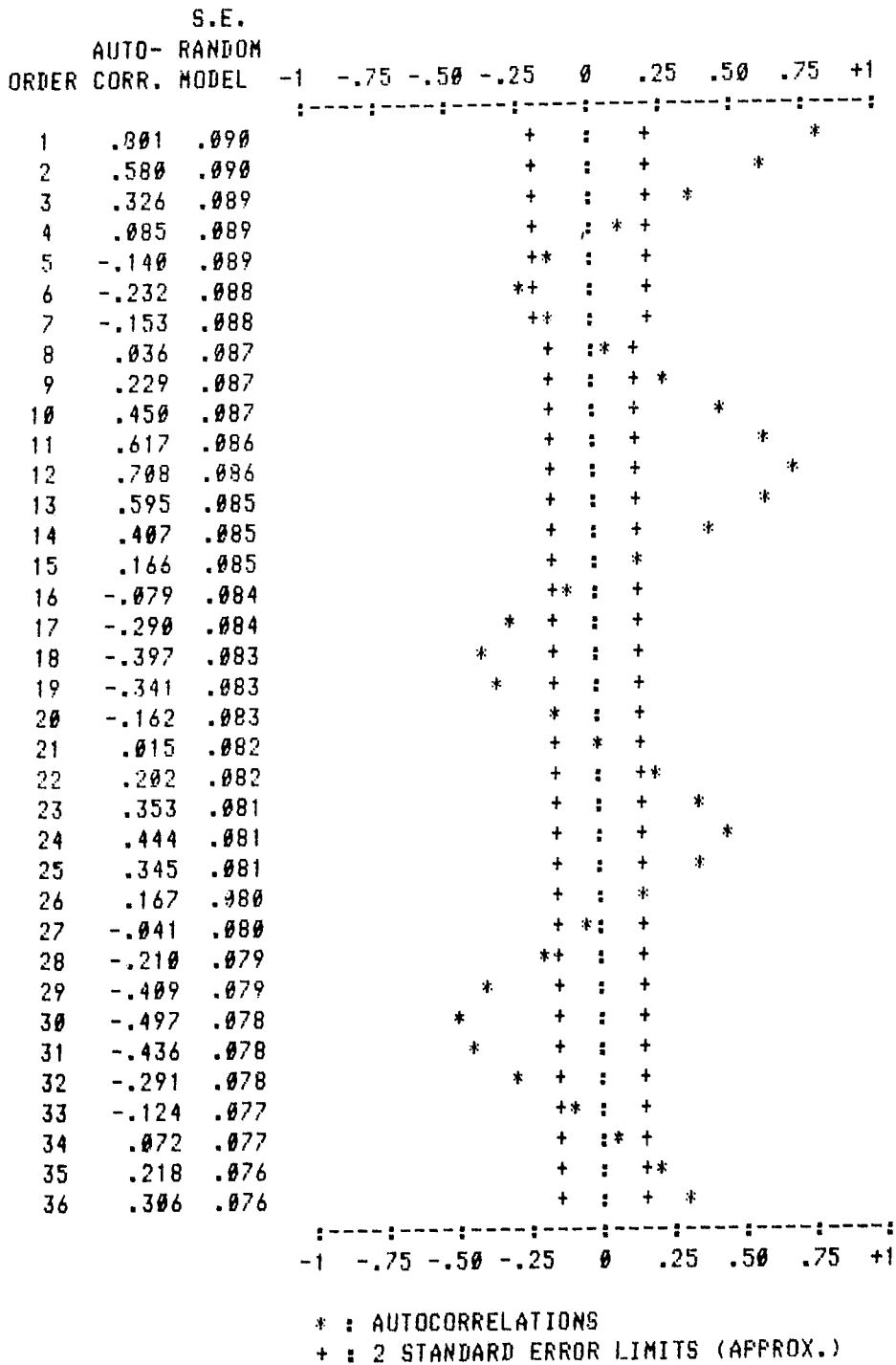
Identifying the Process of a Series

There is no way to measure past error directly.³⁵ How, then, can we differentiate a moving average process from an autoregressive process? An important diagnostic for identifying the process of a series is the correlogram. Moving average and autoregressive processes produce different "autocorrelation" patterns. Autocorrelation refers to the correlation between the observations of a time series. A correlation between each observation and the neighboring observation is a first-order autocorrelation, or an autocorrelation at lag 1. A correlation between each observation and the observation two months away is a second-order autocorrelation, or an autocorrelation at lag 2. A correlogram is a chart of the autocorrelations of a series at various lags. The correlogram in figure 12 shows each autocorrelation from lag 1 to lag 36 for Illinois larceny/theft. The first-order autocorrelation is .801, which means that observations in this series tend to be closely related to their neighboring observations. If one observation is high, the observations before and after it are likely to be high, and vice versa.

In a moving average process, as discussed above, observations are correlated with one or more previous error(s). Although we cannot observe correlation with an error directly, a correlation with a previous error results in a correlation with the corresponding previous observation. For example, in an MA(12)

³⁵Error may be estimated by using the residuals of a regression (the difference between an observation and the equilibrium or mean level of the series). Roberts (1974,1975) calls this the Durbin/ARIMA method.

Figure 12
Correlogram of Illinois Larceny/Theft



series, the current observation is correlated with the twelfth previous observation. This is also true of an autoregressive process. However, because errors are independent of each other by definition, the second or greater previous observations in an MA(1) series, or the third or greater previous observations in an MA(2) series, or the twenty-fourth or greater previous observations in an MA(12) series are not correlated with the current observation. This is not true of an autoregressive process.

In an autoregressive process, neighboring observations are correlated with each other. For example, in an AR(1) process, observation 1 is correlated with observation 2, and observation 2 is correlated with observation 3. Therefore, observation 1 and observation 3 are correlated. The correlation of observations one time period apart produces geometrically decreasing correlations of observations two time periods apart, three periods apart, and so on.

Because the second or greater previous observation is not correlated with the present observation in an MA(1) series, but is correlated with the present observation in an AR(1) series, autocorrelations provide a useful clue as to what stochastic model would best describe a series. A high autocorrelation at lag 1 that disappears at higher lags (for example, see figure 13) suggests an MA(1) model. A high autocorrelation at lag 1 that decreases exponentially at higher lags (for example, see figure 14) suggests an AR(1) model.

We distinguish between seasonal MA and AR processes in a similar way. In both kinds of series, observations twelve months apart are correlated with each other. That is, the January observations are similar to each other, the February observations are similar to each other, and so on. Therefore, both seasonal MA and seasonal AR processes have significant twelfth-order autocorrelations. However, in a seasonal MA series, the 24th-order and 36th-order autocorrelations are small, while in a seasonal AR series, they are significant. A high autocorrelation at lag 12 that is still high but decreasing exponentially at lags 24, 36, and so on suggests a seasonal autoregressive model. A high autocorrelation at lag 12 that disappears at higher seasonal lags suggests a seasonal moving average model.

Stationarity

In general, seasonal series have a significant autocorrelation at lag 12. However, the opposite is not always true. Some series that are not seasonal may have a large correlation between observations twelve months apart. This can happen if there is an overall trend in the series. For example, figure 15 is the correlogram of a nonseasonal series with a decided increase over time. Observations twelve months apart are correlated. Figure 16 is the correlogram of the same series with the increasing trend removed. Observations twelve months apart are not correlated.

