HOW TO HANDLE SEASONALITY

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

May 1983
second edition, July 1984

ILLINOIS
CRIMINAL JUSTICE
INFORMATION AUTHORITY

96376

HOW TO HANDLE SEASONALITY

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

May 1983
second edition, July 1984

by Carolyn Rebecca Block
Statistical Analysis Center

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

Printed by authority of the State of Illinois
July 1984
Number of copies: 100
Printing order number: 85-5

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
Two Traditional Approaches....9
The Component Definition of Seasonality....11
The ARIMA Definition of Seasonality....16
TOOLS FOR DETECTING AND ANALYZING SEASONALITY....19
Component Methods....19
Definitions....20
Moving Average....20
Additive/Multiplicative Assumption....21
Rules of Thumb....22
F of Stable Seasonality, and Relative Contribution of the Irregular....23
Average Duration of Run of the Irregular....25
Months for Cyclical Dominance....26
Pattern Consistency....28
Trading Day Option....30
Appropriate Applications....31
Extremes....31
Series Length....31
Discontinuities....32
Moving Seasonality....32
ARIMA Methods....35
Definitions....35
Moving Average and Autoregressive Processes....35
Identifying the Process of a Series....36
Stationarity....41
Differencing....44
Rules of Thumb for Evaluating a Model....48
Correlogram of Residuals....48
Cumulative Periodogram of Residuals....50
Appropriate Applications....52
Extremes....54
Series Length....54
Discontinuities....54
Moving Seasonality....55
ANNOTATED BIBLIOGRAPHY....57

Preceding page blank

EXECUTIVE SUMMARY

This report is an introduction to the fundamentals of seasonal analysis, with an emphasis on applications to applied social research, especially criminal justice. Administrators, policy makers, researchers, and others who make decisions based on social indicators 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. Other social sciences 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 describing 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.

It is impossible to give a brief, standard definition of seasonality. None exists. Although the major methods of detecting and analyzing seasonal fluctuation imply different underlying concepts of seasonality, these conceptual definitions are seldom stated explicitly. As a result, researchers analyzing the same data may come to confusingly different conclusions. Therefore, the first section of this report introduces the reader to the concepts of seasonality that underlie seasonal analysis and seasonal adjustment. It gives the reader the background information necessary to choose the appropriate method for a given situation and to interpret the results of analysis conducted by others.

This report discusses the two major approaches to defining and detecting seasonality. Although the two approaches are mathematically similar, there are practical differences in emphasis. One approach, called "seasonal adjustment" or "Census X-11 adjustment," emphasizes a separate description of seasonal fluctuation; the other approach, the most common example of which is known as "ARIMA," emphasizes forecasting the future with a model that incorporates seasonality. The first approach focuses on seasonality itself, while the second 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 approach offers a simple, objective, yes-or-no criterion for detecting the presence of seasonality in time series. Both 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 (more than 130 references) of literature about seasonal analysis and reports analyzing the seasonality of crime.

ACKNOWLEDGEMENTS

This report has been a "working paper" for several years. During that time, many people commented on its various drafts. The staff of the Statistical Analysis Center, especially James Coldren and Paul Fields, made helpful suggestions. The comments of Richard L. Block of Loyola University of Chicago were also extremely helpful.

The staff of the U.S. 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 time series modeling. They were never too busy to offer hints and practical suggestions for using the package and for interpreting its results.

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.

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

The second edition benefits from the comments and suggestions of Robert J. McIntire of the Bureau of Labor Statistics and Estela Bee Dagum of the Seasonal Analysis and Time Series Branch of Statistics Canada. Richard E. Barrett of the University of Illinois used the first edition as a manual for seasonal analysis with SAS/ETS, and made helpful suggestions.

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. The concept seems, at first glance, to be simple. Criminologists, for example, have traditionally believed (see Wolfgang,1966:96-106) that more crimes occur during some months of the year than others.[2] However, 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 was 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.

---

[1] A complete review of all seasonal analysis methods would fill at least one book. This report is limited to the two most commonly used methods. 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.(1978), Levenbach and Cleary (1981), or Velleman and Hoaglin (1981) for resistant methods, Rosenblatt (1965) for spectral analysis; and Glass, et al. for regression of a seasonal covariate. For a technical guide to using the seasonality and other time series computer programs that are available at SAC, see Miller (1982).

[2] For a discussion of issues particularly relevant to the analysis of seasonal fluctuation in crime, a review of research on seasonality of crime, and the results of seasonal analysis of 135 Index crime series, see Block (1984).

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 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. 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, even if criminal victimizations occur without seasonal fluctuation, these victimizations may become known to the police more frequently during certain seasons of the year. The number of aggravated assault offenses known to the police in the United States (figure 1) is seasonal, but the number of aggravated assault victimizations (figure 2) is not seasonal. The comparison of these series tells us as much about police reporting practices as it does about the seasonal nature of violent crime.[3]

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 research, administrative or policy decisions?

If we make the first error, to ignore the question of seasonality in a series that is seasonal, we ignore information that may be useful in making decisions. Descriptions of the pattern of change over time in a series, including the pattern of seasonal fluctuation, 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.

---

[3]For details of the analysis of these two series, see Block (1984:9-13).

Figure 1

ASSAULT KNOWN TO THE POLICE, UNITED STATES, 1973-1979

SOURCE: UNIFORM CRIME REPORTING PROGRAM, FBI

THESE ARE RAW FIGURES. THEY ARE NOT WEIGHTED TO ACCOUNT FOR JURISDICTIONS THAT DID NOT REPORT. THE POPULATION REPORTING INCREASED 13% FROM 1973 TO 1979.

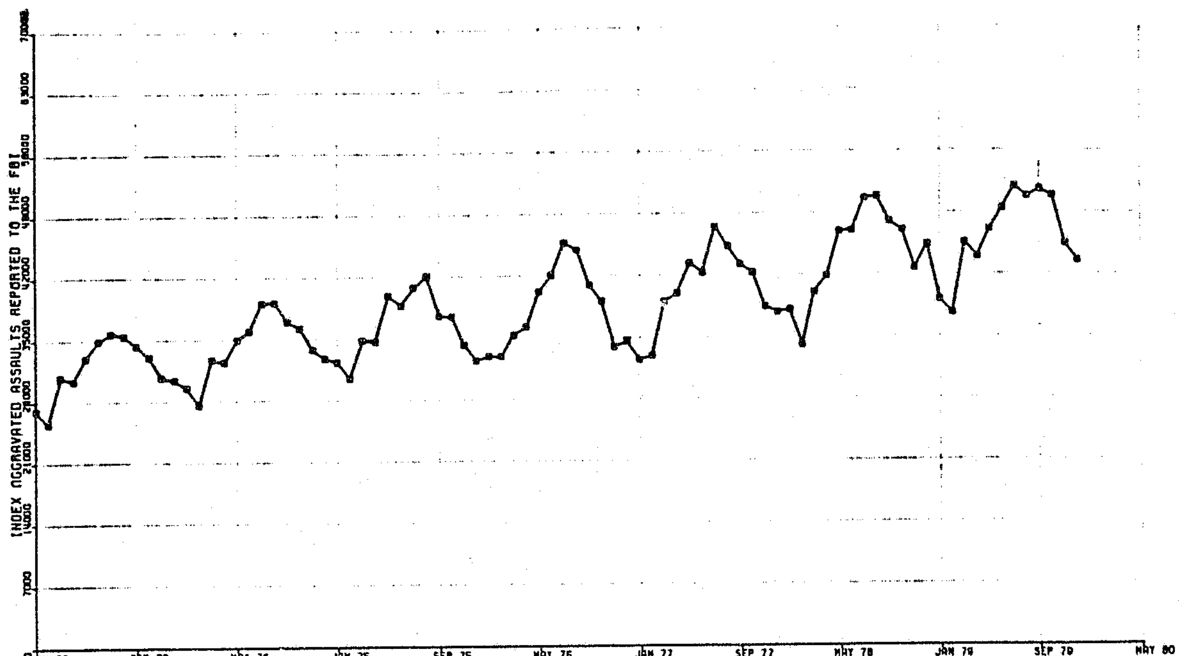
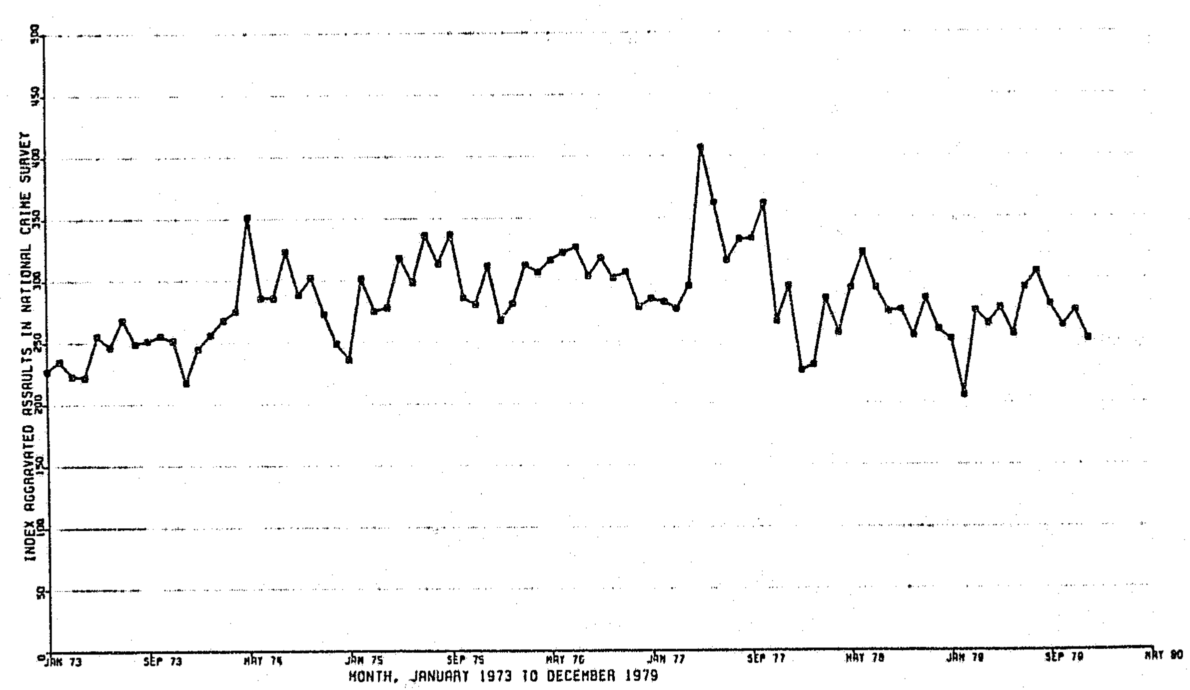


Figure 2

ASSAULT VICTIMIZATIONS, UNITED STATES, 1973-1979

SOURCE: NATIONAL CRIME SURVEY

NUMBER OF ASSAULTS IN SAMPLE, WEIGHTED TO CORRECT FOR SAMPLING BIAS.
DATA PROVIDED BY RICHARD BLOCK AND WESLEY SKOGAN.



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 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 description -- description of the pattern of seasonal fluctuation and description of the pattern of the variable with the seasonal fluctuation removed -- can make analysis results easier to communicate to a general audience (see Granger 1978:38-39). Seasonal fluctuation may be so great that it obscures any other pattern. Removing variation due to a known cause, seasonality, makes these other patterns easier to see.

Suppose that a reporter or a member of the City Council asks the crime analysis unit of the local police department 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, than if it is accompanied by a graph of the original data. Compare figure 3 and figure 4.[4] There is much less variation in the seasonally adjusted larceny series than in the original larceny series. With seasonal fluctuation removed, the general pattern of larcenies over time appears much more clearly. The raw data seem to climb steadily to mid-1975, and then level off. With seasonal fluctuation removed, it becomes apparent that the climb began later and lasted longer. In addition, departures from the general pattern become clear. The extremely low observation in May, 1979 is much more apparent in figure 4 than in figure 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). It results in the same descriptive mistakes discussed above: an erroneous assumption that all Mays are higher than average, for example, might lead to a misallocation of May resources.

---

[4]The lines superimposed on the raw data in figures 3 and 4 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 Block (1983).

