ILLINOIS
CRIMINAL JUSTICE
INFORMATION AUTHORITY

MANUAL FOR THE PATTERN DESCRIPTION OF TIME SERIES

Part I
GUIDE TO PATTERN DESCRIPTION

September, 1982
Second edition, July 1983

by Carolyn Rebecca Block
and Louise S. Miller
with the assistance of Doug Hudson

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

Printed by authority of the State of Illinois
July 1983
Number of copies: 200
Printing order number: 84-01

ACKNOWLEDGEMENTS AND A SHORT HISTORY

The time series pattern description method, to which this manual is a guide, was built upon a foundation laid by others. Hudson (1966) developed, and Fox (1978) wrote a computer algorithm for, the two segment method we adopted. Ertel and Fowlkes (1976) developed and wrote an algorithm for an n segment spline method. We are very grateful to Edward B. Fowlkes of Bell Laboratories, who helped us adapt the program to our system.

Like all Statistical Analysis Center (SAC) publications, the Pattern Description Manual is the result of the cooperation of a number of SAC staff members. Carolyn Rebecca Block (Becky) used the method for the first time in 1979 in "Descriptive Time Series Analysis for Criminal Justice Decision Makers: Local Illinois Robbery and Burglary." For that study, Jean Fearington adapted Fox's program to our system, and L. Edward Day wrote the initial version, which we still use, of the graphics component. "Patterns of Change in Chicago Homicide: The Twenties, The Sixties, and The Seventies" expanded the concept of pattern description to include time series specification, and compared the descriptive results to stochastic time series analysis results.

With the Ertel/Fowlkes program, we were able to expand pattern description beyond the two segment Hudson/Fox capability. Becky and Doug Hudson worked together closely to adapt the program to our system, and Doug added a number of improvements to the SAC graphics package, many of which are described in this manual. The study, "Explaining Patterns of Change Over Time in Chicago Homicide with a Gun," provided a practical test of the new package, and occasioned further expansion and definition of the concept. Louise S. Miller produced the graphs for that analysis, and later used her experience to answer requests for hundreds of pattern descriptions, and to suggest a number of improvements in the package. Louise was chiefly responsible for Part II of the Pattern Description Manual, the "Technical Manual."

A grant from the Bureau of Labor Statistics, "Time Series Pattern Description for Criminal Justice Decision Makers," is currently supporting the transfer of the Pattern Description package to agencies throughout the United States.

TABLE OF CONTENTS

Part I: Guide to Pattern Description

Acknowledgements and a Short History i
Table of Contents iii
Figures v
Introduction 1

The Concept of Pattern Description 3
  Criteria: Simplicity and Accuracy 3
  Pattern Description by Spline Regression 5
  Appropriate Applications 7

The Method of Pattern Description 13
  The Line Segment Package 13
    Hudson/Fox Two Segment Lines 14
    Ertel/Fowlkes n Segment Lines 15
  Choosing the Best Pattern Description 21
    Indicators of Accuracy 22
      C_p Plot 23
      The F-Test 25
    Indicators of Simplicity 26
      Length and Number of Segments 26
      Difference between Neighboring Segments 27

Examples of Pattern Description 31
  A Series Containing Seasonal Fluctuation 32
  Initial Search versus Final Search Criteria 34
  Comparing Patterns of Two Series 38
  Time Series Specification 40
  Simplicity versus Accuracy 42
  Fewer Segments may be More Accurate 44
  Short Line Segment 48
  Rates versus Raw Numbers 50
  Series with Extreme Values 52
  Long and Short Series 54
    Long Series 54
    Short Series 55
  Presence of a Possible Discontinuity 56

Bibliography 59

Part II: Technical Manual for Pattern Description

Acknowledgements and a Short History........................ i
Note to the User...iii
Table of Contents.. v
Figures.. vi

Introduction... 1

How to Produce Ertel/Fowlkes Plots....................... 5
  Using the Ertel/Fowlkes Program........................ 5
  The Complete Calculation Output........................ 6

How to Produce a Hudson/Fox Graph........................ 9
  Introduction... 9
  Using the Hudson/Fox Program........................... 11
  The Complete Calculation Output........................ 13

Appendix... 24
  Error Messages and Other Problems...................... 24
  How to Graph Series With Missing Data.................. 30
  How to Edit Files for Pattern Description.............. 33
  TI30... 33
  COMBINE.. 35
  Combining Separate Files............................... 39