Figure 13

Correlogram, First Difference
Chicago Homicide with a Gun: 1965-1978

			S.E.										
			AUTO- RANDOM										
ORDER	CORR.	MODEL	-1	-.75	-.50	-.25	0	.25	.50	.75	+1		
:-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:													
1	-.412	.077				*	+	:	+				
2	.000	.076					+	*	+				
3	-.078	.076					++	:	+				
4	.050	.076					+	:	+				
5	-.088	.076					++	:	+				
6	.081	.076					+	:	++				
7	-.011	.075					+	*	+				
8	.011	.075					+	*	+				
9	-.043	.075					+	++	:	+			
10	-.046	.075					+	++	:	+			
11	.021	.074					+	*	+				
12	.072	.074					+	:	++	:			
13	-.059	.074					+	++	:	+			
14	.017	.074					+	*	+				
15	-.057	.073					+	++	:	+			
16	.095	.073					+	:	++	:			
17	-.035	.073					+	++	:	+			
18	-.028	.073					+	++	:	+			
19	.055	.072					+	:	++	:	+		
20	.000	.072					+	*	+				
21	-.028	.072					+	++	:	+			
22	.037	.072					+	:	++	:	+		
23	-.011	.071					+	*	+				
24	.022	.071					+	*	+				
25	-.092	.071					++	:	+				
26	.100	.071					+	:	++	:			
27	-.012	.070					+	*	+				
28	-.005	.070					+	*	+				
29	-.130	.070					*	:	+				
30	.100	.070					+	:	++	:			
31	-.059	.069					+	++	:	+			
32	.015	.069					+	*	+				
33	-.046	.069					+	++	:	+			
34	.161	.069					+	:	*				
35	-.161	.068					*	:	+				
36	.153	.068					+	:	*				
:-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:													
			-1	-.75	-.50	-.25	0	.25	.50	.75	+1		

* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

Figure 14

Correlogram, Chicago Assault Homicide: 1965-1978

ORDER	S.E.												
	AUTO-	RANDOM	-1	-.75	-.50	-.25	0	.25	.50	.75	+1		
1	.384	.076					+	:	+		*		
2	.279	.076					+	:	+		*		
3	.203	.076					+	:	++				
4	.238	.076					+	:	+	*			
5	.094	.076					+	:	++				
6	.171	.075					+	:	*				
7	.149	.075					+	:	*				
8	.205	.075					+	:	++				
9	.115	.075					+	:	++				
10	.195	.074					+	:	++				
11	.153	.074					+	:	*				
12	.177	.074					+	:	+				
13	.184	.074					+	:	+				
14	.196	.073					+	:	++				
15	.074	.073					+	:	+				
16	.091	.073					+	:	++				
17	.109	.073					+	:	++				
18	.058	.072					+	:	+				
19	.106	.072					+	:	++				
20	.174	.072					+	:	*				
21	.145	.072					+	:	*				
22	.158	.071					+	:	*				
23	.150	.071					+	:	*				
24	.144	.071					+	:	*				
25	.136	.071					+	:	*				
26	.092	.071					+	:	++				
27	.075	.070					+	:	++				
28	-.057	.070					+	:	+				
29	-.122	.070					++	:	+				
30	-.033	.070					+	:	+				
31	.009	.069					+	:	+				
32	.025	.069					+	:	++				
33	.069	.069					+	:	++				
34	.120	.068					+	:	++				
35	.044	.068					+	:	+				
36	.045	.068					+	:	++				

* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

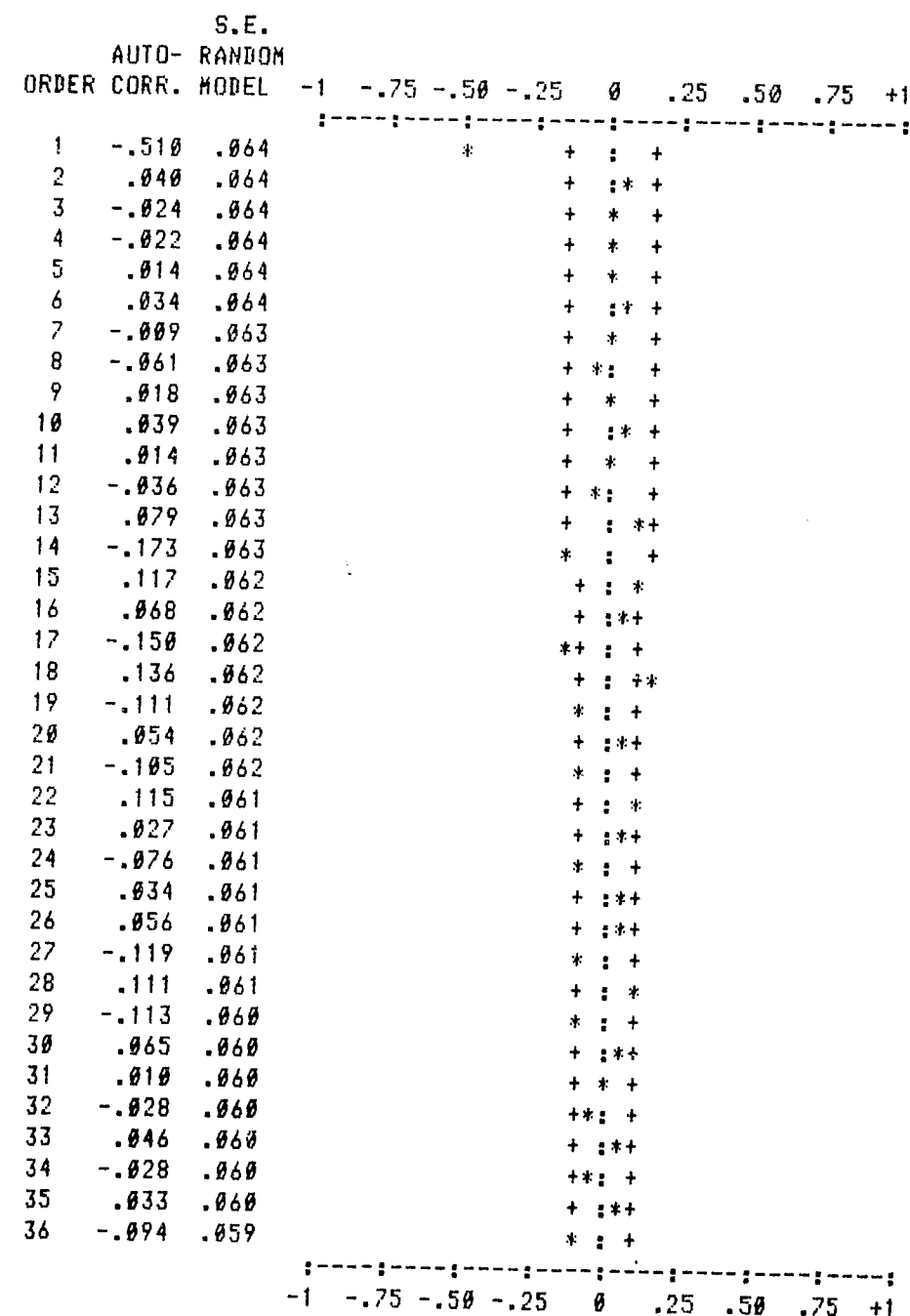
Figure 15

Correlogram, Canadian Homicide: 1961-1980

ORDER	S.E.												
	AUTO-	RANDOM	-1	-.75	-.50	-.25	0	.25	.50	.75	+1		
1	.759	.064					+	:	+		*		
2	.760	.064					+	:	+		*		
3	.743	.064					+	:	+		*		
4	.737	.064					+	:	+		*		
5	.741	.064					+	:	+		*		
6	.738	.063					+	:	+		*		
7	.720	.063					+	:	+		*		
8	.706	.063					+	:	+		*		
9	.720	.063					+	:	+		*		
10	.728	.063					+	:	+		*		
11	.716	.063					+	:	+		*		
12	.692	.063					+	:	+		*		
13	.689	.063					+	:	+		*		
14	.648	.062					+	:	+		*		
15	.687	.062					+	:	+		*		
16	.673	.062					+	:	+		*		
17	.625	.062					+	:	+		*		
18	.654	.062					+	:	+		*		
19	.613	.062					+	:	+		*		
20	.629	.062					+	:	+		*		
21	.617	.061					+	:	+		*		
22	.655	.061					+	:	+		*		
23	.638	.061					+	:	+		*		
24	.609	.061					+	:	+		*		
25	.613	.061					+	:	+		*		
26	.605	.061					+	:	+		*		
27	.563	.061					+	:	+		*		
28	.581	.060					+	:	+		*		
29	.550	.060					+	:	+		*		
30	.570	.060					+	:	+		*		
31	.560	.060					+	:	+		*		
32	.543	.060					+	:	+		*		
33	.538	.060					+	:	+		*		
34	.516	.060					+	:	+		*		
35	.502	.059					+	:	+		*		
36	.474	.059					+	:	+		*		

* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

Figure 16
Correlogram, First Difference
Canadian Homicide: 1961-1980



* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

This emphasizes an additional complication of identifying the stochastic processes of a series: the method we have described for identifying a model only works for stationary series. A series is stationary if its mean and its variance are the same at every part of the series. A stationary series thus shows no trend. Because most series of crime data do show some trend, they may not, therefore, be analyzed by stochastic time series methods unless they are first transformed to remove the trend. First, the series is transformed to make it stationary. Second, a model is identified for the transformed series.