Figure 3

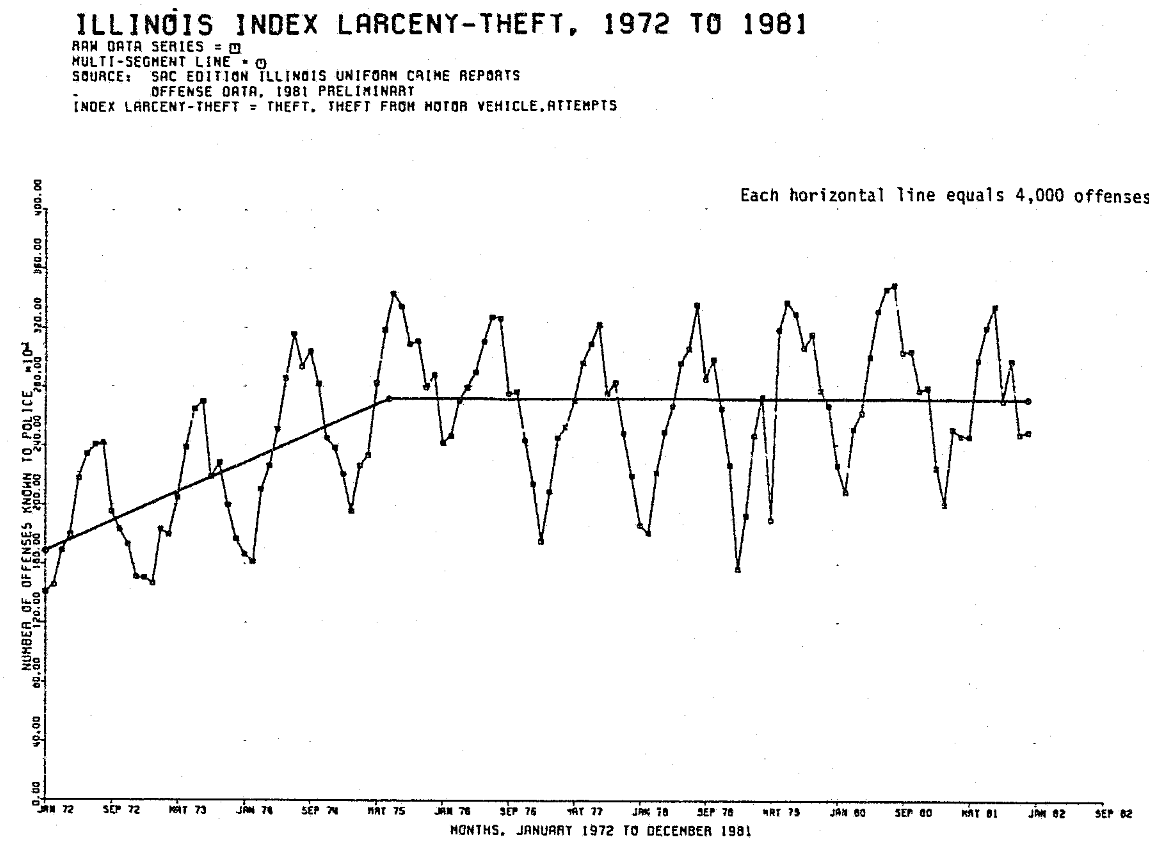
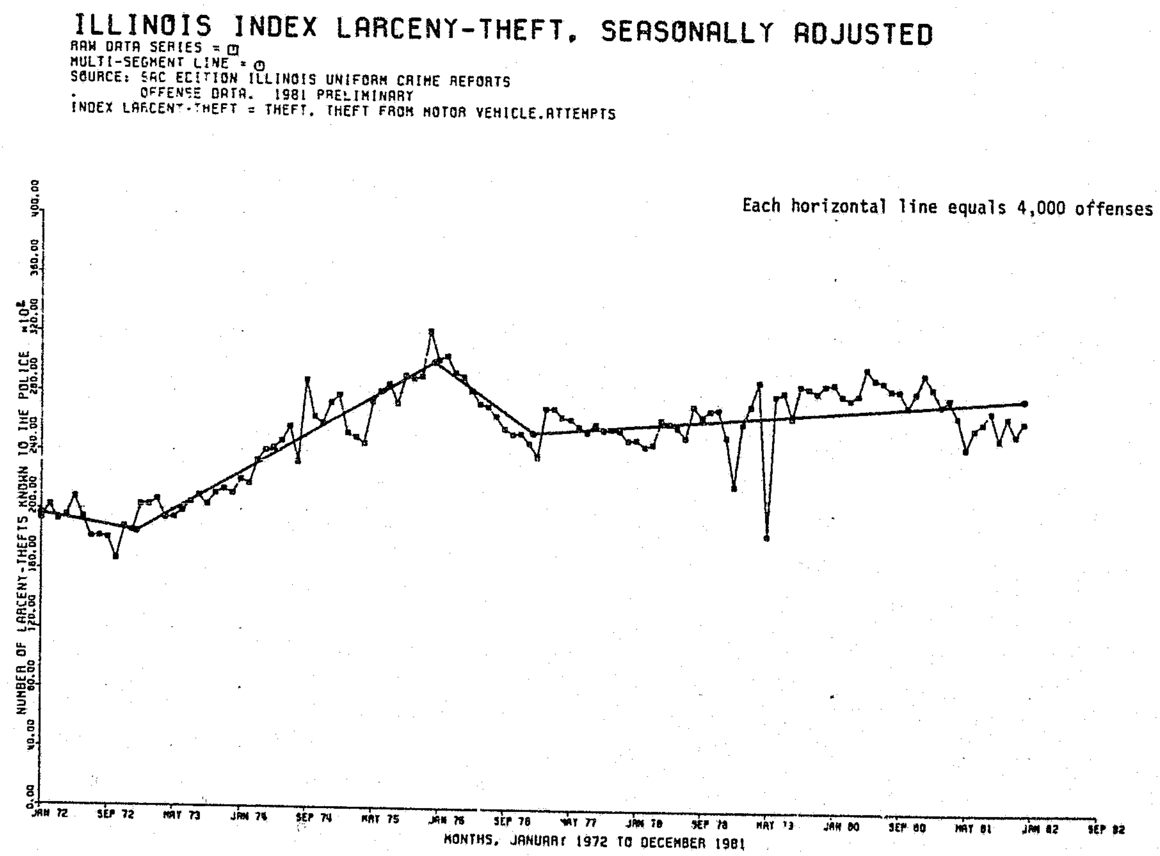


Figure 4



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. Such a misspecified model "overadjusts" for seasonality; it removes or otherwise controls for seasonal fluctuation that never existed.[5] This transformed series 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.[6] 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?

---

[5]For discussions of the problem of overadjustment, see page 47 below, Nettheim (1965), Rosenblatt (1965), Grether and Nerlove (1970:682-683), Kalleck (1978), or Dagum (1981:135). In their forecasting competition, Makridakis, et al. (1982:127) conclude that one reason that simple methods do well in comparison to statistically sophisticated methods is that the sophisticated methods "extrapolate too much trend which can cause overestimation." For a discussion of other errors that may result from erroneous assumptions about seasonality in a regression model, see Wallis (1974).

[6]Even if we know a particular series is seasonal, some decisions would 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 sea- sonal 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.

Two Traditional Approaches

There are two major empirical approaches to defining and detecting the presence of seasonality. 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. However, there are practical differences in emphasis. One approach emphasizes a separate description of seasonal fluctuation. The other emphasizes forecasting the future with a model that incorporates seasonal fluctuation. The first approach is primarily interested in seasonality itself. The second is interested in seasonality as it affects the accuracy of a forecast.

Both approaches model seasonal fluctuation.[7] They are both descriptive, in the sense that any statistical model is a description of reality. However, the first model emphasizes separate descriptions of the seasonal component and the rest of the series; the second does 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

---

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

Preceding page blank

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 approach that emphasizes the separation of the series into seasonal and nonseasonal components is commonly referred to as "seasonal adjustment" or "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, Canada, 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.

The most commonly used example of the second approach, incorporating seasonal fluctuation in a unified model, is ARIMA, a mnemonic for AutoRegressive Integrated Moving Average.[9] ARIMA is also known as the "Box/Jenkins" method, in reference to George Box and Gwilym Jenkins who, with George Tiao, developed it into a comprehensive theory.[10] The method has become so popular that authors sometimes use "ARIMA" and "time series analysis" as if they were synonymous (for example, see McCleary and Hay,1980; Nelson,1973; Glass, et al. 1975).

The division of the world of seasonal analysis into component and ARIMA approaches is somewhat arbitrary. The two actually have close mathematical similarities (Pierce,1980:126-128). The components of a series can be estimated with ARIMA methods, and an ARIMA model may contain seasonal and nonseasonal terms that can be thought of as components.[11] The difference between the two approaches to seasonal analysis exists more in the way they are

[8]For more information on the Census X-11 and other seasonal component methods, see Shiskin (1957), Shiskin, et al. (1967), Plewes (1977), Grether and Nerlove (1970), Hannon (1960,1963), Lovell (1963), Levenbach and Cleary (1981), Willson (1973), Nettheim (1965), and Rosenblatt (1965). In addition, more than 20 papers on aspects of seasonal adjustment and analysis are contained in Zellner (1978).

[9]These terms are explained in the "ARIMA Methods" section, below.

[10]The present discussion is only a brief guide to ARIMA, especially as it pertains to seasonal analysis. For a more complete, but still elementary, treatment, see Nelson (1973) or McCleary and Hay (1981). Also see the bibliography, below, for brief reviews of more advanced texts.

[11]In fact, the X-11/ARIMA computer package uses ARIMA models to produce better estimates of the components of a series. See page 19, below.

used and interpreted by analysts, rather than in their mathematical properties. Think of them as photographs of the same object, taken from different angles. They describe the same phenomenon, but from different perspectives.

Despite their technical similarity, social researchers and policy analysts, especially in fields other than economics, tend to use one method to the exclusion of the other. Actually, 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 (Grether and Nerlove,1970; Kuiper,1978; Armstrong,1978; Granger,1978; Makridakis, et al.,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. Therefore, instead of being tied to a single approach, analysts should become familiar with all the alternatives, and choose the particular method that suits the question at hand.

In the rest of this report, we provide information and practical examples to assist analysts in making a rational choice among seasonal analysis methods. First, we review the conceptual definitions of seasonality that underlie the two most commonly used approaches. Then, we discuss and compare the tools for detecting and analyzing seasonality that the two approaches offer, and give the analyst some basic rules of thumb for using these tools in various practical situations.

The Component Definition of Seasonality

The component concept of seasonality is expressed in Kallek's 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.