Bibliography... 41

FIGURES

Part I: Guide to Pattern Description

1. Example of a Pattern Description: Illinois Felony Commitments from Court, 1965 to 1981 5

2. Pattern Description of a Seasonal Series: Illinois Index Aggravated Assault Offenses, 1972 to 1981 10

3. Example of a Hudson/Fox Pattern Description: Illinois Felony Commitments from Court, 1965 to 1981 15

4. Example of Complete Ertel/Fowlkes Graphic Output 18

5. Different Minimum Segment Lengths for the Same Series: Chicago Homicides without a Gun 28

6. A Typical Seasonal Series: Illinois Index Burglary 1972 to 1981 33

7. Different Initial and Final Search Criteria: Index Larceny/Theft in Quincy, Illinois 35

8. Comparing Patterns of Two Series: Index Burglary Offenses 1972 to 1981, Champaign and Danville, Illinois 39

9. Example of Time Series Specification: Chicago Homicides with a Gun and without a Gun 41

10. Simplicity versus Accuracy: Alternative Pattern Descriptions of Homicide Rates in the East South Central States. 43

11. Homicide Rates in the United States Metropolitan Counties, 1950 to 1977, Five Segment and Four Segment Descriptions. 45

12. Canadian Homicides Attributed to a Stranger, 1961 to 1980, Four Segment and Three Segment Descriptions 46

13. C_p Plots for United States Metropolitan County Homicides and Canadian Stranger Homicides 47

14. People Released, Illinois Department of Corrections: Six Segment and Five Segment Alternative Descriptions 49

15. Patterns of Rates versus Patterns of Raw Numbers: Arrests and Arrest Rates in Racine, Wisconsin 51

16. An Extreme that Revealed a Data Definition Problem: Lost and Stolen Firearms, Illinois except Chicago 52

17. An Extreme that Actually Occurred: Chicago Homicides 53

18. Yearly Patterns in a Long Series: Releases from the Illinois Department of Corrections 54

19. Patterns in a Monthly Series: Releases from the Illinois Department of Corrections 55

20. A Short Series: California Homicides Attributed to an Acquaintance 56

21. An Apparent Discontinuity: People Receiving Unemployment Insurance Benefits in Illinois 57

Part I

GUIDE TO PATTERN DESCRIPTION

INTRODUCTION

Time series pattern description provides concrete and readily understood answers to simple descriptive questions about the general pattern of change over time in a variable. It tells the user, in nonstatistical language, whether the variable generally increased, decreased or stayed the same during the period in question; whether there was a change in the pattern (for example from an increase to a decrease); and if there was a change, roughly when the change occurred.

For the first time, a substantial number of criminal justice time series data sets are available. Researchers, criminal justice administrators, and policy makers often ask the Illinois Statistical Analysis Center (SAC) for a simple description of the pattern of change over time in one of these data sets. We find that conventional methods of time series analysis cannot answer many of these simple descriptive questions. In response to this situation, SAC developed the time series pattern description method.

Time series pattern description consists of the conceptual method and a package of computer programs that enables analysts to use the method. SAC is continuing to test both the concept and the package in practical criminal justice situations, and to improve them based on that experience. This manual, therefore, is a working draft, distributed for use, testing and comments. We hope that people who have had various degrees of experience with data analysis will use pattern description in a variety of practical situations, to produce reports addressed to a variety of audiences. We encourage users to tell the authors of any problems they encounter, and welcome their comments and suggestions for improvement.

The manual consists of two parts. Part I, "Guide to Pattern Description," is an introduction to the concept of pattern description and a guide to applying that concept in practice. It includes detailed instructions for interpreting a pattern description graph, and illustrates the instructions with many examples of real criminal justice applications. Part II, "Technical Manual for Pattern Description," tells the user how to produce pattern description graphs on the Illinois Criminal Justice Information Authority computer system.

THE CONCEPT OF PATTERN DESCRIPTION

Although a time series may be complex, not every research or administrative decision based on that time series calls for an equally complex method of analysis. Many decisions require only a description of the overall pattern of a series. Using a more complex or more abstract method than the decision at hand requires is time consuming and expensive, may not answer the simple descriptive questions an analyst needs to know, and produces results that may be difficult to explain to a general audience.