Just as there are seasonal MA and AR processes, there can be seasonal lack of stationarity. In such a case, each month is systematically higher (or lower) than the same month one year ago.³⁶ In addition, just as it is possible to have a combination of serial and seasonal processes in the same series, it is possible to have a combination of serial and seasonal lack of stationarity. How do we decide whether or not a series is stationary, and if we decide it is not, how do we transform it?

To decide whether or not a series is stationary, first look at a graph of the series.³⁷ Does the level of the series seem to increase or decrease over time? Second, look at a correlogram. In a series with trend, the correlogram shows a pattern of high autocorrelations that do not decrease with lag. In contrast, in a series that is stationary, but follows an autoregressive process, the autocorrelations decrease geometrically (see above section). Similarly, in a series with seasonal trend, the autocorrelation at the first seasonal lag is high, and the autocorrelations at successive seasonal lags do not decrease. For example, figure 17 is the correlogram of the Illinois larceny/theft series with serial trend removed (by first differencing; see below). The autocorrelations are .567 at lag 12, .520 at lag 24, and .416 at lag 36. The seasonal autocorrelations do decrease a little with lag, but the decrease is certainly less than geometrical. This suggests seasonal lack of trend instead of an AR(12) process.

³⁶Another cause of lack of stationarity is a change in the variance from the beginning to the end of the series. We do not discuss this kind of seasonal lack of stationarity, because it is difficult to imagine a seasonal change of variance. However, with serial change in variance, transforming the series with a log or a square root may produce a stationary series. For more detail, see Nelson (1976), McCleary and Hay (1981), or the other stochastic time series analysis sources listed in the bibliography.

³⁷It is easier to see a trend in a graph of a standardized series than in a graph of the raw data. In a standardized series, each observation is converted to its Z score, or its standard deviation above or below the mean. This useful option is available in the IDA package (Ling and Roberts, 1982).

Figure 17

Correlogram, First Difference
Illinois Index Larceny/Theft: 1972-1981

S.E.													
AUTO- RANDOM			-1	-.75	-.50	-.25	0	.25	.50	.75	+1		
ORDER	CORR.	MODEL	-----										
1	.065	.091					+	+	+				
2	.118	.090					+	+	+				
3	-.033	.090					+	+	+				
4	-.003	.089					+	+	+				
5	-.356	.089				*	+	+	+				
6	-.472	.089			*		+	+	+				
7	-.308	.088				+	+	+	+				
8	-.053	.088					+	+	+				
9	-.089	.087					+	+	+				
10	.136	.087						+	+				
11	.193	.087					+	+	+				
12	.567	.086					+	+	+		*		
13	.204	.086					+	+	+				
14	.182	.085					+	+	+				
15	.006	.085					+	+	+				
16	-.070	.085					+	+	+				
17	-.258	.084				+	+	+	+				
18	-.430	.084			*		+	+	+				
19	-.340	.083				*	+	+	+				
20	-.044	.083					+	+	+				
21	-.020	.082					+	+	+				
22	.073	.082					+	+	+				
23	.142	.082					+	+	+				
24	.520	.081					+	+	+		*		
25	.217	.081					+	+	+				
26	.127	.080					+	+	+				
27	-.094	.080					+	+	+				
28	.109	.079					+	+	+				
29	-.279	.079				+	+	+	+				
30	-.385	.079			*		+	+	+				
31	-.260	.078				+	+	+	+				
32	-.048	.078					+	+	+				
33	-.102	.077					+	+	+				
34	.104	.077					+	+	+				
35	.152	.076					+	+	+				
36	.416	.076					+	+	+		*		

			-1	-.75	-.50	-.25	0	.25	.50	.75	+1		

* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

Differencing is a transformation intended to produce a stationary series. An overall trend can usually be removed by a first difference; a seasonal trend can usually be removed by a twelfth difference. In a first difference, each observation is subtracted from the following observation.³⁸ In a twelfth difference, each observation is subtracted from the observation twelve months away. The differenced series is interpreted as the change from one observation to the next for a first difference, or the change from one year to the next for a twelfth difference. If a series has both a serial and a seasonal trend, you would transform it into a stationary series by taking a twelfth difference of the first difference.

For example, figure 17 is the correlogram of Illinois larceny/theft after first differencing. Figure 18 is the correlogram of the series after first and twelfth differencing. In other words, each observation was subtracted from the following observation, which produced a series of first differences. These first difference values were then subtracted from the first difference value twelve months away. Compare these correlograms to the correlogram of the original series, figure 12 above.

A drawback of differencing is that the differenced series has fewer observations than the original series. If the original series has 144 observations, for example, a first difference has 143. Even more observations are lost with twelfth differencing. However, if your series is not stationary, analyzing a difference transformation of it may be the only alternative if you want to identify a stochastic process.

Model Evaluation

After obtaining a stationary transformation, or determining that the original series is already stationary, the next step is to estimate the stochastic process or processes that best describe the stationary series. In the Illinois larceny/theft example, the first and twelfth-order autocorrelations are significant, but the second and 24th-order autocorrelations do not differ significantly from zero. This pattern of autocorrelations suggests a combination of MA(1) and MA(12) processes. Therefore, to model larceny/theft, we applied MA(1) and MA(12) processes to the differenced series. This produced the modelled series graphed in figure 10 (above). How do we decide whether or not this model accurately describes the stochastic process of the series?

³⁸Some curves require two successive first differences to make them stationary. That is, each observation is subtracted from the next observation. This produces a series of first differences, which will be a straight line with a trend. Then this differenced series is differenced again. The second differencing produces a stationary series.

Figure 18

Correlogram, First and Twelfth Difference
Illinois Index Larceny/Theft: 1972-1981

ORDER	CORR.	MODEL	S.E.									
			AUTO-	RANDOM								
				-1	-.75	-.50	-.25	0	.25	.50	.75	+1
1	-.388	.095					*	+	:	+		
2	-.130	.095						++	:	+		
3	.015	.094						+	+	+		
4	.178	.094						+	:	+		
5	-.133	.094						++	:	+		
6	.009	.093						+	+	+		
7	.080	.093						+	:	+	+	
8	-.052	.092						+	+	+		
9	-.133	.092						++	:	+		
10	.101	.091						+	:	+		
11	.243	.091						+	:	++		
12	-.499	.090						+	:	+		
13	.128	.090						+	:	++		
14	.130	.089						+	:	++		
15	.180	.089						+	:	+		
16	-.314	.088						+	:	+		
17	.144	.088						+	:	++		
18	.015	.087						+	+	+		
19	-.054	.087						+	+	+		
20	-.026	.086						+	+	+		
21	.123	.086						+	:	++		
22	-.051	.085						+	+	+		
23	-.089	.085						++	:	+		
24	.052	.084						+	+	+		
25	.076	.084						+	:	++		
26	-.068	.083						+	+	+		
27	-.191	.083						++	:	+		
28	.209	.082						+	:	++		
29	-.004	.082						+	+	+		
30	-.055	.081						+	+	+		
31	-.003	.081						+	+	+		
32	.091	.080						+	:	++		
33	-.122	.080						++	:	+		
34	.010	.079						+	+	+		
35	.049	.079						+	+	+		
36	.026	.078						+	+	+		

* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

Correlogram of Residuals

One way to evaluate a stochastic model is to analyze the residuals, the discrepancy between the modelled values and the actual series. Residuals of a good model vary randomly over time.³⁹ The correlogram of the residuals of an MA(1) and MA(12) model for Illinois larceny/theft (figure 19) does indeed indicate such a random pattern.⁴⁰ Compare this pattern to the correlograms of the original series (figure 12), the series transformed by first differencing (figure 17), and the series transformed by both first and twelfth differencing (figure 18). Clearly, the residuals look most like a random series.