Figure 5

CHICAGO LONG GUN REGISTRATIONS, JANUARY 1969-JULY 1980

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

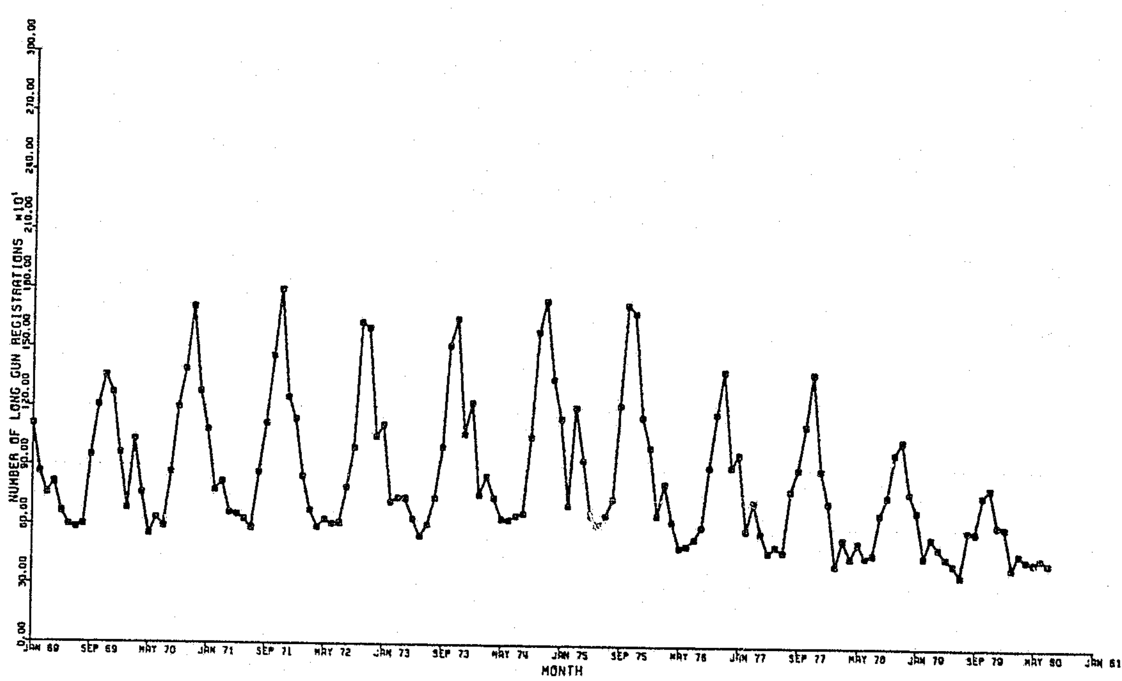
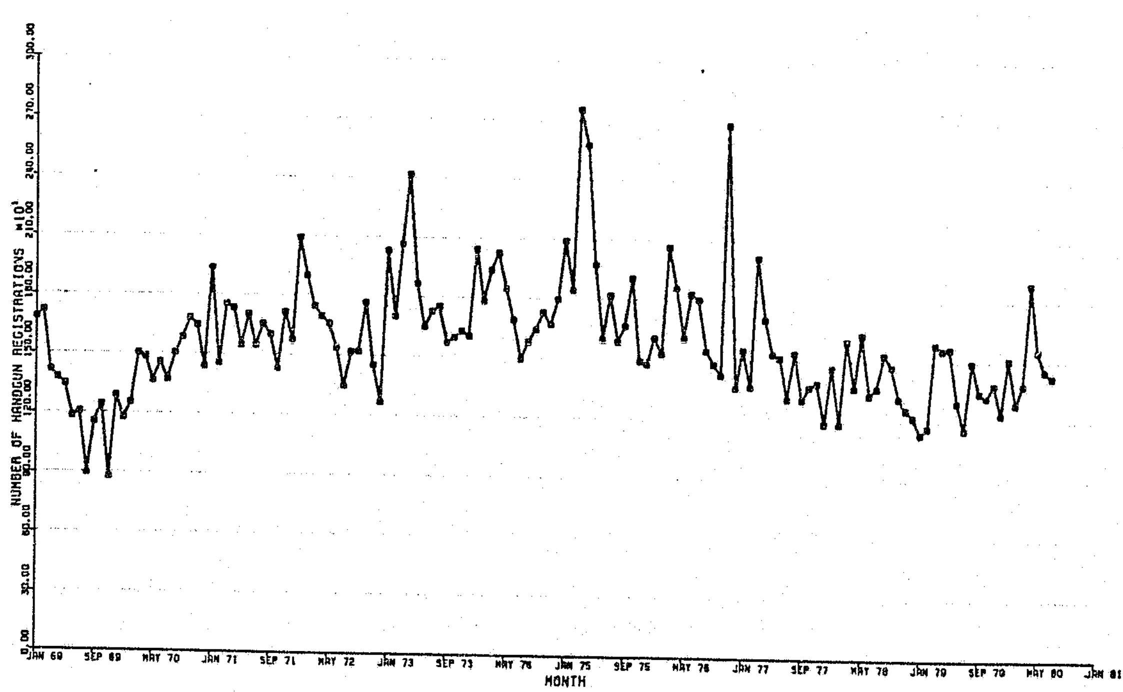


Figure 6

CHICAGO HANDGUN REGISTRATIONS, JANUARY 1969-JULY 1980

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



Although this concept seems simple, the measurement ("operationalization") of the concept is not. 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 other 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 to 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?

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 non-seasonal 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, including error, 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.[12] 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.

Thus, the problem of detecting seasonality, using the component approach, 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.[13] (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 series is seasonal, according to the component approach.

---

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

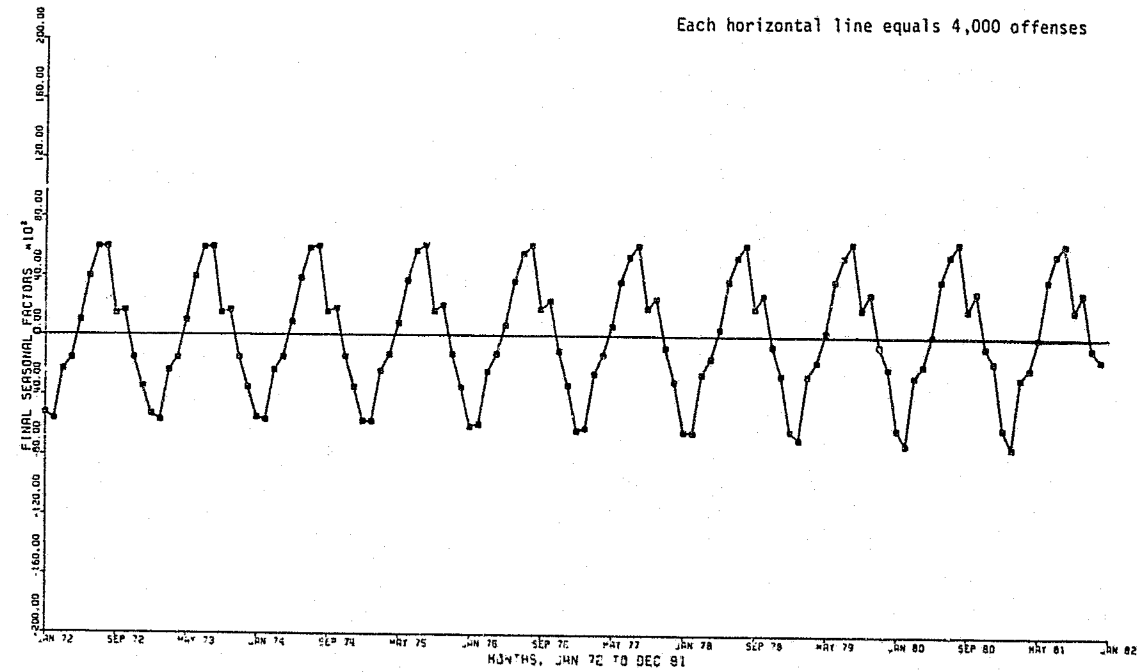
[13]A moving average replaces each observation with the average of that observation and the observations that occur just prior to it and just after it. See pages 20-21, below.

For example, the three components of the larceny/theft series discussed above (figures 3 and 4) are shown in figures 7, 8, and 9. Figure 7 shows the seasonal fluctuation, figure 8 shows the irregular, and figure 9 shows the trend/cycle.[14]

Figure 7

ILLINOIS INDEX LARCENY-THEFT, FINAL SEASONAL FACTORS

SOURCES: ORIGINAL SERIES, SAC EDITION ILLINOIS UNIFORM CRIME REPORTS OFFENSE DATA. SEASONAL FACTORS, SAC EDITION, U.S. BUREAU OF THE CENSUS X11 PROGRAM.



[14]These 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

SOURCES: ORIGINAL SERIES, SAC EDITION ILLINOIS UNIFORM CRIME REPORTS OFFENSE DATA. IRREGULAR SERIES, SAC EDITION, U.S. BUREAU OF THE CENSUS X-11 PROGRAM, ADDITIVE ASSUMPTION.

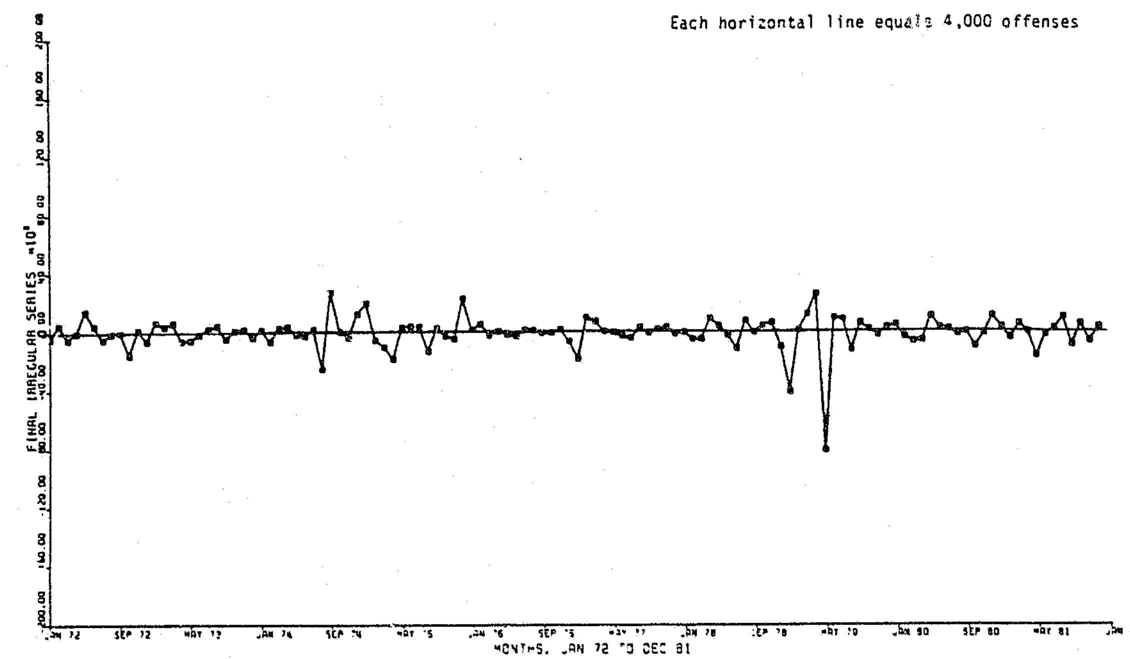
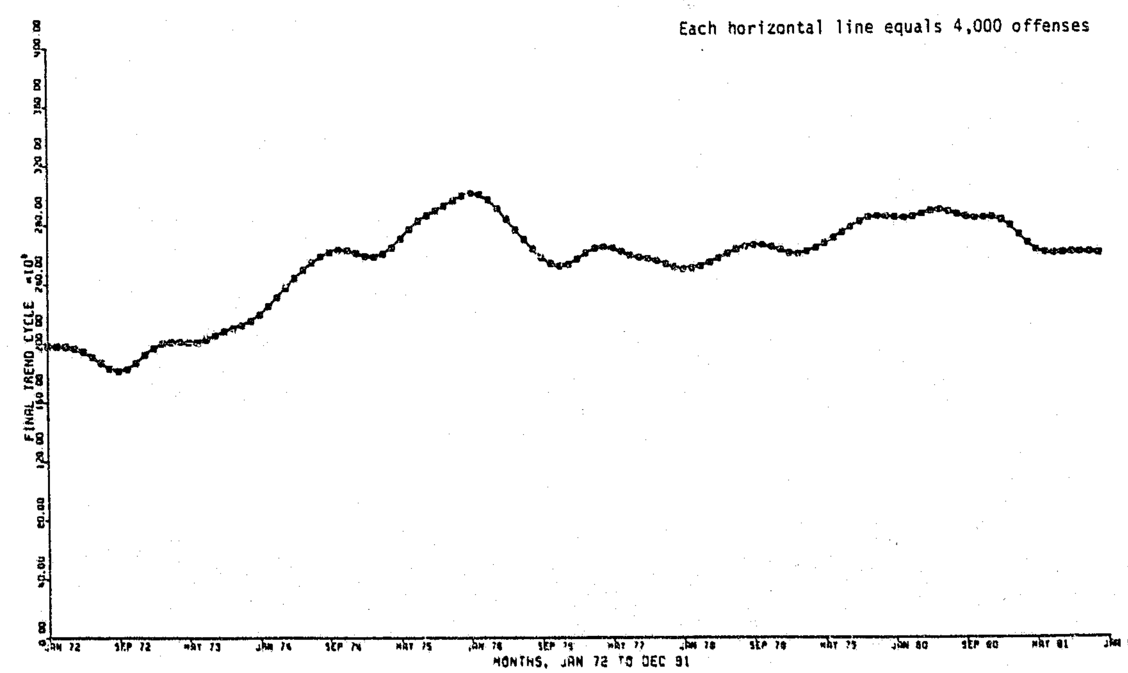


Figure 9

ILLINOIS INDEX LARCENY-THEFT, FINAL TREND CYCLE

SOURCES: ORIGINAL SERIES, SAC EDITION ILLINOIS UNIFORM CRIME REPORTS OFFENSE DATA. TREND CYCLE, SAC EDITION, U.S. BUREAU OF THE CENSUS X11 PROGRAM.



The ARIMA Definition of Seasonality

In the ARIMA 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.

Like Kallek's component definition (page 11, above), the ARIMA definition of seasonality emphasizes the existence of regular periodic fluctuation. However, unlike the component definition, the ARIMA definition does not emphasize separating this fluctuation from the rest of the series.

The ARIMA approach is not so much concerned about describing the past as it is about forecasting the future. Of course, any forecast of the future must begin with a description of the past. However, Box and Jenkins (1976:301) emphasize that descriptions of each of the components of a series, even though each separate description may be good, will not necessarily produce a good forecast of the whole. Therefore, the ARIMA approach does not describe the series by describing each separate component. Instead, it describes the "stochastic processes" of the entire series.