Simple time series pattern description is useful for describing and exploring the general pattern of change over time in a data set. Such a general description may be all that is needed for many practical applications. If the problem at hand requires an explanatory model, forecast, or other exact explanation, an initial pattern description will provide a foundation for the more detailed analysis.

Criteria: Simplicity and Accuracy

A time series pattern description should provide an accurate answer to general descriptive questions. The results of pattern description should be concrete, have a straightforward interpretation, and be easy to communicate to a general audience.

The two criteria of simplicity and accuracy seem to be antithetical. If we increase simplicity, it may be at the expense of a decrease in accuracy. What degree of accuracy do we require, and what degree of complexity do we accept? How much accuracy must we sacrifice to achieve simplicity? To answer these questions, we first need to define our terms. It is relatively easy to define accuracy. Techniques of defining and measuring accuracy have been a part of the field of statistics since it began. However, the question of measuring simplicity has been considered only recently. A relatively new field, the study of the communication of quantitative relationships to a general audience, relates statistics and cognitive psychology.

Ehrenberg (1978,1981) suggests that a graphic description of the pattern of change over time in a variable is simple enough for effective communication if it has these characteristics:

- The pattern is presented in a simple visual structure.
- The description presents a small amount of information. Because it tries to communicate only one or two things, it makes few demands of the audience's short-term memory.

- The descriptive method is familiar to the audience. They have seen it used before, will recognize it, and will readily enter this instance of it into their long-term memory.

For example, a straight line is a simpler visual structure than a curve. Raw data are simpler than re-expressed or transformed data, because the raw data contain a smaller amount of information. The third criterion, familiarity to the audience, depends upon the audience. In our experience with an audience of criminal justice administrators, we have found that a straight least squares regression line is quickly recognized, but that a stochastic model or a polynomial line is not.

Some apparently simple statistical analysis methods are not simple at all, according to these criteria. For example, the "resistant lines" used in Exploratory Data Analysis (Tukey,1977; Velleman and Hoaglin,1981;Emerson and Hoaglin,1982) rely on involved summary calculations and transformations. Another method, picturing the data as a familiar object, such as a castle or a tree (Kleiner and Hartigan,1981) or human faces (Chernoff,1971, 1973; Flury and Riedwyl,1981), is entertaining and attracts the audience's attention, but does not present a simple visual structure to the audience. General audiences may be familiar with castles, trees, and faces, but they are not familiar with a quantitative concept being represented as a turret or as quizzical eyebrows.

Our criteria for accuracy are grounded in the general, descriptive decisions for which pattern description is intended to be used. If our objective were forecasting or model building, the necessary degree of accuracy would be greater than it is. However, our objective is not to set exact parameters, but rather, in an exploratory sense, to describe generally the variable's pattern over time.

For this objective, a time series pattern description is accurate enough if it gives the audience an idea of whether the variable has increased, decreased, or stayed the same, and whether or not there was a change or discontinuity in the direction of the series. It should also direct the user's eye to possible discrepancies from the overall pattern, such as unusually high or low occurrences (extremes), and seasonal or other cycles. A pattern description should present data to an audience simply and concretely, answer a few basic questions in a general descriptive way, and suggest to the audience more detailed questions that it may want to ask about the data.