However, like component analysis, stochastic analysis is open to alternative interpretations. The Illinois larceny/theft series exemplifies a common situation requiring interpretation: is the series non-stationary, or is it an autoregressive process with a very high correlation between one observation and the next? The series transformed by first and twelfth differencing (figure 18) has negative autocorrelation at lag 12. One interpretation of this is that it suggests the model described above, a moving average process with a negative relation between observation and error. On the other hand, the differencing may have overadjusted the series, adding a systematic pattern that was not in the original series. A simpler twelfth difference without the first difference produces a transformed series that has the autocorrelations in figure 20. This pattern of autocorrelations suggests an AR(1) or AR(2) process.

Cumulative Periodogram of Residuals

Another diagnostic, the cumulative periodogram, is very useful in evaluating a tentative model, especially when the series may contain seasonal fluctuation. The cumulative periodogram is based on the assumption that a series is made up of sine and cosine waves. The analysis of the period, phase, and amplitude of these waves is known as examining the series in the "frequency domain," in contrast to the "time domain," which is the kind of analysis we have discussed so far in this report. Period is the time required for a full cycle. Frequency is the number of cycles per time unit. Because frequency is the reciprocal of period, the meaning of "high frequency" and "low periodicity" are the same, and "power domain" means the same thing as "frequency domain." Phase is the position of the cosine function relative to the starting point of the series. The measure of amplitude, or power over the frequency domain, is the spectrum, or "power spectrum." (See Rosenblatt 1965:1-2 for more detail.) A periodogram measures the intensity of the spectrum at a certain

³⁹This random variation in a time series is sometimes called "white noise."

⁴⁰Two of the 36 autocorrelations in figure 19 are slightly outside of the two standard deviation limit, but a small percentage of significant autocorrelation is expected by chance.

Figure 19

Correlogram, Residuals of (0,1,1) (0,1,1) Model
Illinois Index Larceny/Theft: 1972-1981

ORDER	S.E.											
	AUTO-	RANDOM	-1	-.75	-.50	-.25	0	.25	.50	.75	+1	
1	.000	.090					+	*	+			
2	-.037	.090					+	*	+			
3	-.024	.089					+	*	+			
4	.124	.089					+	*	+			
5	-.076	.089					+	*	+			
6	.022	.088					+	*	+			
7	-.047	.088					+	*	+			
8	-.027	.087					+	*	+			
9	-.170	.087					*	*	+			
10	-.030	.087					+	*	+			
11	.003	.086					+	*	+			
12	.006	.086					+	*	+			
13	.026	.085					+	*	+			
14	.151	.085					+	*	+			
15	.066	.085					+	*	+			
16	-.050	.084					+	*	+			
17	.036	.084					+	*	+			
18	.013	.083					+	*	+			
19	-.114	.083					+	*	+			
20	.028	.083					+	*	+			
21	-.025	.082					+	*	+			
22	-.106	.082					+	*	+			
23	-.122	.081					+	*	+			
24	-.010	.081					+	*	+			
25	.000	.081					+	*	+			
26	-.089	.080					+	*	+			
27	-.214	.080					+	*	+			
28	.195	.079					+	*	+			
29	-.004	.079					+	*	+			
30	-.022	.078					+	*	+			
31	-.031	.078					+	*	+			
32	.080	.078					+	*	+			
33	-.076	.077					+	*	+			
34	-.024	.077					+	*	+			
35	-.012	.076					+	*	+			
36	-.065	.076					+	*	+			

* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

Figure 20

Correlogram, Twelfth Difference
Illinois Index Larceny/Theft: 1972-1981

ORDER	S.E.											
	AUTO-	RANDOM	-1	-.75	-.50	-.25	0	.25	.50	.75	+1	
1	.603	.095					+	*	+			
2	.509	.094					+	*	+			
3	.524	.094					+	*	+			
4	.519	.094					+	*	+			
5	.381	.093					+	*	+			
6	.343	.093					+	*	+			
7	.295	.092					+	*	+			
8	.177	.092					+	*	+			
9	.113	.091					+	*	+			
10	.158	.091					+	*	+			
11	.122	.090					+	*	+			
12	-.102	.090					+	*	+			
13	.069	.089					+	*	+			
14	.135	.089					+	*	+			
15	.096	.088					+	*	+			
16	-.080	.088					+	*	+			
17	-.004	.088					+	*	+			
18	-.043	.087					+	*	+			
19	-.094	.087					+	*	+			
20	-.066	.086					+	*	+			
21	-.063	.086					+	*	+			
22	-.150	.085					+	*	+			
23	-.193	.085					+	*	+			
24	-.148	.084					+	*	+			
25	-.155	.084					+	*	+			
26	-.230	.083					+	*	+			
27	-.250	.083					+	*	+			
28	-.120	.082					+	*	+			
29	-.163	.082					+	*	+			
30	-.190	.081					+	*	+			
31	-.177	.081					+	*	+			
32	-.198	.080					+	*	+			
33	-.254	.079					+	*	+			
34	-.214	.079					+	*	+			
35	-.187	.078					+	*	+			
36	-.213	.078					+	*	+			

* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

frequency, and the "normalized cumulative periodogram" (Box and Jenkins 1976:295) is a good tool for detecting periodic patterns in the residuals of a model.

For example, figure 21 shows two cumulative periodogram graphs side by side for comparison. The first, a graph for the original larceny/theft data, indicates a distinct departure from linearity at about a twelve-month period. The graph of the residuals of the MA(1) MA(12) model, on the other hand, do not indicate any significant periodicity.

A cumulative periodogram gives you the same sort of information that a correlogram gives you, but from a different perspective. The spectrum is mathematically equivalent to the autocorrelation function (Box and Jenkins 1976:39-45). It is simply an alternative way of describing the pattern of relationships among the observations. However, Box and Jenkins (1976:294) recommend it over the correlogram in evaluating departures from randomness in the residuals of a model. When we fit a model to a series containing seasonal fluctuation, we want to be sure that the model accounts for all of the seasonality. We do not want the residuals of the model to contain periodicity. As Box and Jenkins point out (1976:294):

Therefore, we are on the lookout for periodicities in the residuals. The autocorrelation function will not be a sensitive indicator of such departures from randomness, because periodic effects will typically dilute themselves among several autocorrelations. The periodogram, on the other hand, is specifically designed for the detection of periodic patterns in a background of white noise.