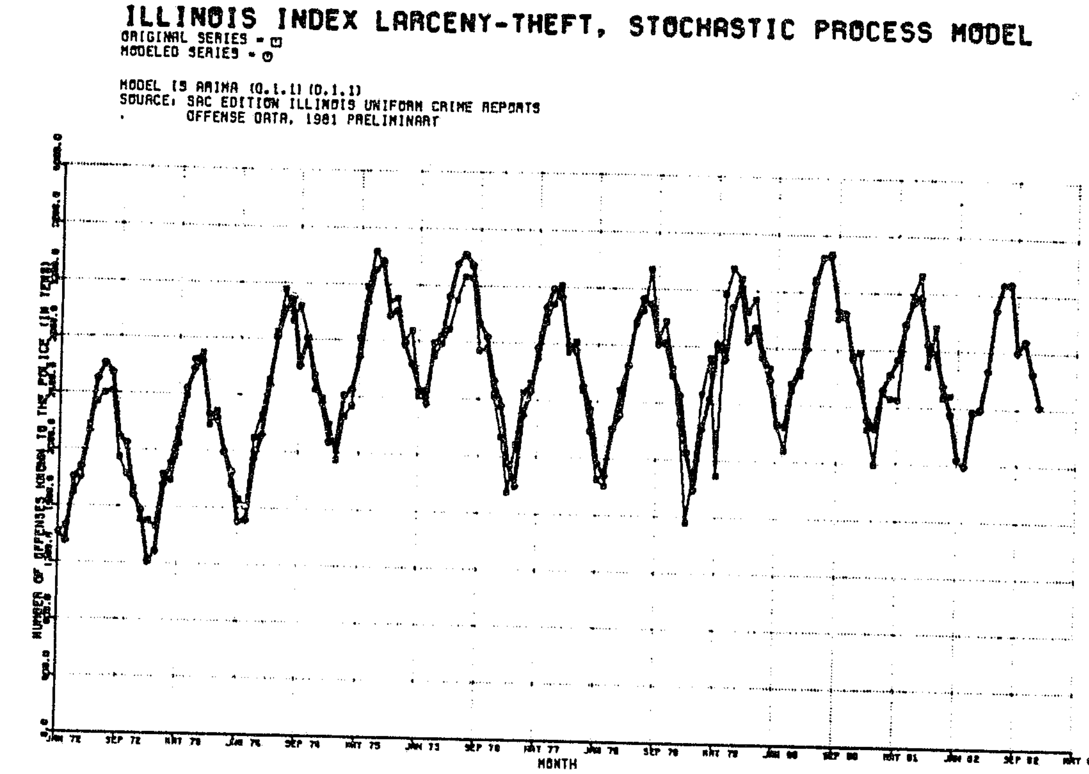
In a stochastic process, one observation follows the next with a certain probability. In a monthly series with seasonal fluctuation, observations 12 months apart are correlated, which means that they follow each other with a certain probability.[15] Thus, seasonality may be part of a stochastic process.

ARIMA assumes that a time series has followed some unknown but identifiable pattern in the past. If we can determine the probability, or set of probabilities, under which observations followed one another in the past (the stochastic processes), and if the same processes continue unchanged, then we can forecast the future accurately. The analyst's problem is to identify, or model, the processes of a series. One of the processes in a series could be a seasonal process. If the analyst's model includes such a seasonal process, the ARIMA approach concludes that the series is seasonal.

---

[15]The opposite is not always true. If observations 12 months apart are correlated, the series is not necessarily seasonal. See figures 15 and 16 and accompanying discussion, below.

Figure 10



For example, the best ARIMA model for Illinois larceny/theft suggests that the series follows a seasonal process: the current observation is related to the error of the observation 12 months ago, and, in addition, the current observation is related to the error of the preceding observation.[16] Figure 10 shows the original larceny/theft series (light line) and the modeled series (dark line). The original series ends in December, 1981, but the model includes a forecast through 1982.

How was this forecast calculated? The model states that each observation is related to, and can be calculated from, the immediately preceding observation and the observation one year ago. Therefore, the actual number of larceny/thefts in the years 1972 through 1980 are used to calculate the modeled values for 1973 through 1981. For example, the December, 1981 modeled value is calculated from the actual November, 1981 and December, 1980 values. The January, 1982 modeled value is calculated from the actual December, 1981 and January, 1981 values. To forecast for February, 1982, the model uses the actual February, 1981 observation and the modeled value for January, 1982. By continuing this process, the model forecasts all 1982 values.

---

[16]The model realized in figure 10 is a (0,1,1)(0,1,1) ARIMA process (a first-order serial and seasonal moving average process with serial and seasonal differencing). For definitions and details, see "ARIMA Methods," page 35 below.

TOOLS FOR DETECTING AND ANALYZING SEASONALITY

Neither the component nor the ARIMA 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

This section is a guide to using and interpreting component seasonal analysis, with particular emphasis on the Census X-11 seasonal adjustment program.[17] Although there are other computer packages that partition a series into components, most of these have options and results that are similar to the X-11. Thus, someone who is able to use the X-11 should have little difficulty using the other packages.

There are, in general, three kinds of component packages other than the X-11. The X-11/ARIMA program identifies an ARIMA model for the series (if possible), forecasts one year, and then uses the forecasted values to calculate the component part of the program. Therefore, to interpret X-11/ARIMA results, you must be familiar with the interpretation of standard X-11 results. Most econometric packages contain some sort of moving average routine. If you understand X-11 moving average options, you will also be able to use these packages. The SABL program (Seasonal Adjustment-Bell Laboratories; see Cleveland, et al.1978) differs from traditional component programs in several ways, such as its use of "resistant" smoothers instead of a moving average.[18] If you understand the general concept of component analysis, you will find it easier to understand SABL.

---

[17]See page 10, above. The Census X-11 is available from the U.S. Bureau of the Census and from the Bureau of Labor Statistics, U.S. Department of Labor. It is also a part of the SAS/ETS (Econometric and Time Series) package. The X-11/ARIMA package is available from Statistics Canada. It is now the official method of the Bureau of Labor Statistics for the seasonal adjustment of household and establishment survey data on labor force, unemployment, employment, and hours (McIntire,1983). A quick, abbreviated component program, developed by Statistics Canada, is a useful screener for the presence of seasonal fluctuation in a series (see Block,1984).

[18]A resistant smoother resists the effect of extreme values. A moving median is resistant, in contrast to a moving mean. See Velleman and Hoaglin (1981) and Velleman (1980,1982).

Definitions

Moving Average[19]

A moving average "smooths" a time series.[20] It separates the series into two parts -- a sequence of values that follow each other smoothly, with relatively little variation from one to the next, and another sequence of values that vary relatively more from one to the next. There are many kinds of smoothers; a moving average is just one of them. There are also many kinds of moving averages, each with a different effect on the data.

A moving average replaces each observation with a weighted average of that observation and observations occurring before it and after it in the sequence. For example, to calculate a moving average with a five-month span, you would calculate the average of observations 1 through 5, then the average of observations 2 through 6, the average of observations 3 through 7, and so on to the end of the series. The result would be a transformed series in which random and periodic fluctuation occurring within a five-month frequency is "averaged out." The new series would be shorter than the original series, by two observations at the beginning and two observations at the end. This is called the "end effect," and may be important if we are particularly interested in the most recent past of the series.[21]

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. These other patterns will be more clearly discernible in the smoothed series than in the original series. However, not every kind of moving average

[19]The concept and calculation of moving average in the context of component methods is very different from the moving average process in ARIMA (see page 35, below). 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 (Nelson 1973:33). In the context of spectral analysis, a moving average is called a "filter," specifically, a "linear filter" (Hamming, 1983). For example, a "low-pass filter" removes high frequency periodicity (see Kendall 1976:44 and "Cumulative Periodogram of Residuals," page 50, below).

[20]Moving averages are also called "running means." For an elementary review, see Macaulay (1931). For a more advanced discussion, see Kendall (1976:53) or Dagum (1983c). For a review that emphasizes nonlinear smoothing, see Velleman (1982).

[21]There are various statistical techniques to handle the end effect. See Kendall (1976) for a review. The X-11/ARIMA method uses an ARIMA forecast to estimate end values. See Dagum (1983a).

meets this goal for every series.[22] The extent (span) of the average, and the amount each observation is weighted in the average, affect the kind of variation that the transformation removes. For example, a moving average with a five-month span would smooth random variation from month to month and periodic fluctuation occurring every four or five months, but not periodic fluctuation occurring at a 12-month frequency. 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 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.

The X-11 program conducts iterative approximations of the best moving average. The program also offers the user a choice of types of moving average, as well as a choice of treatments for extreme values. Before attempting to use these options, become familiar with the effects of various moving averages (Kendall, 1976). Then, consult Shiskin, et al. (1967) for specific moving average options offered by the package.

Additive/Multiplicative Assumption[23]

The three components--seasonal, trend/cycle, and irregular--may be related to each other in two ways. They may be independent or dependent. 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 whether the total number of larcenies were

[22]For 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). For the "Gibbs" phenomenon, in which linear filters treat some frequencies inconsistently, see Velleman (1983:143) and Hamming (1983: 93-101).

[23]For a discussion of additive versus multiplicative relationships in ARIMA models, see Box and Jenkins (1976:322-324). ARIMA models, like component models, usually assume that the relationship is multiplicative.

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 (Block,1984).

A good general procedure 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 diagnostic tests discussed below.[24] The two assumptions usually produce very similar results, but, when they do not, assume that the better adjustment, additive or multiplicative, reflects the true underlying relationship among the three components.[25]

Rules of Thumb

The 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 that were used to arrive at the decision, and the results of each test.

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

---

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

[25]For an important series in which the multiplicative or additive relationship is not clear, it may be necessary to use more complex analytical methods. The literature on seasonal adjustment contains may 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.[26] 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, we cannot necessarily assume that successive observations are independent (Anderson,1950). Therefore, there is some question as to the proper interpretation of this F value.[27]

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.

The "relative contribution of the irregular" varies from 0 percent to 100 percent, 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.[28] It indicates the absolute importance of each component to the variation in the total series. 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:

---

[26]Stable seasonality assumes that seasonal fluctuation is constant from year to year. For "moving seasonality," see pages 32-33, below.

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

[28]In X-11 printed results, we also find the relative contributions of each of the three components to variation over a two-month span, three-month span, etc., up to a 12-month span.

Stable Seasonality F Value	Percent Contribution of Irregular	Decision
0.00-2.40	any percent	no stable seasonality
2.41-15.00	greater than 14%	no stable seasonality
15.01-50.00	greater than 25%	no stable seasonality
50.01 and up	greater than 30%	no stable seasonality

To this rule of thumb, we would add a qualification.[29] 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 greater than 15 and a percent contribution of the irregular about 30, before rejecting the stable seasonality hypothesis, consider 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 about the presence of seasonality in the series.

Thus, 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 (page 14), 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

[29]Kathryn Beale (Bureau of Labor Statistics) pointed this out. There is an additional complication in the case of a series that has a very weak trend/cycle component and a strong seasonal component (McIntire,1983). The seasonally adjusted series (the original series with seasonal fluctuation removed) may contain little else than random error. With such series, calculation and analysis of a seasonally adjusted series is fairly useless.

60 percent. Therefore, other diagnostics should be consulted before making the final decision.[30]

Average Duration of Run of the Irregular

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 quality of the seasonal adjustment should be questioned.

The ADR is the mean length of runs of values consecutively higher (or lower) than the preceding value. The higher the ADR of a series, the fewer the total number of runs. 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:

> 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 quality of each of the three adjustments can be trusted. Each of the irregular components varies randomly over time, as we would expect them to do. These random 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 13 police periods per year. This moving average probably has less irregular variation than the original series, resulting in an overly smooth irregular.

[30]The 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.

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 event, do not accept the adjustment unless other indicators, especially the F of stable seasonality and the percent contribution of the irregular, are unequivocal.

<u>Months for Cyclical Dominance</u>

Months for cyclical dominance (MCD) compares the relative contribution of the trend/cycle to the relative contribution of the irregular. Like the ADR of the irregular, the MCD is an indication of the quality of the adjustment (the extent to which you should trust the results).

As discussed above (note 28), 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 percent contribution of the irregular over a one-month span. 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 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 12-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. The MCD of the Index aggravated assault series is 6 (table 2). The MCD of the Index homicide series (not shown) is over 12 months (the trend/cycle contribution does not exceed the irregular contribution at any span from one to 12 months).