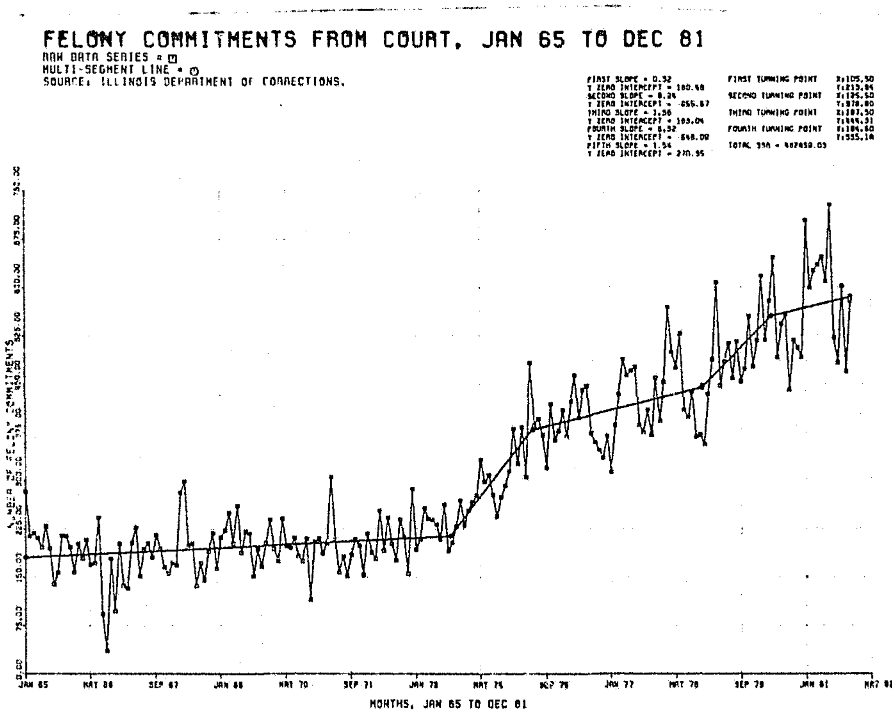
If it is to remain simple and understandable, a pattern description should not attempt to answer complex questions. Pattern descriptions are not exact, explanatory statistics. Such questions as confidence limits for the time periods in which a change in the pattern of the series occurred, forecasts, or the details of seasonal and other cyclical patterns are better left to the statistical methods that have been developed to handle them. On the other hand, to avoid a misspecified model, it is wise to preface any detailed explanatory analysis with a simple time series pattern description.

Pattern Description by Spline Regression

We have found that a linear spline regression line superimposed on a graph of the raw data describes the general pattern of a time series in a way that meets the dual criteria of simplicity and accuracy. Linear spline regression fits a least squares regression line with two or more connected segments to the data.[1] For example, the linear spline shown in figure 1 describes the pattern of change from 1965 through 1981 in the number of felony commitments per month to the Illinois Department of Corrections (IDOC). This five segment line fits the data better than other five segment lines, and better than the best-fitting four, three,

Figure 1

Example of a Pattern Description



[1]For an overview of linear spline regression, see Poirier (1976), who defines spline regression as follows (p. 2): "In the simplest sense a spline function is a piecewise function in which the pieces are joined together in a suitably smooth fashion." Also see note 3, page 6 below. For a review of the spline literature, see Wegman and Wright (1983).

two, or one segment line. It describes the number of commitments as remaining fairly steady between 1965 and late 1973, but rising sharply in 1974 and again in 1979. The pattern contains turning points in mid-1973, mid-1975, late 1978, and early 1980.[2] There were about 200 felony commitments per month in 1973. By 1981, there were more than 500.

The most obvious characteristic of a linear spline regression line is that every segment is connected to the next segment. Although there may be an abrupt difference between the slope of one segment and the slope of the next, there is no discontinuous gap between them. The line may change direction, but it remains unbroken. Instead of fitting separate regression lines to sections of the series, as a "piecewise regression" does, a linear spline regression fits one continuous line to the entire series.[3] Because every segment is connected to the next, the best fit for one segment is affected by the best fit for the adjoining segment.

Thus, a linear spline regression line consists of connected segments that differ from each other in their analytical definition (slope and intercept).[4] The segmented line is continuous in the sense that there is no gap between segments (they are connected). It is analytically discontinuous, because the definition of one segment is not the same as the definition of the next. Most other statistical time series descriptions are analytically continuous; that is, they describe the entire series with the

---

[2]The terms "join point," "knot," and "break point" are used synonymously with "turning point" in the spline function literature. We have found that "turning point" more clearly connotes, to people who are not statisticians, a change in the pattern of a variable over time. Although the term is sometimes used in a more exact sense (such as a "turning point error," Nelson 1973: 211), pattern description never uses "turning point" as a predictor of the future, or as an exact estimate of a past change.

[3]A piecewise regression line contains a discontinuous gap, an instantaneous jump in the pattern from one observation to the next. In some cases, such a jump may describe the actual situation, but it seems more reasonable to assume that the effect of a change in the structure of most social or economic series will not be instantaneous (Poirier 1976:1-3). For a discussion, see the section, "Presence of a Possible Discontinuity," page 56 below.