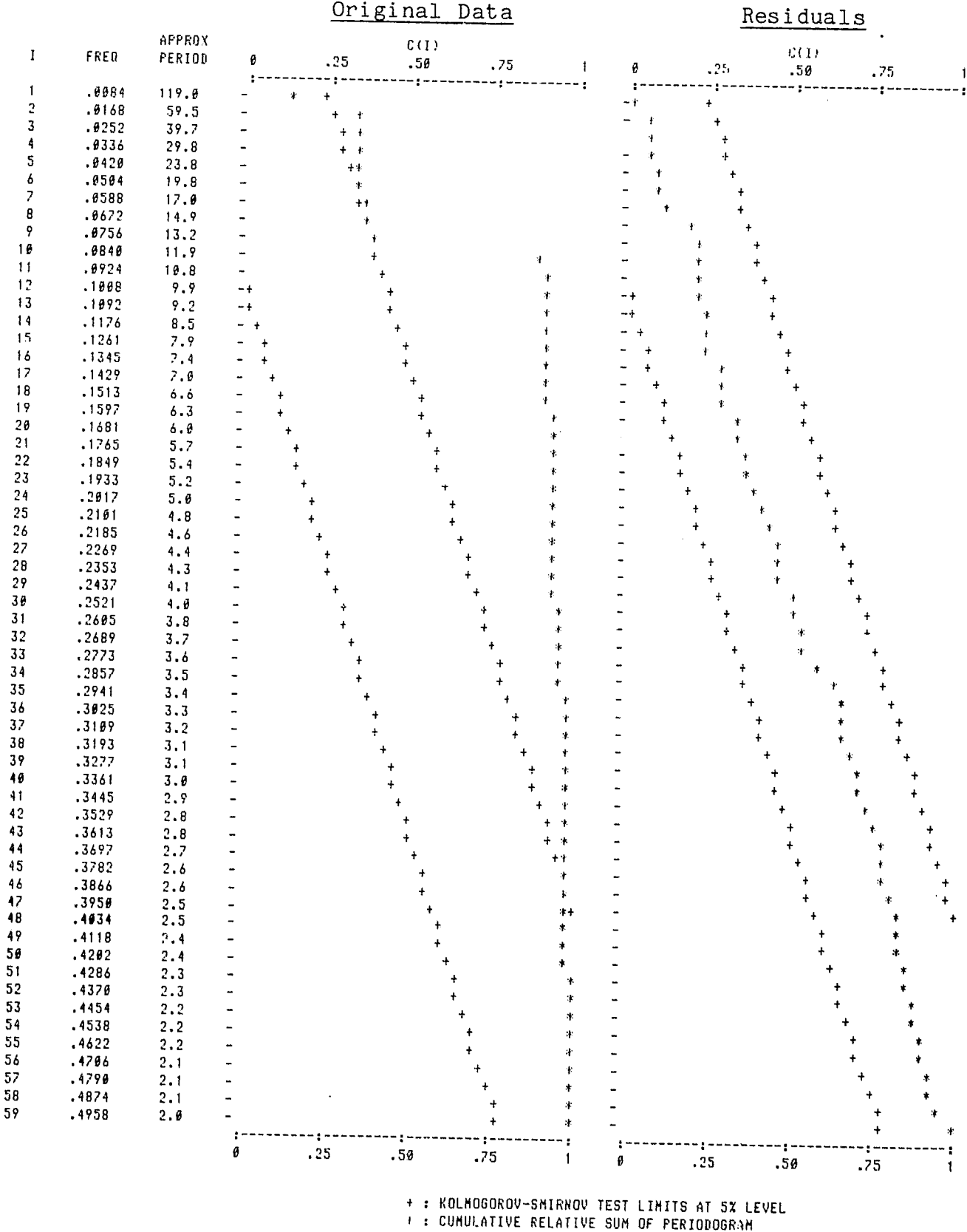
Appropriate Applications

Neither stochastic nor component methods are appropriate for highly irregular series, short series (six or fewer years), or series containing an abrupt change or discontinuity. A general rule of thumb in stochastic time series analysis is that a minimum of 50 observations are necessary to estimate the stochastic process of a series (see Hartman et al.,1980). However, with seasonal stochastic processes, even more observations are necessary. Also, keep in mind that, if a twelfth difference is necessary to make the series stationary, twelve observations will be lost.

Discontinuities

A model of a stochastic process, like a component model, is analytically continuous. Analytic continuity means that the behavior of the series in one small region is the same as the behavior of the series everywhere (Cox,1971:36). A stochastic process describes the relationship of each observation to preceding observation(s). This relationship is the same throughout the series. If there is an abrupt change or discontinuity in the definition of the series, a single stochastic model is not appropriate. If you suspect that this is the case, first inspect the

Figure 21
Cumulative Periodograms
Original Data and Model Residuals
Illinois Index Larceny/Theft: 1972-1981



series carefully, and check the original data source for possible changes in definition or data collection practices. Based on your knowledge of the series, you may want to hypothesize that some intervention changed the behavior of the series after a certain date. Such an hypothesis can be tested (see Glass et al. 1975; Shine, 1980, 1982). Your final model may be complex, including a change in level or stochastic process after the occurrence of the hypothesized intervention. In any case, do not try to fit a continuous stochastic process to a series containing a discontinuity.

Extremes

Stochastic methods, in contrast to component methods, are not resistant to the effect of extremes (see Chernick et al. 1982). Therefore, they are not appropriate for series containing extreme values. However, stochastic methods can, of course, be used if the series is first transformed to remove or re-weight the extremes.

Moving Seasonality

On the other hand, stochastic methods are more appropriate than component methods for series containing moving seasonality. The stochastic process concept is based on the assumption that the current observation is more strongly related to recent observations than it is to observations in the distant past. The whole purpose of identifying an autoregressive or moving average process is to describe this decreasing relationship. Thus, the stochastic approach was developed to allow for gradual change over time in the seasonal pattern.

Summary

Obviously, the combinations of moving average processes and autoregressive processes, serial processes and seasonal processes, can become quite complicated. Identifying the stochastic processes that define a series is not entirely objective, nor is it simple for an analyst to state these subjective decisions in a published report. It is not uncommon for two statisticians using the same stochastic time series analysis methods to identify different models for the same series. As Pierce (1980:130) argues, "Theoretically incompatible models can produce results uncomfortably close to each other and uncomfortably far from the truth." Unlike the Census X-11 program, which can be used easily and quickly for a large number of series, and which has standard options and criteria that can be explicitly stated, stochastic methods require a lengthy analysis and re-analysis of each individual series. Therefore, they are most appropriate when one or two important series must be analyzed, and not as the standard method of analyzing all of an agency's data (see Kuiper, 1978: 59-60).

ANNOTATED BIBLIOGRAPHY

ANNOTATED BIBLIOGRAPHY

- Alt, Frank B., Stuart J. Deutsch and Jamie J. Goode
1977 Estimation for the multi-consequence intervention model. Proceedings of the Statistical Computing Section, American Statistical Association. Presents an algorithm for an interrupted time series experiment.
- Anderson, R.L.
1950 Tests of significance in time-series analysis. Pp. 352-355 in Statistical Inference in Dynamic Economic Models. Tjalling C. Koopmans (ed.) New York: John Wiley & Sons. Discusses the F of stable seasonality.
- Ascher, William
1978 Forecasting: An Appraisal for Policy-Makers and Planners. Baltimore: Johns Hopkins University Press.
- Beaton, Albert E. and John W. Tukey
1974 The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. Technometrics 16(May,2): 147-185.
- Bliss, C. I.
1958 Periodic regression in biology and climatology. Connecticut Agricultural Experimental Station Bulletin 615:1-56.
1970 Statistics in Biology. Vol. II. New York: McGraw-Hill. Two sources for periodic regression analysis (PRA). For an example of the use of PRA, see Warren, et.al.
- Block, Carolyn Rebecca
1976 Cross-sectional and longitudinal analysis of developmental data. Social Science Research 5:137-151.
1979 Descriptive Time Series Analysis for Criminal Justice Decision Makers: Local Illinois Robbery and Burglary. Chicago: Statistical Analysis Center, Illinois Law Enforcement Commission. An appendix reviews "methods of determining the presence of seasonality."
1980 Is Crime Seasonal? Working paper, Illinois Law Enforcement Commission, Statistical Analysis Center.
- Block, Carolyn Rebecca and Richard L. Block
1980 Patterns of Change in Chicago Homicide: The Twenties, The Sixties, and The Seventies. Illinois Law Enforcement Commission, Statistical Analysis Center.
- Block, Carolyn Rebecca, Craig McKie, and Louise S. Miller
1983 Patterns of change over time in Canadian and United States Homicide. Policy Perspectives, forthcoming.

Block, Carolyn Rebecca and Louise S. Miller
1982 Manual for the Pattern Description of Time Series. Chicago: Statistical Analysis Center, Illinois Law Enforcement Commission.

Block, Carolyn Rebecca, Louise S. Miller, Richard Block, Douglas Hudson
1981 Explaining patterns of change over time in Chicago homicides with a gun. Manuscript. Statistical Analysis Center Illinois Law Enforcement Commission.

Box, George E.P., and Gwilym M. Jenkins
1970 Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
The classic treatment of ARIMA.

Box, George E.P., Gwilym M. Jenkins and D.W. Bacon
1967 Models for forecasting seasonal and nonseasonal time series. Pp. 271-311 in B. Harris (ed.) Advanced Seminar on Spectral Analysis of Time Series. New York: John Wiley and Sons, Inc.