Table 1

Relative Contributions of Components to Variance
<u>Illinois Larceny/Theft, Additive Adjustment</u>

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
<u>Illinois Aggravated Assault, Additive Adjustment</u>

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

Notice that the contribution of the seasonal component drops close to zero over a 12-month span. This makes sense, because, by definition, seasonal values 12 months apart are similar to each other. The seasonal differencing technique (see "Stationarity," page 41 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 percent, 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 <u>and</u> 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 percent 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 "Pattern Consistency").

<u>Pattern Consistency</u>

Consistency in the seasonal pattern is another important consideration in determining whether or not a series is seasonal. Both the component and ARIMA approaches include consistency, or regularly evolving fluctuation, in their conceptual definition of seasonality. While a gradual change from year to year may indicate moving seasonality (see pages 32 to 33 below), abrupt change or change in sign from one year to the next argue against the hypothesis that the series is seasonal, by the usual definition.[31]

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 10-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 10-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.

---

[31]See Warren, <u>et al</u>. (1981) for an example of an analysis of seasonality that does <u>not</u> include year-to-year consistency in the definition.

The "seasonal factor" table, produced by the X-11 program, allows us to examine year-to-year and within-season consistency. A seasonal factor table contains one value per month or quarter. For example, there are 144 seasonal factors for a 12-year monthly series (table 3). Each seasonal factor indicates the degree to which the month in question is relatively high or low due to seasonal fluctuation.

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, a seasonal factor of 20 for a certain month in a homicide series would indicate that that month was seasonally high by about 20 homicides. With the standard deviation, which the table also includes, you can decide whether 20 homicides should be considered high or within the normal range.[32]

Table 3

Final Seasonal Factors, Multiplicative Adjustment
<u>Male Homicide Victims, Chicago: 1965-1978</u>

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

---

[32]For a similar diagnostic check for consistency, but using ARIMA methods, see Thompson and Tiao (1971:540-541).

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 discernibly high and others low in the number of male homicide victims, there is no consistent pattern from year to year.[33] If we define a "high month" as 10 percent high, and a low month as 10 percent low, 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. The seasonal factors of July, August, September, October, and November all change from year to year. Only one month, February, is consistently high or low throughout the 14 years, although some argument could be made for July being high. If we consider all the evidence, including this lack of seasonal consistency, as well as the low stable seasonality F (4.00), the high relative contribution of the irregular (63 percent), an irregular ADR of 1.52 (indicating that the irregular does not contain any seasonal fluctuation), and an MCD higher than 12 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.

Therefore, analysts who are primarily interested in diurnal periodicity might want to analyze daily data, if available, in preference to trading-day 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 both, 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. For an example of a model of daily variation that combines ARIMA and regression, see Bell and Hillmer (1983).

---

[33]We chose this homicide series as an example because it contains the _most_ seasonality of any homicide series we have analyzed (Block and Block,1980; Block,_et al_.1983; Block,1984). The common assumption that homicide is seasonal (Wolfgang,1966; President's Commission,1967; Warren,_et al_.1981) is due to differing definitions and measures of both homicide and seasonality (Block,1984).

However, there are limits to the use of the trading-day option. It will not provide accurate estimates when the average absolute month-to-month change of the irregular component is 8 percent or more (Shiskin _et al_.1967; McIntire,1983). 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 it is necessary to describe the patterns in a large number of series (for example, seven Index crimes in Illinois's 102 counties -- 714 series). The X-11 also provides standard measures that can be compared from series to series (see, for example, Block, 1984). Component analysis is especially 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 a few extreme values. Because it is based on a moving average, it is not resistant in the same way that a nonlinear smoother is (see note 18, above; Velleman,1982; Cleveland, _et al_.1978). It does, however, resist the effect of extremes with a graduated weighting system (Pierce,1980:131). 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.

In practice, we have found that X-11 results vary with the length of the series. For example, an analysis of the seven Index crimes in Illinois (Block,1984:3-5) found that the stable seasonality F values and the percent contributions of the seasonal and of the irregular differed with the addition of three years to the series. For several crimes, the F value increased with the longer series, but the percent contribution of the seasonal component decreased. One explanation for this apparent contradiction is that a longer series allows a more accurate description of seasonal activity. This more accurate description tells us that the seasonal contribution, in these crimes, is less.

Discontinuities

If there is an abrupt change or discontinuity in the series, no continuous method, component or ARIMA, will work. Smoothers such as a moving average are analytically continuous. They are defined in the same way throughout the series.[34] Therefore, X-11 cannot accurately describe discontinuities or abrupt changes in the direction of a series.

If you suspect that a series contains a discontinuity, Shiskin, et al. (1967:5) suggest that you break it into segments for analysis. Also, investigate the data source to determine whether there was a change in definition or data collection practices (see Block,1982:56-58).

Moving Seasonality

The X-11 program assumes that seasonal fluctuation follows a stable or gradually evolving pattern from year to year (see "Pattern Consistency," page 28 above). If there is a large amount of year-to-year change in the seasonal pattern, the series is said to contain "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. An added attraction of the X-11/ARIMA is a combined test for stable and moving seasonality (Dagum,1983a).

---

[34]Macaulay (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 (Block,1983:7,8,58,59).

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, divide 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, question the results of an adjustment in which the moving seasonality F value is significant.

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.[35] Between October, 1974 and January, 1975, the number of unemployed people in Illinois tripled (Block, et al. 1981). The seasonal factors (table 4) reflect this drastic change. The seasonal factor of many months changes in 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.

The distinction between gradual change (moving seasonality) and abrupt change (discontinuity) requires subjective interpretation and an intimate knowledge of the data source. In this Illinois unemployment example, we thoroughly investigated the source of the data to determine if some change in data collection practices had occurred, and finally concluded that there was no change in the definition of the data. Thus, the apparent discontinuity may represent a real increase in the number of unemployment insurance recipients. Such an hypothesis may be tested, using "time series intervention" methods (Glass, et al. 1975).

---

[35]Moving seasonality 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

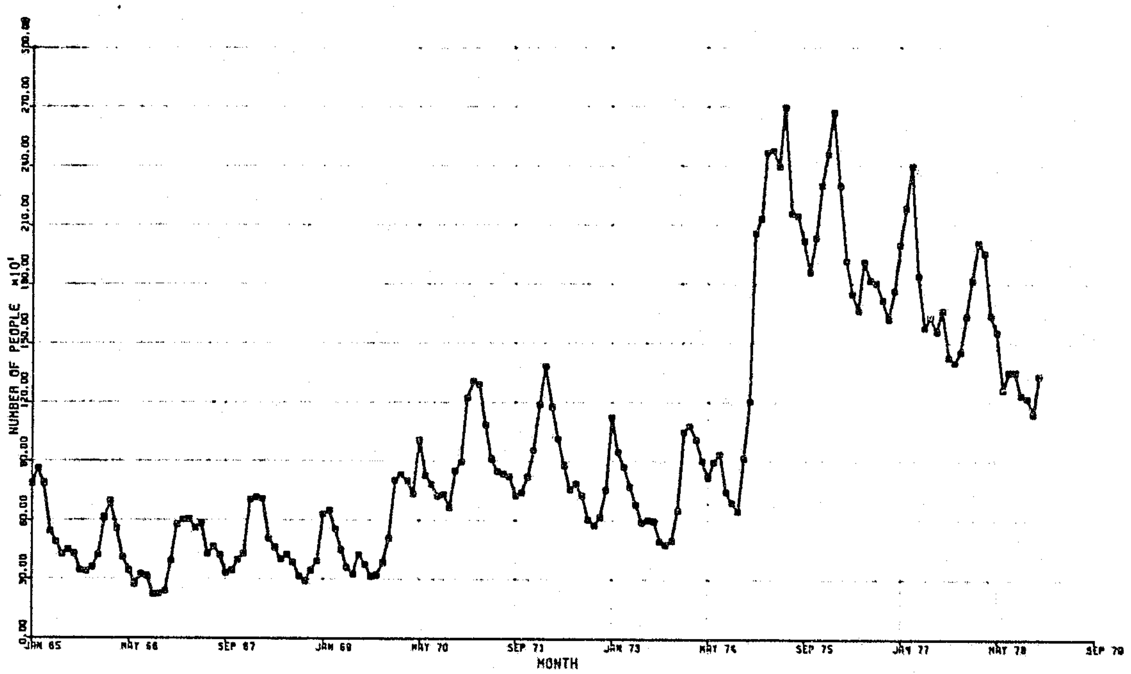


Table 4

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	-36	-71	-153	-164	-117	-52
1968	199	247	179	63	-15	-65	-40	-79	-161	-177	-117	-50
1969	210	248	188	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

ARIMA Methods

ARIMA time series models are a sophisticated way of forecasting future observations from past observations. Box and Jenkins (1970) suggested that most time series encountered in practice follow either (or a combination of) two types of stochastic processes: moving average or autoregressive.[36] 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 an ARIMA model for a series, the analyst must decide whether or not a seasonal process should be part of the model.

The method Box and Jenkins developed to identify the processes of a series uses trial and error ("iterative decisions"). The analyst begins with an initial diagnosis of the series. This diagnosis uses descriptions of the relationship between observations at one time period and another to discover any systematic movement. If some of this movement appears to be seasonal, the analyst considers a number of alternative seasonal processes that may account for the diagnostic results. Each set of alternative processes becomes a tentative model. The analyst then evaluates the "goodness of fit" of each tentative model by calculating the "residuals," the difference between the modeled values and the actual data. Eventually, the analyst reaches a model that appears to describe the series better than alternative models. This model may or may not contain a seasonal process.

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, ARIMA also relies on the subjective interpretation of a number of diagnostic tests. In this section, we explain the most important of these diagnostics.

Definitions

Moving Average and Autoregressive Processes

In a moving average (MA) process, the current observation is a function of a past error.[37] 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 related to the observations themselves. This happens in a moving average process.

---

[36] For a definition of "stochastic process," see page 16, above.

[37] A moving average process is not the same as the "moving average" of the component method. See note 19, above.

An MA(1) moving average process means that the current observation is affected by the error of the previous observation.[38] An MA(2) moving average process means that the current observation is affected by the error of the second previous 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.

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.[39] 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.

---

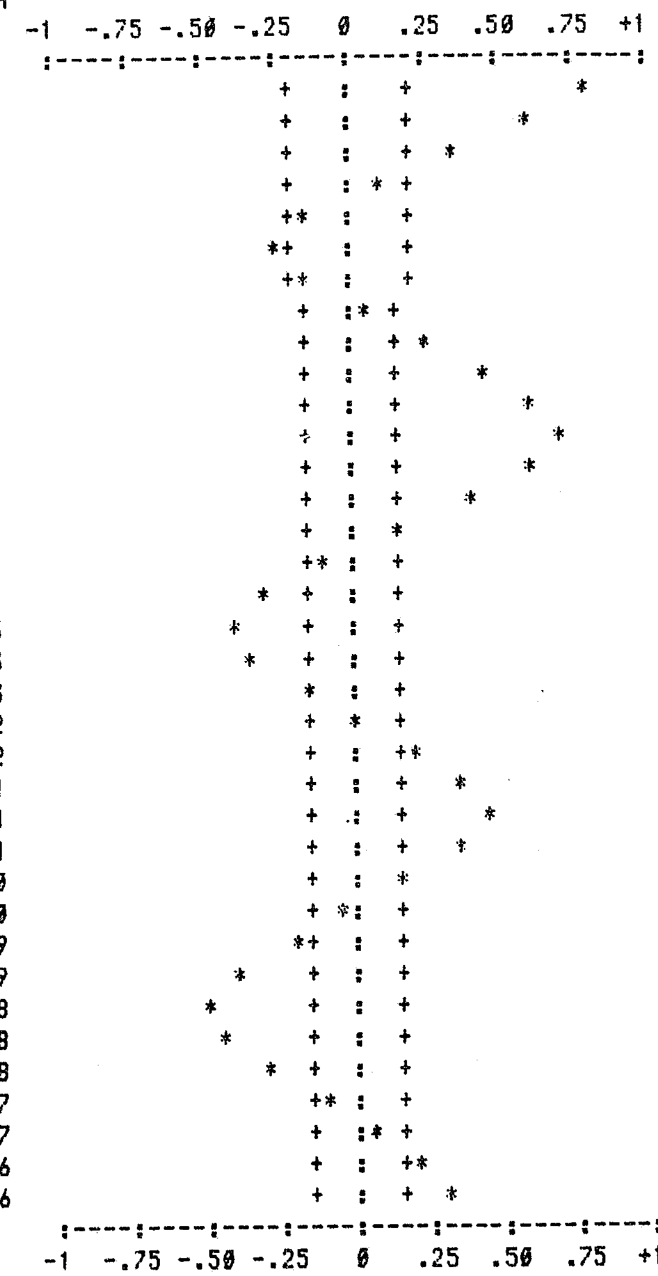
[38]The word "affected" here simply means correlated. Although the current observation may be predicted from the past, the past does not directly "cause" the current observation.