[4]A spline regression line does not necessarily consist of straight line segments. One or more of the segments could be a curve. See Wecker and Ansley (1983) for a method of fitting polynomial splines. In practice, however, we have found that connected straight line segments communicate better than either a curved line or connected curved line segments. A straight line segment provides a simple answer to one of the most commonly asked questions in practical situations, "Did the variable decrease or increase during a certain time period?"

same definition (a curved or straight line, a model of a stochastic process, a moving average or resistant smoother, and so on). A single, continuous definition assumes that the pattern of the series is constant over time. As Cox (1971:36; also see Kendall, 1976:29) argues in respect to polynomials, "the behavior of a polynomial in an arbitrary small region defines, through the concept of analytic continuity, its behavior everywhere." This assumption is open to question (Brown, Durbin and Evans, 1975: 149). A continually unchanging equation or process may describe many mathematically generated artificial series, but it may not describe practical empirical data, which may contain an abrupt structural change.

Of course, an analyst who has already discovered a change or discontinuity in the series, or who hypothesized it a priori, may write an equation accommodating that change, or may add an intervention term to the model (see Glass, et al, 1975). In this regard, McCleary and Hay (1980:143) argue that analytic continuity is an advantage of stochastic model-building, and that an exploratory "blind search," such as a spline pattern description, is "uninterpretable" in a test of an intervention hypothesis. Their argument is certainly true, but irrelevant in the present context. Pattern description is not designed to be used for explanatory purposes. It is descriptive. While analytic continuity, as compared to a blind, empirical description, is an advantage in the explanatory stages of analysis, it is a disadvantage in the initial, exploratory, descriptive stage. Description must precede explanation.

In summary, a linear spline regression line is not analytically continuous. The description changes with each line segment. In addition, it is concrete and straightforward, and can be presented to an audience in a simple visual structure (a succession of connected straight lines). It thus is a simple description of the pattern of change over time that is sufficiently accurate for general descriptive purposes.

Appropriate Applications

The degree of simplicity inherent in a linear spline pattern description limits its use to the initial description of patterns in data, especially in raw, untransformed data. More complex descriptions and transformations are necessary to answer more detailed questions. This section reviews the descriptive questions that can, and cannot, be answered with time series pattern description.

Like all least squares regression methods, pattern description is affected by the presence of extreme values (outliers). An unusually high or low observation will pull a line segment up or down and possibly cause the program to find different turning points than it would have otherwise. In addition, pattern description will not distinguish trend from drift. It is possible that what appears to be an increasing or decreasing trend, even over a long period of time, is actually a "random walk," due only to the tendency of one observation to move a random distance from the previous observation. Finally, pattern description ignores autocorrelation, which occurs when observations in a sequence are correlated with each other, and seasonality, which occurs when observations 12 months apart are correlated with each other.[5] Autocorrelation and seasonality may affect the line segment fit, especially the turning points.

Other, more complex, statistical methods have been developed to produce exact explanations and descriptions of such data.[6] Resistant smoothing methods (Velleman and Hoaglin,1981; Emerson and Hoaglin,1982) are called "resistant" because they resist the effect of extreme values, and the Census X-11 seasonal adjustment program has a built-in routine plus a number of options for weighing extremes (see Pierce,1980). Stochastic time series analysis can distinguish trend from drift (see McCleary and Hay, 1980:35-45 for a detailed discussion), and the time series literature abounds with ways to handle autocorrelated data (see Kendall,1976 or McCleary and Hay,1980 for an introduction) and seasonality (see Kendall, 1976 and Pierce,1980 for reviews). These complex methods, however, do not produce general descriptions of the actual, "raw," observations. Although time series analysis should not necessarily end with a description of the raw data, it should always begin there.

For example, the question of extremes is, to some extent, a subjective choice of whether to look at the forest or the trees. Should an extreme be considered accidental in the forest of values, and therefore be eliminated, or should it be pinpointed for special consideration? Both choices may be appropriate, but at different stages of the analysis. By summarizing the pattern over time of the raw data, a pattern description draws the observer's attention to exceptions from that general pattern.[7] For example, the felony commitment series (figure 1, page 5) contains

---

[5]These are extremely simplified definitions of seasonality and autocorrelation, but a more detailed discussion would be beyond the scope of this manual. For more information, see the references given in the following paragraph. The Statistical Analysis Center publication, "How to Handle Seasonality," guides the non-statistician to the various methods of detecting, measuring, and adjusting for seasonality.