Buchin, Stanley I.
1982 Sales forecasting for the marketing function. Paper prepared for the Info/Manufacturing 82 Conference by Temple, Barker & Sloane, Inc., 33 Hayden Ave., Lexington, Massachusetts 02173. Journal of Forecasting, forthcoming.
Describes a forecasting competition, performed on 1000 time series by eight statisticians using 24 alternative methods. Forecasts using adjusted series were more accurate than forecasts using raw data, with all methods. Box-Jenkins forecast was almost never more accurate than a forecast equal to the actual (seasonally adjusted) observation last period. Also see Kendall (1976), Hibbs (1977), Willson (1973), Kuiper (1978), for other method comparisons.

Campbell, Donald T. and H. Laurence Ross
1968 The Connecticut crackdown on speeding. Law and Society Review 3:33-53.
One of the first time series experiments.

Campbell, Donald T. and Julian C. Stanley
1966 Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Co.
This is the classic reference on time series experiments.

Chernick, Michael R., Darryl J. Downing, and David H. Pike
1982 Detecting outliers in time series data. Journal of the American Statistical Association 77 (December, 380):743-747.

Cleveland, William S., Douglas M. Dunn and Irma J. Terpenning
1978 SABL: A resistant seasonal procedure. Graphical methods for interpretation and diagnosis. Pp. 201-241 in Zellner (ed.) 1978.
SABL is like the X-11 in its component approach, but differs in its treatment of extremes, its choices regarding multiplicative versus additive adjustment, and its graphical displays.

Cohen, Lawrence E., Marcus Felson and Kenneth C. Land
1980 Property crime rates in the United States: A macrodynamic analysis, 1947-77, with ex ante forecasts for the mid-1980s. American Journal of Sociology 86(1, July):90-118.
Also see Felson and Land (1977).

Cox, M. G.
1971 Curve fitting with piecewise polynomials. Journal of the Institute of Mathematics and its Applications 8(1):36-52.

Dagum, Estela Bee
1978 A Comparison and Assessment of Seasonal Adjustment Methods for Employment and Unemployment Statistics. Background paper No. 5, National Commission on Employment and Unemployment Statistics, Washington, D.C. 20006.
1980 The X-11-ARIMA Seasonal Adjustment Method. Seasonal Adjustment and Time Series Staff, Statistics Canada, Ottawa, K1A 0T6.
Dagum has done, or inspired others to do, much of the advanced work in seasonal adjustment today, including the X-11/ARIMA method.

Deutsch, Stuart Jay
1978 Stochastic models of crime rates. International Journal of Comparative and Applied Criminal Justice 2(2):127-151.
Builds stochastic models for seven index crimes in each of ten U.S. cities. Finds that robbery, burglary, aggravated assault, larceny, and motor vehicle theft are seasonal, but homicide and forcible rape are not.
1979 Lies, damn lies and statistics: A rejoinder to the comment by Hay and McCleary. Evaluation Quarterly 3(2, May):315-328.

Deutsch, Stuart Jay and Francis B. Alt
1977 The effect of Massachusetts' gun control law on gun-related crime in the city of Boston. Evaluation Quarterly 1(4, November):543-568.
Finds that assault with a gun and armed robbery are seasonal, but homicide is not. See Hay and McCleary (1979) for a criticism, and Deutsch (1979) for the rejoinder. Also see Pierce and Bowers (1979) for an analysis of the same data.

Deutsch, Stuart Jay and Lu Ann Sims

- 1978 An identification algorithm for dynamic intervention modeling with application to gun control. Series no. J-78-29, Georgia Institute of Technology, Atlanta 30332. Mimeographed.
See Alt, Deutsch and Goode (1977).

Dutta, M.

- 1975 Econometric Methods. Cincinnati: South-Western Publishing Co.
See Chapter 6 for an elementary discussion of analyzing seasonality by regressing dummy variables.

Edgerton, Julie, Linda Phelps, Karen Boley-Chang, Constance Osgood
1978 Ecology of Rape, Kansas City Metropolitan Area: Summary Report of the Rape Data Bank. Institute for Community Studies, University of Missouri, Kansas City. Report Prepared prepared for the Metropolitan Organization to Counter Sexual Assault.
"No definite seasonal pattern" in 1971 and 1975 rape offenses in Kansas City, Missouri, Kansas City, Kansas, and Independence, Missouri. The method used was simple inspection of two years of monthly data.

Felson, Marcus and Kenneth C. Land

- 1977 Social, demographic and economic interrelationships with Educational trends in the United States. Research in Population Economics: An Annual Compilation of Research, Vol. I. Julian Simon (ed.)
Example of time-inhomogenous analysis method. Also see listings under Land, Land and Felson, and Cohen et al.

Fromm, Gary

- 1978 Comment on "An Overview of the Objectives and Framework of Seasonal Adjustment," by Kallek. Pp.26-29 in Zellner (1978)

Glass, Gene V.

- 1968 Analysis of data on the Connecticut speeding crackdown as a time series quasi-experiment. Law and Society Review 3 (August):55-76.
See Stanley and Ross (1968.)
1971 Estimating the effects of intervention into a non-stationary time series. University of Colorado, Laboratory of Educational Research, Report No. 51.

Glass, Gene V., Victor L. Willson, and John M. Gottman

- 1975 Design and Analysis of Time-Series Experiments. Boulder: Colorado Associated University Press.
With Campbell and Stanley (1966,) this is the classic time series experiment literature. For time series intervention also see Shine (1980,1982), Tyron (1982).

Granger, Clive W.J.

- 1978 Seasonality: Causation, interpretation and implications. Pp. 33-46 in Zellner (1978).

Grether, D.M., and M. Nerlove

- 1970 Some properties of "optimal" seasonal adjustment. Econometrica 38(5,September):682-703.
Clearly written. For other discussions of criteria, see Lovell (1963), Willson (1973), Granger (1978).

Hannon, E. J.

- 1960 The estimation of seasonal variation. The Australian Journal of Statistics 2(1, April):1-15.
1963 The estimation of seasonal variation in economic time series. American Statistical Association Journal 58(March): 31-44.

Hartmann, D.P., J.M. Gottman, R.R. Jones, W. Gardner, A.E. Kazdin and R. Vaught

- 1980 Interrupted time-series analysis and its application to behavioral data. Journal of Applied Behavior Analysis 13: 543-559.
Review of literature on necessity of 50-100 observations for fitting a stochastic model. For a simplified intervention analysis for shorter series, see Tyron (1982).

Hauser, Robert M.

- 1978 Some exploratory methods for modeling mobility tables and other cross-classified data. University of Wisconsin-Madison: Center for Demography and Ecology.
Also see Land (1980), Felson and Land (1977).

Hay, Richard A., Jr. and Richard McCleary

- 1979 Box-Tiao time series models for impact assessment: A comment on the recent work of Deutsch and Alt. Evaluation Quarterly 3(2,May):277-314.
The two analyses disagree on the seasonality of the armed robbery series. Also see Deutsch's (1979) rejoinder.

Hibbs, Douglas A., Jr.

- 1974 Problems of statistical estimation and causal inference in time-series regression models. Pp. 252-308 in Sociological Methodology 1973-1974.
1977 On analyzing the effects of policy interventions: Box-Jenkins and Box-Tiao vs. structural equation models. Pp. 137-179 in Sociological Methodology 1977. David R. Heise (ed.) San Francisco: Jossey-Bass.
Also see Buchin (1982), Willson (1973).

Hickman, J.P. and J.G. Hilton

- 1971 Probability and Statistics. Scranton, Pa: Intext.
See Chapter 19 for an explanation of the ratio-to-moving-average method.

Hurwicz, Leonid

- 1950 Variable parameters in stochastic processes: trend and seasonality. Statistical Inference in Dynamic Economic Models Tjalling C. Koopmans (ed.) New York: John Wiley & Sons.

Kallek, Shirley
1978 An overview of the objectives and framework of seasonal adjustment. Pp. 3-25 in Zellner (1978).