[39]Error 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 (1974,1975,1982) calls this the Durbin/ARIMA method.

Figure 12

Correlogram of Illinois Larceny/Theft

ORDER	AUTO-CORR.	S.E. RANDOM MODEL
1	.801	.090
2	.580	.090
3	.326	.089
4	.085	.089
5	-.140	.089
6	-.232	.088
7	-.153	.088
8	.036	.087
9	.229	.087
10	.450	.087
11	.617	.086
12	.708	.086
13	.595	.085
14	.407	.085
15	.166	.085
16	-.079	.084
17	-.290	.084
18	-.397	.083
19	-.341	.083
20	-.162	.083
21	.015	.082
22	.202	.082
23	.353	.081
24	.444	.081
25	.345	.081
26	.167	.080
27	-.041	.080
28	-.210	.079
29	-.409	.079
30	-.497	.078
31	-.436	.078
32	-.291	.078
33	-.124	.077
34	.072	.077
35	.218	.076
36	.306	.076



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

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) series, the current observation is correlated with the 12th 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 24th 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 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 12 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 12th-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.

The identification of the process of a series, especially combinations of AR and MA or serial and seasonal, can become quite complex. Correlograms provide clues, but are subject to varied interpretation. The partial correlogram (McCleary and Hay,1980:75-79;Nelson,1973:82-84) is helpful, and we have also found the Durbin/ARIMA technique (Roberts,1982) to be helpful.

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	AUTO-CORR.	S.E. RANDOM MODEL	-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	+ :* +

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

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

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 12 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 12 months apart are correlated. Figure 16 is the correlogram of the same series with the increasing trend removed. Observations 12 months apart are not correlated. Such a series is said to follow an "integrated" process. The "I" in ARIMA stands for integrated.

This emphasizes an additional complication of identifying an ARIMA model: the method we have described works only 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.[40] Because many series do show some trend, they may not, therefore, be analyzed by ARIMA 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 (seasonal trend). In such a case, each month is systematically higher (or lower) than the same month one year ago. 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 its graph.[41] Does the level of the series seem to increase or decrease over time? Second, look at a correlogram. In a series

---

[40]A stationary series not only has no change in level over time, but also has no change in variance from the beginning to the end of the series. We do not discuss this kind of seasonal lack of stationarity here, 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. The "ladder of powers" (Velleman and Hoaglin,1981: 48-50) is a useful tool for the novice in determining the expected effect of various power transformations. Also see Nelson (1976) or McCleary and Hay (1981).

[41]It 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 15

Correlogram, Canadian Homicide: 1961-1980

ORDER	AUTO-CORR.	S.E. RANDOM MODEL
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

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

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

Figure 16

Correlogram, First Difference
Canadian Homicide: 1961-1980

ORDER	AUTO-CORR.	S.E. RANDOM MODEL
1	-.510	.064
2	.040	.064
3	-.024	.064
4	-.022	.064
5	.014	.064
6	.034	.064
7	-.009	.063
8	-.061	.063
9	.018	.063
10	.039	.063
11	.014	.063
12	-.036	.063
13	.079	.063
14	-.173	.063
15	.117	.062
16	.068	.062
17	-.150	.062
18	.136	.062
19	-.111	.062
20	.054	.062
21	-.105	.062
22	.115	.061
23	.027	.061
24	-.076	.061
25	.034	.061
26	.056	.061
27	-.119	.061
28	.111	.061
29	-.113	.060
30	.065	.060
31	.010	.060
32	-.028	.060
33	.046	.060
34	-.028	.060
35	.033	.060
36	-.094	.059

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

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

with either serial or seasonal 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 "Identifying The Process of a Series," above). 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. Be very cautious, however, in seasonal decomposition of a series that has been log transformed. Dagum (1981:133) has shown that, in an additive assumption, "logarithmic transformation destroys the underlying properties of the series and that the ARIMA models adequate for the non-transformed data are no longer applicable."

Differencing

Differencing is a transformation intended to produce a series that is stationary in level. As discussed in the section just above, a series that is not stationary in level cannot be modeled with ARIMA methods. Unfortunately, there is no "cookbook" test to determine whether or not a series is stationary (see Rauma,1981;Blumstein,et al.,1981). Autocorrelations that do not die out with increasing lag, such as the correlogram in figure 15 above, should make you suspect a non-stationary series. However, ultimately, you must use trial and error. Transform a series to remove trend; then analyze the transformed series.

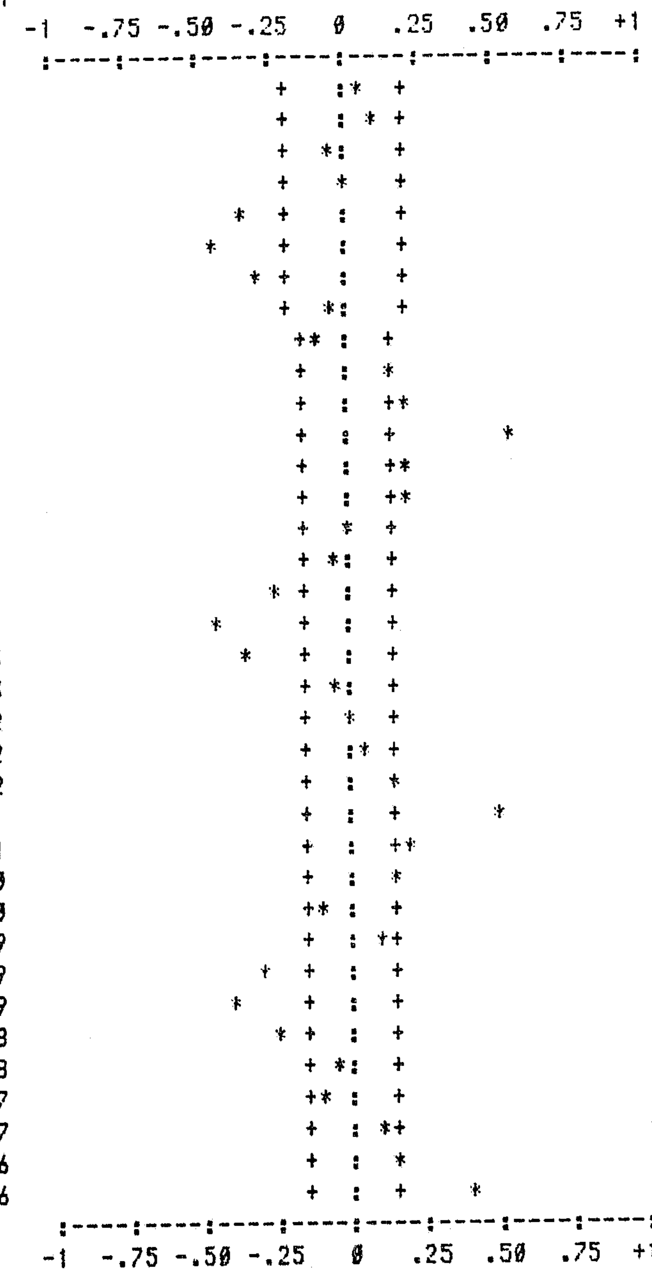
An overall trend can usually be removed by a first difference; a seasonal trend can usually be removed by a 12th difference.[42] In a first difference, each observation is subtracted from the following observation. In a 12th difference, each observation is subtracted from the observation 12 months away. 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 12th difference. If a series has both a serial and a seasonal trend, you would transform it into a stationary series by taking a 12th difference of the first difference.

[42] Some series require two successive first difference transformations 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 17

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

ORDER	AUTO-CORR.	S.E. RANDOM 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

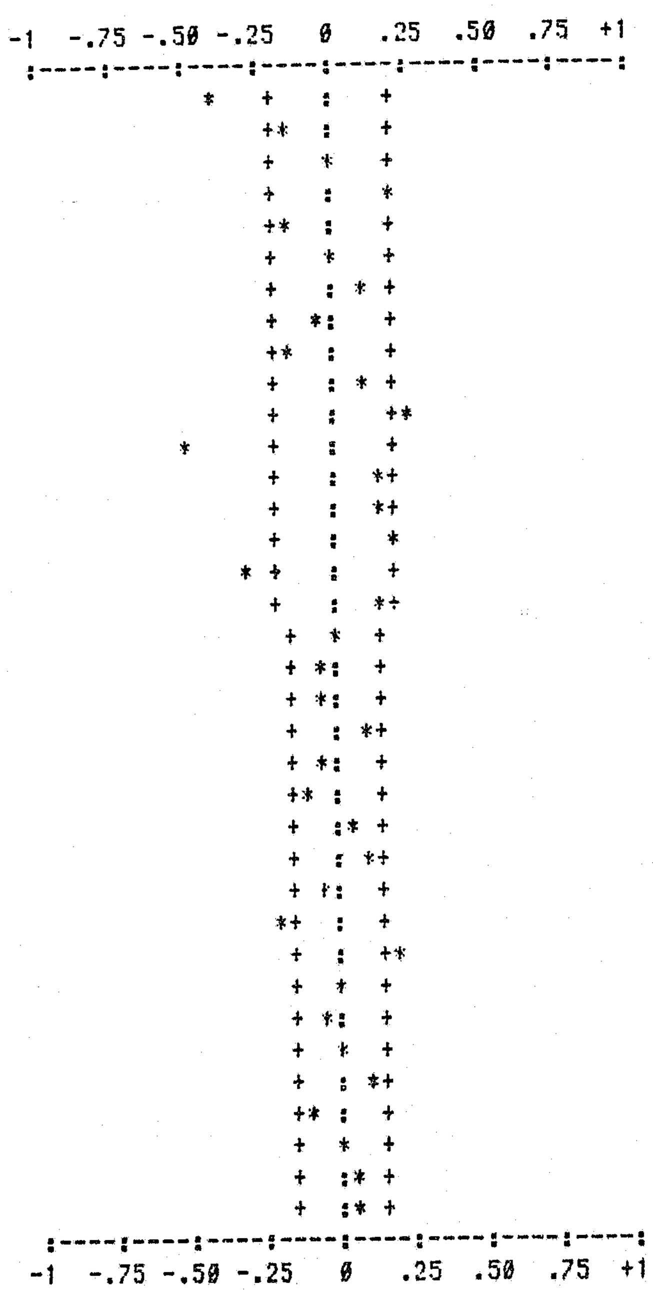


* : AUTOCORRELATIONS
+ : 2 STANDARD ERROR LIMITS (APPROX.)

Figure 18

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

ORDER	AUTO-CORR.	S.E. RANDOM MODEL
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.)

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 12th 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 12 months away. Compare these correlograms to the correlogram of the original series (figure 12, above).

Although differencing may be necessary to produce a stationary series, it may also produce some problems.[43] One possible 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 12th differencing. We have occasionally found this to produce confusing results. For example, if the initial year of a series contains more (or less) trend or more (or less) seasonality than the rest of the series, a 12th difference, by eliminating the initial year, would change the diagnostic results for the series. This is one of the many reasons that, before attempting to estimate a model, it is wise to inspect the pattern of the raw data (see Block, 1983).

Another drawback of differencing is the danger of overdifferencing (also called overadjustment; see page 7 and references in note 5, above). For example, removing the trend from the Canadian homicide series (figure 15) produces a transformed series that contains <u>negative</u> autocorrelation between each observation and the next observation (figure 16). We have encountered this phenomenon with some frequency in crime series. Dagum (1981:135) notes that, "Excessive application of the difference operator, to generate a process stationary in the differences, induces a non-invertible moving average process in the residuals."[44] In other words, the correlogram of an overdifferenced series may mimic the correlogram of a moving average process. Dagum's X-11/ARIMA package rejects a moving average ARIMA model if the weight of the moving average process equals .90 or greater.[45] This is a good rule of thumb to use in general situations.

---

[43] If the original data are random, then the first differences will be random also. However, if the original data are aggregated over time (for example, monthly data cumulated to quarterly or yearly data), then the first difference of the aggregate data will <u>not</u> be random. See Kendall (1976:6-7).

[44] "Non-invertible" refers to one of the statistical requirements of moving average processes, equivalent to the requirement of stationarity for autoregressive processes. The weight of a moving average term in an ARIMA model must be less than 1. See Nelson (1973:46-48).