[6]For a discussion of the use of splines in building explanatory models, and, specifically, the use of spline lags to estimate the degree and type of lags between dependent and independent time series variables, see Poirier (1976:85-106).

[7]Pattern description also draws the audience's attention to other kinds of discrepancy from the general pattern, such as seasonal or other kinds of cycles, or an increase or decrease in variation over time.

several observations that could be considered extreme, such as the low number of commitments in September 1968.[8] The decision of whether or not to weight or otherwise eliminate these values from the data should be based not only on a quantitative analysis of their statistical likelihood, but also on a substantive judgment as to their validity as a "true" representation of what happened in that month. To eliminate extremes mechanically, without a prior description and some investigation into their origin, may lead to a misspecified model. (For practical examples, see the section, "Series with Extreme Values," page 52 below.)

Like the issue of extremes, the issue of trend versus drift should be handled differently depending on the exploratory or explanatory stage of the analysis, and the needs of the problem at hand. To again use the felony commitment example, whether or not the rapid climb of commitments after 1973 was due to a "real" trend or to random drift, there were still about three times as many people committed per month in 1981 as in 1965. The resources of IDOC had to handle the larger number of people committed, whatever the cause of the increase. If the problem at hand is only to describe the actual number of commitments that IDOC had to handle over the time period, pattern description would be appropriate and sufficient.

In the same way, adjustment for seasonality or transformation for autocorrelation should follow a description of the raw data. The number of aggravated assault offenses known to the police in Illinois, for example, varies with the seasons. June, July, and August are usually high, and January and February are usually low (figure 2).[9] The large seasonal fluctuation affects the pattern description: the best line segment fit for the raw data (figure 2a) is similar to, but not exactly the same as, the best line segment fit for the seasonally adjusted data (figure 2b). Seasonal fluctuation and autocorrelation affect the line segment fit because they add systematic variation to the series.

---

[8]In our experience, we have found that time series pattern description is robust enough that one extreme value will not cause the program to find an "extra" turning point. However, an extreme may cause the program to place a turning point near the extreme rather than elsewhere in the series. For example, the low September 1968 value in the felony commitment series did not result in a new segment with a turning point in 1968. The program found a straight line from January 1965 through September 1973. For additional examples, see the section, "Series with Extreme Values," page 52 below.

[9]Source: Census X-11 adjustment, additive assumption. We did not use the felony commitment series as the example here, because it is not significantly seasonal. For the details of both analyses of seasonality, contact SAC.