Kendall, Sir Maurice
1976 Time-Series. Second edition. New York: Hafner Press.
This is an excellent introduction to time series analysis. Unlike most other beginning texts, it covers all methods: component, stochastic, etc. It includes an overview of problems relevant to all time series analysis, and discusses the application of various methods to solving these problems. It also describes a forecasting competition by Reid (see Buchin, 1982). Highly recommended as an initial text for someone new to time series analysis.

Kendall, M.G. and A. Stuart
1966 The Advanced Theory of Statistics. Vol.3. New York: Hafner Pub. Co., Inc.
Chapter 46 outlines seasonality and trend. Contains more technical detail than Kendall (1976).

Ku, Richard and Bradford Smith
1977 First Year Evaluation of the Illinois Urban High Crime Reduction Program: Final Report. Manuscript. Abt Associates, Inc., Cambridge, Massachusetts.
1978 Second Year Evaluation of the Illinois Urban High Crime Reduction Program: Final report. Manuscript. Abt Associates, Inc., Cambridge, Massachusetts.
Analysis of 1972 to mid-1978 residential burglary and robbery in Peoria, Champaign, and Joliet, Illinois. Models fitted by polynomial regression. Uses ratio-to-moving-average to adjust for seasonality, but does not address questions of whether the series contain seasonal fluctuation. No diagnostic results given.

Kuiper, John
1978 A survey and comparative analysis of various methods of seasonal adjustment. Pp. 59-76 in Zellner (1978).
For other method comparisons, see Buchin (1982), Kendall (1976).

Lamp, Rainer
1983 Jahreszeit und Kriminalitat. (Time of year and criminality) Paper presented at the International Congress on Criminology, Vienna. Max-Planck-Institut, Freiburg.

Land, Kenneth C.
1978 Modelling macro social change. Social Science Quantitative Laboratory, University of Illinois at Urbana-Champaign. Mimeographed.
1980 Modeling macro social change. Ch.8, pp. 219-278 in Sociological Methodology 1980.

Land, Kenneth C. and Marc Felson
1976 A general framework for building dynamic macro social indicator models: including an analysis of changes in crime rates and police expenditures. American Journal of Sociology 82:565-604.
Also see Cohen, et.al. (1980).

Leinhardt, Samuel and Stanley S. Wasserman
1979 Exploratory Data Analysis: An Introduction to Selected Methods. Sociological Methodology.
For another introduction to EDA, see Velleman and Hoaglin (1981).

Lester, David
1972 Why People Kill Themselves. Springfield, Illinois: Charles Thomas.
Contains a review of literature on seasonality of suicide. Also see Vigderhous (1978).

Leuthold, R.M., A.J.A. MacCormick, A. Schmitz and D.G. Watts
1970 Forecasting daily hog prices and quantities: A study of alternative forecasting techniques. Journal of the American Statistical Association 65(March):90-107.
Example of an econometric model with day of the week and season of the year as predictors. Uses Theil's (1966) inequality coefficient to measure the accuracy of prediction.

Ling, Robert F. and Harry V. Roberts
1979 Exploring Statistics with IDA. Clemson University and University of Chicago. Mimeographed.
1980 Users Manual for IDA. Palo Alto, California: The Scientific Press.
1982 IDA: A User's Guide to the IDA Interactive Data Analysis and Forecasting System. New York: Scientific Press and McGraw-Hill.
IDA is an easy-to-use, very "friendly" interactive package for time series analysis. It was originally developed by the University of Chicago Graduate School of Business, but is now supported by SPSS, Inc. The stochastic analysis examples in this report were done on IDA. Also see listings under Roberts.

Lovell, Michael C.
1963 Seasonal adjustment of economic time series and multiple regression analysis. American Statistical Association Journal 58(304, December):993-1010.
An excellent, clearly written review of criteria for seasonal adjustment methods. Also see Willson (1973), Kuiper (1978), Grether and Nerlove (1970), and Buchin (1982) for other critical reviews.

- Macaulay, Frederick R.
1931 The Smoothing of Time Series. New York: National Bureau of Economic Research.
An early, classic review of smoothing, including moving average. For detecting seasonality, see pp. 121-129.
- McCain, Leslie J. and Richard McCleary
1979 The statistical analysis of the simple interrupted time-series quasi-experiments. Pp. 233-293 in Quasi-experimentation: Design and Analysis Issues for Field Settings, by Thomas D. Cook and Donald T. Campbell. Chicago: Rand McNally.
A practical guide to stochastic seasonal analysis models, especially with respect to intervention analysis.
- McCleary, Richard, and Richard A. Hay, Jr.
1980 Applied Time Series Analysis for the Social Sciences. Beverly Hills: Sage Publications.
With Nelson (1973), this is an excellent introduction to stochastic time series analysis.
- Makridakis, Spyros and Steven C. Wheelwright
1978 Forecasting: Methods and Applications. Santa Barbara: John Wiley and Sons.
A basic forecasting textbook.
- Marshall, Clifford W.
1977a Application of Time Series Methodology to Crime Analysis. The Polytechnic Institute, 33 Jay St. Brooklyn, 11201. Law Enforcement Assistance Administration Grant #76-TA-99-0028
1977b The State Space Forecasting Technique Applied to Reported Crime Data. Supplement to 1977a, above.
Uses X-11 with crime data for Cincinnati, 1967-1974. Finds robbery and aggravated assault, but not burglary, to be seasonal. Rape has too much irregular variation to tell.
- Nelson, Charles R.
1973 Applied Time Series Analysis. San Francisco: Holden-Day, Inc.
With McCleary and Hay (1980), this is an excellent introduction to stochastic time series analysis.
- Nettheim, Nigel F.
1965 A Spectral Study of "Overadjustment" for Seasonality. U.S. Department of Commerce, Bureau of the Census. Working Paper No. 21.
- Pfeifer, Phillip E. and Stuart Jay Deutsch
1980 Identification and interpretation of first order space-time ARMA models. Technometrics 22(August,3):397-408.
An extension of stochastic methods into the spatial domain.

- Pierce, David A.
1980 A survey of recent developments in seasonal adjustment. The American Statistician 34(August,3):125-134.
This is a relatively simple review and update.
- Pierce, Glenn L. and William J. Bowers
1979 The impact of the Bartley-Fox gun law on crime in Massachusetts. Unpublished manuscript, Center for Applied Social Research, Northeastern University, Boston, 02115.
Found aggravated assault with and without a gun to be seasonal. Also see Deutsch and Alt (1977).
- Pittman, David J. and William Handy
1964 Patterns in criminal aggravated assault. Journal of Criminal Law, Criminology, and Police Science 55:462-470.
Random sample of 25% of aggravated assaults known to police in St. Louis, 1961. Found no seasonal pattern, no relation between indoor-outdoor location and season.
- Plewes, Tom
1977 Criteria for judging the accuracy of a seasonal adjustment. Technical paper, U.S. Department of Labor, Bureau of Labor Statistics, Washington, D.C.
See discussion in text, under "Component Methods."
- Plosser, Charles I.
1978 A time series analysis of seasonality in econometric models Pp. 365-397 in Zellner (1978).
States the argument for incorporating seasonal fluctuation into a model.
- President's Commission on Law Enforcement and the Administration of Justice
1967 The Challenge of Crime in a Free Society. U.S. Government Printing Office.
"Murder is a seasonal offense. Rates are generally higher in the summer, except for December, which is often the highest month and almost always 5 to 20 percent above the yearly average. In December 1963, following the assassination of President Kennedy, murders were below the yearly average by 4 percent, one of the few years in the history of the UCR that this occurred."(p.27). Also see Wolfgang (1966).
- Quetelet, Adolphe
1842 A Treatise on Man and the Development of his Faculties. English translation 1968. New York: Burt Franklin.
Quetelet, a Belgian statistician, states, "The seasons, in their course, exercise a very marked influence: thus, during summer, the greatest number of crimes against persons are committed and the fewest against property; the contrary takes place during the winter," (p. 90). Also see Sylvester (1982).