[45] In other words, the X-11/ARIMA assumes overdifferencing in a model in which the current observation is related more strongly than .90 to the error of the previous observation.

Despite these drawbacks, if your series is not stationary, analyzing a difference transformation of it may be your only alternative for identifying an ARIMA model. You may have to choose between several alternatives (overdifferencing versus not identifying a model), none of which is entirely satisfactory.

Rules of Thumb for Evaluating a Model

After first obtaining a stationary transformation, or determining that the original series is already stationary, and identifying a tentative model, we must evaluate this model. For example, for Illinois larceny/theft, the first-order and 12th-order autocorrelations of the transformed series (figure 18, above) are significant, but the second order 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, we modeled larceny/theft by applying MA(1) and MA(12) processes to the differenced series. This produced the modeled series graphed in figure 10 (page 17, above). We now must determine whether or not this model accurately describes the stochastic process of the series.

One way to evaluate a model is to analyze the residuals, the discrepancy between the modeled values and the actual series. Residuals of a good model vary randomly over time.[46] This section explains two tools for analyzing the residuals of a model -- the correlogram, which has been introduced above, and a new tool, the cumulative periodogram.

This section is only a quick overview of model evaluation. For more detail, see Nelson (1973) or McCleary and Hay (1980). In addition, we have found the criteria developed by Statistics Canada (Dagum 1981,1983a; Lothian and Morry,1975) to be useful in model evaluation.

Correlogram of Residuals

The correlogram of the residuals of an MA(1) and MA(12) model for Illinois larceny/theft (figure 19) does indeed indicate a random pattern.[47] 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 12th differencing (figure 18). Clearly, the residuals look most like a random series.

---

[46]This random variation in a time series is sometimes called "white noise."

[47]Two of the 36 autocorrelations in figure 19 are slightly larger than two standard deviations, but a small percentage of autocorrelations might be expected to be significant, only by chance.

Figure 19

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

ORDER	AUTO-CORR.	S.E. RANDOM MODEL	-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	+ *: +

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

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

If inspecting a correlogram of residuals seems too subjective, there is an objective criterion for randomness in a set of sample autocorrelations. This is the "Box-Pierce" statistic, also known as the "Q" statistic (Box and Pierce, 1970; Nelson, 1973:115). ARIMA programs such as IDA calculate the Box-Pierce statistic whenever the user requests a correlogram. It is distributed approximately as chi-square, with degrees of freedom equal to the number of lags in the correlogram minus the number of autoregressive and/or moving average processes in the model.

However, like component analysis, ARIMA 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 12th differencing (figure 18) has negative autocorrelation at lag 12. One interpretation of this is that it suggests the model we have identified, 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 12th 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. The decision as to which alternative is best is subjective. Therefore, analysts should be careful to mention their decisions, and the diagnostics that led to these decisions, in published results.

Cumulative Periodogram of Residuals

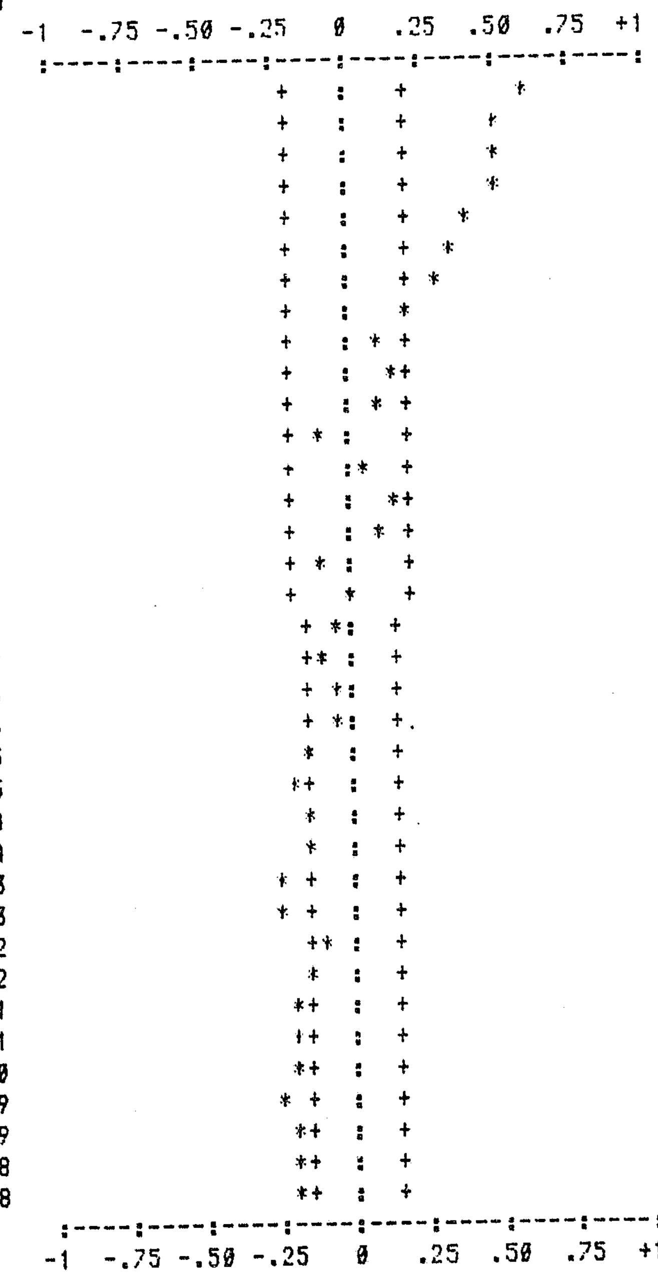
The cumulative periodogram is also very useful in evaluating a tentative model, especially when the series may contain seasonal fluctuation.[48] A cumulative periodogram gives you the

[48]For more detail about analysis of series in the frequency domain, see Rosenblatt (1965:1-2) or Glass, et al.(1975:205-216). A periodogram is a tool for analyzing of the spectrum or the harmonics of a series. It 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 analysis 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." A periodogram measures the intensity of the spectrum at a certain frequency. The "normalized cumulative periodogram" (Box and Jenkins, 1976:295) is a good tool for detecting periodic patterns in the residuals of a model.

Figure 20

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

ORDER	AUTO-CORR.	S.E. RANDOM MODEL
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.)

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.

Box and Jenkins (1976:294) recommend a periodogram analysis in preference to a correlogram for 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):

> . . . 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.

For example, figure 21 shows two cumulative periodograms side by side for comparison. The periodogram for the original larceny/theft data indicates two distinct departures from linearity. The jump at an interval of 2 is often seen in non-stationary series. The jump at an approximate period of 11.9 months is often seen in a seasonal series. The graph of the residuals of the MA(1) MA(12) model, on the other hand, does not indicate any significant periodicity. The cumulative relative sum of periodogram (asterisks) moves in a well-behaved fashion, well within the 5 percent confidence limits (crosses).

Appropriate Applications

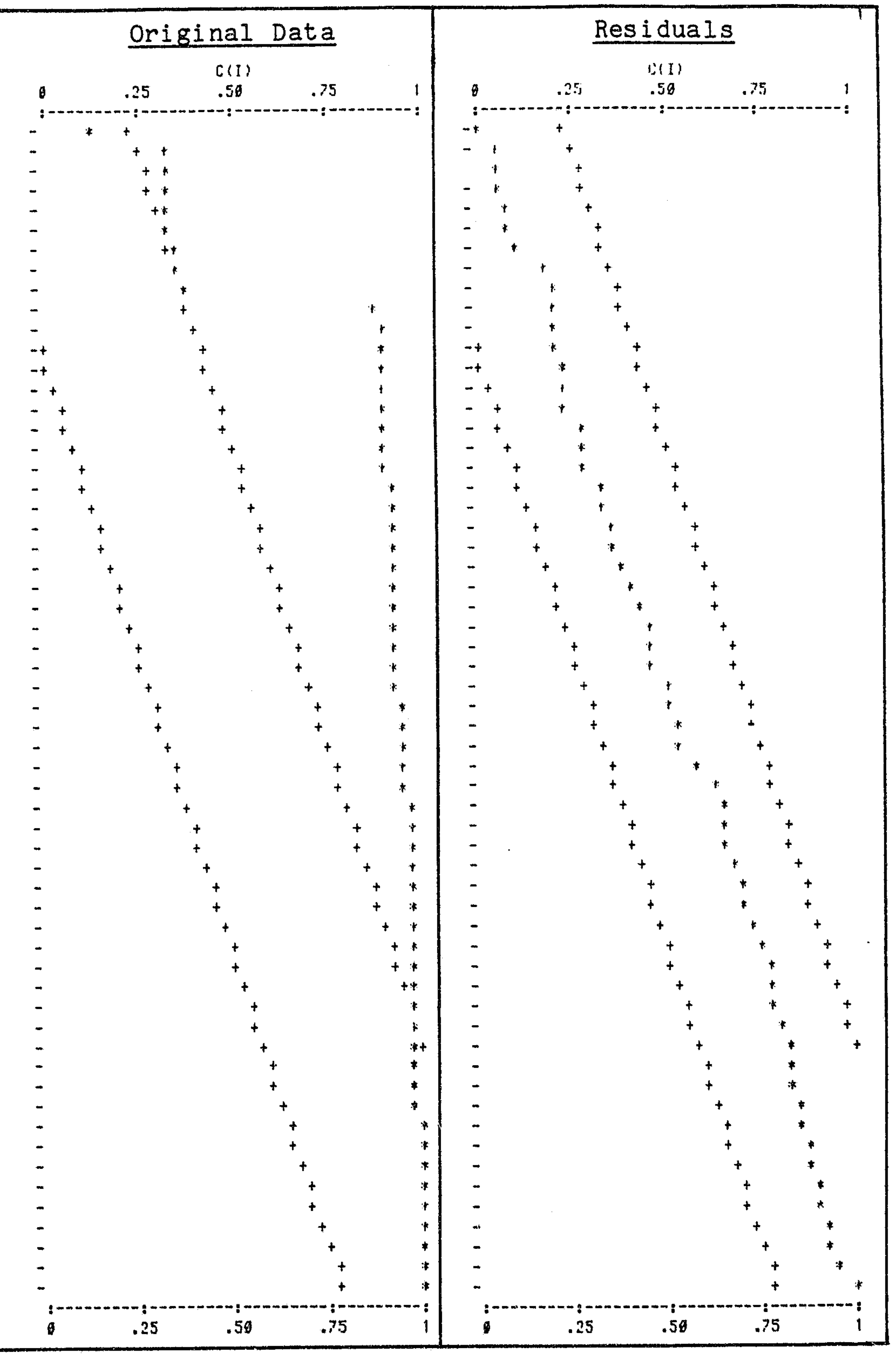
Obviously, the combinations of moving average processes and autoregressive processes, serial processes and seasonal processes, can become quite complicated. Identifying the ARIMA model that best describes a series is not entirely objective, nor is it simple for an analyst to state these subjective decisions in a published report. It is not uncommon for two statisticians to identify different models for the same series, even though they use the same methods. As Pierce (1980:130) argues, "Theoretically incompatible models can produce results uncomfortably close to each other and uncomfortably far from the truth."

In contrast to the X-11, 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, ARIMA methods require a lengthy analysis and re-analysis of each individual series. Therefore, they are most appropriate for analyzing one or two important series, rather than as the standard method of analyzing all of an agency's data (see Kuiper,1978: 59-60).

Figure 21

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

I	FREQ	APPROX PERIOD
1	.0084	119.0
2	.0168	59.5
3	.0252	39.7
4	.0336	29.8
5	.0420	23.8
6	.0504	19.8
7	.0588	17.0
8	.0672	14.9
9	.0756	13.2
10	.0840	11.9
11	.0924	10.8
12	.1008	9.9
13	.1092	9.2
14	.1176	8.5
15	.1261	7.9
16	.1345	7.4
17	.1429	7.0
18	.1513	6.6
19	.1597	6.3
20	.1681	6.0
21	.1765	5.7
22	.1849	5.4
23	.1933	5.2
24	.2017	5.0
25	.2101	4.8
26	.2185	4.6
27	.2269	4.4
28	.2353	4.3
29	.2437	4.1
30	.2521	4.0
31	.2605	3.8
32	.2689	3.7
33	.2773	3.6
34	.2857	3.5
35	.2941	3.4
36	.3025	3.3
37	.3109	3.2
38	.3193	3.1
39	.3277	3.1
40	.3361	3.0
41	.3445	2.9
42	.3529	2.8
43	.3613	2.8
44	.3697	2.7
45	.3782	2.6
46	.3866	2.6
47	.3950	2.5
48	.4034	2.5
49	.4118	2.4
50	.4202	2.4
51	.4286	2.3
52	.4370	2.3
53	.4454	2.2
54	.4538	2.2
55	.4622	2.2
56	.4706	2.1
57	.4790	2.1
58	.4874	2.1
59	.4958	2.0



+ : KOLMOGOROV-SMIRNOV TEST LIMITS AT 5% LEVEL
* : CUMULATIVE RELATIVE SUM OF PERIODOGRAM

As discussed above (page 31), neither ARIMA nor component methods are appropriate for highly irregular series, short series (six or fewer years), or series containing an abrupt change or discontinuity.

Extremes

ARIMA 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, ARIMA can, of course, be used if the series is first transformed to remove or re-weight the extremes.

Series Length

A general rule of thumb in ARIMA 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 processes, even more observations are necessary. Also, keep in mind that, if a 12th difference is necessary to make the series stationary, 12 observations will be lost.

Discontinuities

An ARIMA model, like a component model, is analytically continuous (see note 34, above). 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). In other words, an ARIMA model 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 continuous model is not appropriate. If you suspect that this is the case, first inspect the 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 model to a series containing a discontinuity.

Moving Seasonality

ARIMA methods are particularly helpful for describing series that contain moving seasonality (see page 32, above). The ARIMA 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. However, as discussed above (page 33), the distinction between moving seasonality and discontinuity is not always easy to make. ARIMA methods can handle the first, but not the second.

ANNOTATED BIBLIOGRAPHY

Preceding page blank

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.

Armstrong, J. Scott
1978 Forecasting with econometric methods: Folklore versus Fact. Journal of Business 51(4):549-564. Responses by Chow, Kosobud, McNees, Miller, Wecker, and Zellner, pp 565-600, same issue.
One of the first serious criticisms of econometric forecasting.

Ascher, William
1978 Forecasting: An Appraisal for Policy-Makers and Planners. Baltimore: Johns Hopkins University Press.

Banks, Jerry and David Vatz
1976 Sinusoidal pattern analysis in criminal justice. Criminology 14(2):251-258.
Uses a multiple regression with trend and sine-cosine components. For an introduction to spectral analysis, see Rosenblatt (1965).

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.
A review of spectral analysis, with emphasis on robust and resistant techniques.

Bell, W. R. and S. C. Hillmer
1983 Modeling time series with calendar variation. Journal of the American Statistical Association 78(383):526-534.
Combination of regression and ARIMA models handles trading-day and Easter holiday variations.

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 (1981).

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.
Seasonal analysis of Index robbery and Index burglary in 77 Illinois law enforcement jurisdictions.
1983 Manual for the Pattern Description of Time Series: Guide to Pattern Description. Statistical Analysis Center, Illinois Criminal Justice Information Authority. Revised from 1982 edition.
Presents a simple method, easy to communicate to a general audience, of describing patterns of change over time. Also see Miller (1982).
1984 Is Crime Seasonal? Statistical Analysis Center, Illinois Criminal Justice Information Authority.
Contains a discussion of issues particularly relevant to analysis of seasonal fluctuation in crime, a review of research literature on seasonality of crime, and the results of seasonal analysis of 135 Index crime series.

Block, Carolyn Rebecca and Richard L. Block
1980 Patterns of Change in Chicago Homicide: The Twenties, The Sixties, and The Seventies. Statistical Analysis Center, Illinois Law Enforcement Commission.

Block, Carolyn Rebecca, Craig McKie and Louise S. Miller
1983 Patterns of change over time in Canadian and United States Homicide. Policy Perspectives 3(2):121-180.

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.

Blumstein, Alfred, Jaqueline Cohen, Soumyo Moitra, and Daniel Nagin
1981 On testing the stability of punishment hypothesis: a reply. The Journal of Criminal Law & Criminology 72(4):1799-1808.
Reply to Rauma (1981). Discusses the issue of stationarity.

Box, George E.P. and Gwilym M. Jenkins
1976 Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
The classic treatment of ARIMA. This is a revision of the first edition, which was published in 1970.

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.

Box, George E.P. and David A. Pierce
1970 Distribution of residual autocorrelations in auto-regressive-integrated moving average time series models. Journal of the American Statistical Association 65(332):509-526.
Introduces the "Box-Pierce" statistic, which has become a standard criterion for randomness in a correlogram of residuals.

Box, George E.P. and George C. Tiao
1965 A change in level of a non-stationary time-series. Biometrika 52:181-192.

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. For a comparison of SABL and X-11, see Levenbach and Cleary (1981). For general theory of nonlinear smoothing, see Emerson and Hoaglin (1983) and Velleman (1980,1982).

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.
1981 Diagnostic checks for the ARIMA models of the X-11-ARIMA seasonal adjustment method. Pp. 133-145 in O.D. Anderson and M.R. Perryman (eds.), Time Series Analysis. Amsterdam: North-Holland Publishing Company.
Discusses the calculation of, and the supporting theory for, the three criteria for ARIMA models built in to the X-11/ARIMA package. Also discusses the effect of log transformation on the models. See Lothian and Morry (1978).
1983 The X-11-ARIMA Seasonal Adjustment Method. Seasonal Adjustment and Time Series Staff, Statistics Canada, Ottawa, K1A OT6.
Dagum has done, or inspired others to do, much of the advanced work in seasonal adjustment today, including this X-11/ARIMA method.
1983 Seasonality. Forthcoming in Encyclopedia of Statistical Sciences, S. Katz and N. Johnson (eds.), vol. 6.
1983 Moving averages. Forthcoming in Encyclopedia of Statistical Sciences, S. Katz and N. Johnson (eds.), vol. 6.

Deutsch, Stuart Jay
1978 Stochastic models of crime rates. International Journal of Comparative and Applied Criminal Justice 2(2):127-151.
Builds ARIMA models for seven Index crimes in each of 10 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, and 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 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.

Emerson, John D. and David C. Hoaglin
1983 Resistant lines for y versus x. Pp.129-165 in Understanding Robust and Exploratory Data Analysis, Hoaglin, Mosteller and Tukey (eds.). New York: John Wiley & Sons, Inc.
Review of the theory and method of resistant lines, including partitioning the series into three groups by various methods, and alternative approaches such as repeated medians. Also see Velleman (1980,1982).

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), and Granger (1978).

Hamming, R. W.
1977 Digital Filters. Englewood Cliffs, New Jersey: Prentice-Hall. (Second edition; first edition 1977.)
This is the classic reference for spectral analysis. See discussion of the "Gibbs" phenomenon, which is an argument against moving averages.

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 an ARIMA 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 Makridakis, et al. (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, autoprojection, ARIMA, 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 describes a forecasting competition by Reid (also see Makridakis, et al.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 Publishing 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 Makridakis, et al. (1982) and 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.
1979 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
1978 Exploratory Data Analysis: An Introduction to Selected Methods. Pp. 311-372 in Sociological Methodology 1979, Karl F. Schuessler (ed.).
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.

Levenbach, Hans and James P. Cleary
1981 The Beginning Forecaster: The Forecasting Process through Data Analysis. Belmont, California: Wadsworth.
A good introduction to component methods. Contains a lot of information on SABL (Cleveland, et al. 1978), including a comparison of SABL and X-11. Although it does not cover ARIMA, a companion volume, The Professional Forecaster, does.

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. Developed by the University of Chicago Graduate School of Business. The ARIMA analyses in this report were done on IDA. Also see Roberts.

Lothian, J. and M. Morry
1978 Selection of models for the automated X-11-ARIMA seasonal adjustment program. Seasonal Adjustment and Time Series Analysis Staff, Statistics Canada.
Reviews the analysis of 175 15-year economic series that provided the basis for choosing the three ARIMA models to be built in to the X-11/ARIMA program.

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 Makridakis, et al. (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.

Makridakis, Spyros, A. Anderson, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen and R. Winkler
1982 The accuracy of extrapolation (time series) methods: results of a forecasting competition. Journal of Forecasting 1:111-153.
Describes a forecasting competition, performed on 1,001 time series by seven experts using 24 alternative methods, to forecast for six to 18 time periods. Expands and enlarges on Makridakis and Hibon (1979). For other method comparisons, see Kendall (1976), Hibbs (1977), Willson (1973), and Kuiper (1978).

Makridakis, Spyros and Michele Hibon
1979 Accuracy of forecasting: An empirical investigation. Journal of the Royal Statistical Society A 142, part 2:97-145.
Concludes that, "(a) Judgmental approaches are not necessarily more accurate than objective methods; (b) Causal or explanatory methods are not necessarily more accurate than extrapolative methods; and (c) More complex or statistically sophisticated methods are not necessarily more accurate than simpler methods."

Makridakis, Spyros and Steven C. Wheelwright
1978 Forecasting: Methods and Applications. Santa Barbara: John Wiley and Sons.
A basic forecasting textbook.

Mallows, C.L.
1980 Some theory of nonlinear smoothers. The Annals of Statistics 8(4):695-715.

Marshall, Clifford W.
1977a Application of Time Series Methodology to Crime Analysis. The Polytechnic Institute, 33 Jay St., Brooklyn, NY 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.

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 seasonal ARIMA 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 ARIMA methods.

McIntire, Robert J.
1983 Comments on "How to Handle Seasonality." Bureau of Labor Statistics, letter, June 30, 1983.

Miller, Louise S.
1982 Manual for the Pattern Description of Time Series: Technical Manual. Statistical Analysis Center, Illinois Criminal Justice Information Authority.
See Block (1983).

Munk, W.H., G.R. Miller, F.E. Snodgrass and N.F. Barber
1962 Directional recording of swell from distant storms. Journal of the Royal Statistical Society A 255:62-583.
Elementary treatment of spectral analysis.

Nelson, Charles R.
1973 Applied Time Series Analysis. San Francisco: Holden-Day, Inc.
With McCleary and Hay (1980), this is an excellent introduction to ARIMA methods.

Nettheim, Nigel F.
1965 A Spectral Study of "Overadjustment" for Seasonality. U.S. Department of Commerce, Bureau of the Census. Working Paper No. 21.
Concludes that, "A situation in which the best procedure is to overadjust and then to correct for the overadjustment is unlikely to be a final resting place."

Pfeifer, Phillip E. and Stuart Jay Deutsch
1980 Identification and interpretation of first order space-time ARIMA models. Technometrics 22(August,3):397-408.
An extension of ARIMA 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 percent 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).

Priestley, M.B.
1981 Spectral Analysis and Time Series. Vol. 1, Univariate Series; Vol. 2, Multivariate Series. London: Academic Press.
Provides an overview. For periodogram analysis, see pages 394-397 of volume 1. Also see Hamming (1977) and Rosenblatt (1965).

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).

Rauma, David
1981 Crime and punishment reconsidered: Some comments on Blumstein's stability of punishment hypothesis. The Journal of Criminal Law & Criminology 72(4):1772-1798. Includes a review of the problem of determining whether or not a series is stationary. Concludes, "In general, the existing tests for stationarity all require some specification of the form that the possible nonstationarity takes" (p. 1779). Also see the reply (Blumstein, et al. 1981), which also focuses on stationarity.

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).
1982 Data Analysis for Managers. Manuscript.
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. For a more detailed text, see Hamming (1977) or Priestly (1981). Also see Munk, et al. (1962), Beaton and Tukey (1973), Nettheim (1965), Bliss (1958,1970).

SAS Institute, Inc.
1982 SAS/ETS User's Guide, 1982 Edition. Cary, N.C.: SAS Institute, Inc.

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 canceling the actual crime report that was unfounded. The effect of this would be to reduce the intensity of any seasonal fluctuation. Also see US/BJS (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-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 percent 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.
Intervention analysis for short series. See Hartmann, et al. (1980).

Velleman, Paul F.
1980 Definition and comparison of robust nonlinear data smoothing algorithms. Journal of the American Statistical Association 75(371):609-615.
1982 Applied nonlinear smoothing. Pp. 141-177 in Sociological Methodology 1982, Samuel Leinhardt (ed.).
This is a review of resistant smoothers, an alternative to the moving average. Also see Tukey (1977), Mallows (1980), Velleman and Hoaglin (1981), Beaton and Tukey (1974).

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). It includes a discussion of resistant time series analysis methods. Also see Tukey (1977), Velleman (1982), Cleveland, et al. (1978), Emerson and Hoaglin (1983).

Vigderhous, Gideon
1977 Forecasting sociological phenomena: Application of Box-Jenkins methodology to suicide rates. Pp. 20-51 in Sociological Methodology 1978.
Good overview of ARIMA 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