- Roberts, Harry V.
 1974 Conversational Statistics. Palo Alto: The Scientific Press
 Hewlett Packard University Business Series.
 1976 Conversational Statistics II. University of Chicago, Graduate School of Business. Mimeographed.
 1978 Comment on "The analysis of single and related time series into components: proposals for improving the X-11" by Raphael Raymond V. Bar On. Pp. 161-170 in Zellner (1978). Also see listings under Ling and Roberts.
- Roberts, Harry V., Robert F. Ling and George R. Bateman
 1979 Exploring Statistics with IDA. Palo Alto, California: The Scientific Press.
 Also see listings under Roberts, Ling.
- Rosenblatt, Harry M.
 1965 Spectral Analysis and Parametric Methods for Seasonal Adjustment of Economic Time Series. U.S. Department of Commerce, Bureau of the Census, Working Paper No. 23.
 A clearly written basic introduction to the spectral analysis of seasonal fluctuation.
- Schlicht, Ekkehart
 1981 Seasonal adjustment principle and a seasonal adjustment method derived from this principle. Journal of the American Statistical Association 76:374-378.
- Schneider, Anne L. and David Sumi
 1977 Patterns of Forgetting and Telescoping in LEAA Survey Victimization Data. Institute of Policy Analysis, 777 High Street, Suite 222, Eugene, Oregon 97401.
 Discusses seasonal patterns in victim survey responses. Suggests that respondents use the season of the year to assist their long-term memory of a victimization. Also notes that police reporting practices may affect the seasonality of crimes "known to the police." The aggregate number of unfounded crimes is subtracted each month from the aggregate number of reported crimes, rather than cancelling the actual crime report that was unfounded. The effect of this would be to reduce the intensity of any seasonal fluctuation. Also see US/BJIS (1980).
- Shine, Lester C., II
 1980 On two fundamental single-subject behavior functions. Educational and Psychological Measurement 40:63-72.
 1981 Integrating the study of Shine's actualized and pure single subject behavior functions. Educational and Psychological Measurement 41:673-685.
 1982 An illustration of how the effects of serial dependencies are handled in analyses of Shine's pure and actualized single-subject behavior functions. Educational and Psychological Measurement 42:87-94.
 An alternative approach to testing intervention hypotheses.

- Shiskin, Julius
 1957 Electronic computers and business indicators. The Journal of Business 30(4, October):219-267.
 This is a good introduction to the logic of the Census X-11 program, as it was originally developed. It reviews the X-11 in comparison to easy methods that are still common, such as same-month-last-year, monthly-means, and ratio-to-ratio-to-moving-average.
 1968 Time series: seasonal adjustment. Pp. 80-88 in International Encyclopedia of the Social Sciences 16. David L. Sills (ed.)
 Also see Tintner, et al. (1968).
 1978 Keynote address: Seasonal adjustment of sensitive indicators. Pp. 97-103 in Zellner (1978).
 Shiskin is considered to be the "father" of the Census X-11 program. The conference recorded in Zellner (1978) was organized to honor him.
- Shiskin, Julius, Allan H. Young, and John C. Musgrave
 1967 The X-11 Variant of the Census Method II Seasonal Adjustment Program. U.S. Department of Commerce, Bureau of the Census. Reprinted 1976.
 This is the Census X-11 user's guide.
- Sims, Christopher A.
 1974 Seasonality in regression. Journal of the American Statistical Association 69(347, September):618-626.
 Discusses bias in regression due to seasonal adjustment and seasonal noise. See Rosenblatt (1965), Wallis (1974).
- Stein, Donald P., Jay-Louise Crawshaw and Algrid R. Barskis
 1967 Computer-Aided Crime Prediction in a Metropolitan Area. Technical Reports 1-202 and 1-202-A, The Franklin Institute Research Laboratories, Philadelphia.
 1966 Part I offenses, 5 per cent sample. Predictors included weather, time of day, day of week, month of year, phase of the moon. Probability that a certain type of crime would occur, given that some crime did occur.
- Sylvester, Sawyer F.
 1982 Adolphe Quetelet: At the beginning. Federal Probation 46 (December, 4):14-19.
- Thiel, Henri
 1966 Applied Economic Forecasting. Amsterdam: North Holland Publishing Co.
 A classic economics forecasting text.
- Thompson, Howard E. and George C. Tiao
 1971 Analysis of telephone data: A case study of forecasting seasonal time series. The Bell Journal of Economics and Management Science 2(Autumn):515-541.
 Contains a diagnostic check for consistency, the implicit assumption "that the same relationship exists between observations 12 periods apart for all 12 months of the year."

- Tintner, Gerhard, P. Whittle, Herman Wold and Julius Shiskin
1968 Time series. Pp. 47-88 in International Encyclopedia of the Social Sciences 16. David L. Sills (ed.)
Also see Shiskin (1968).
- Tukey, John W.
1962 The future of data analysis. Annals of Mathematical Statistics 33:1-67.
1977 Exploratory Data Analysis. Reading, Mass.: Addison-Wesley.
- Tyron, Warren W.
1982 A simplified time-series analysis for evaluating treatment interventions. Journal of Applied Behavior Analysis 15 (3,Fall):423-429.
- Velleman, Paul F. and David C. Hoaglin
1981 Applications, Basics and Computing of Exploratory Data Analysis. Boston:Duxbury Press.
This is a beginner's guide to EDA (Exploratory Data Analysis). See Tukey (1977). It includes a discussion of resistant time series analysis methods. For a resistant seasonal adjustment method, see Cleveland, et.al. (1978).
- Vigderhous, Gideon
1978 Forecasting sociological phenomena: Application of Box-Jenkins methodology to suicide rates. Pp. 20-51 in Sociological Methodology 1978.
Good overview of stochastic time series analysis methods.
- United States, Bureau of Justice Statistics
1980 Crime and Seasonality. National Crime Survey Report SD-NCS-N-15,NCJ-64818. Report written by Richard W. Dodge and Harold R. Lentzner, Crime Statistics Analysis staff, Center for Demographic Studies, U.S. Bureau of the Census.
Although plagued by short series and other problems, this is the most comprehensive seasonal analysis of victim data to date. Uses Census X-11 with 1973-1977 National Crime Survey data. Finds stable seasonality F values of 10.0 or higher for household larceny (under and over \$50), personal larceny without contact (under and over \$50), and unlawful entry without force. Also see the pioneering article on this subject, Schneider and Sumi (1977).
- Wallis, Kenneth F.
1974 Seasonal adjustment and relations between variables. Journal of the American Statistical Association 69(March,345):18-31.
Argues that the use of seasonally adjusted and unadjusted data in the same model may lead to spurious dynamic relationships.

- Warren, Charles W., Jack C. Smith and Carl W. Tyler
1981 Seasonal variation in suicide and homicide: A question of consistency. Unpublished manuscript. Public Health Service, U.S. Centers for Disease Control, Atlanta, 30333.
Although this paper does not explicitly define seasonality, the implicit definition includes the possibility of year-to-year inconsistency. Example of PRA (periodic regression analysis). See Bliss.
- Willson, Victor L.
1973 Estimation of intervention effects in seasonal time-series. University of Colorado, Laboratory of Educational Research, Report No. 63.
Compares four methods of handling seasonality (linear sine term, prior seasonal adjustment, differencing, and ignoring the seasonal component) with seven simulated series. Finds that a sine term "works best in cases where error variance and amplitude are of the same order of magnitude. Seasonal adjustment seems better for situations when the amplitude is much larger than the error variance. Differencing was a poor method in all cases." Also see Hibbs (1977).
- Wolfgang, Marvin E.
1966 Patterns in Criminal Homicide. New York: John Wiley & Sons
See pp. 96 to 106 for a review of research on seasonality of crime, from the early 1800's. Also see Quetelet (1842), Lester (1972), and US/BJS (1980).
- Zellner, Arnold (ed.)
1978 Seasonal Analysis of Economic Time Series. Proceedings of the Conference on the Seasonal Analysis of Economic Time Series, September 9-10, 1976, Washington, D.C. U.S. Department of Commerce, Bureau of the Census, Economic Research Report ER-1.
This is an extremely valuable source book. Unfortunately, it is out of print. The Illinois State Library has a copy.

END