Pattern Description of a Seasonal Series

Figure 2a

ILLINOIS INDEX AGGRAVATED ASSAULT, 1972 TO 1981

Each horizontal line equals 500 offenses

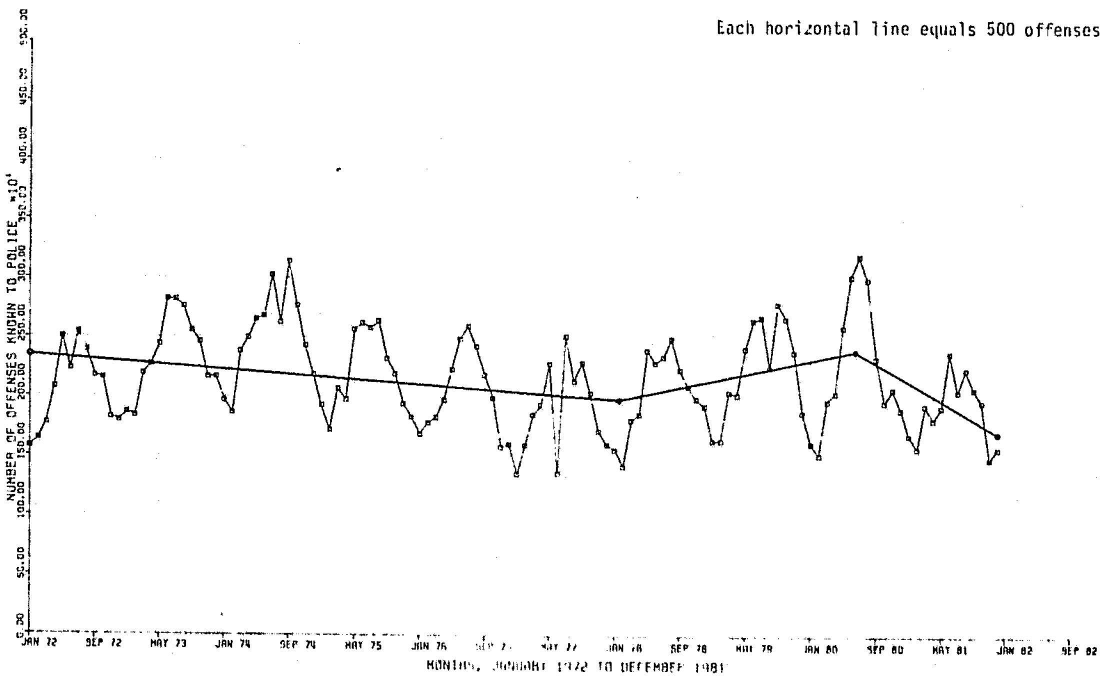


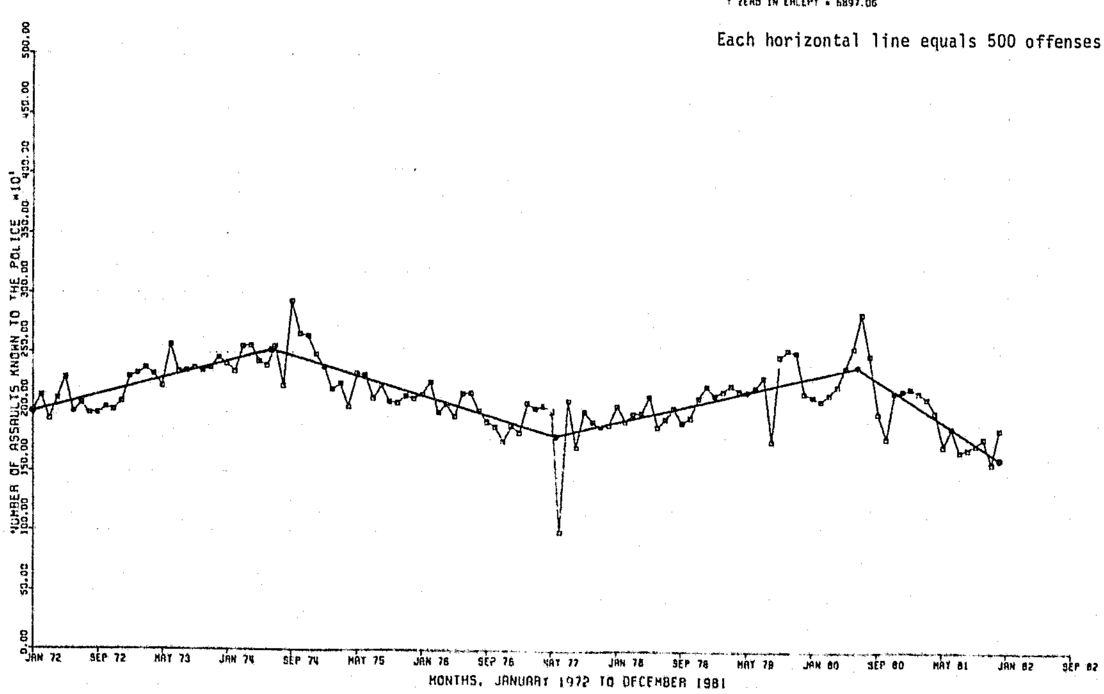
Figure 2b

ILLINOIS INDEX AGGRAVATED ASSAULT, SEASONALLY ADJUSTED

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

FIRST SLOPE = 17.90 FIRST TURNING POINT X:30.50 Y:2527.82
Y ZERO INTERCEPT = 1981.80
SECOND SLOPE = -20.37 SECOND TURNING POINT X:85.50 Y:1814.80
Y ZERO INTERCEPT = 3149.18
THIRD SLOPE = 15.87 THIRD TURNING POINT X:102.50 Y:2402.17
Y ZERO INTERCEPT = 775.00
FOURTH SLOPE = -43.85 TOTAL SSR = 3316435.81
Y ZERO INTERCEPT = 6897.06

Each horizontal line equals 500 offenses



There are many good arguments for removing such known, systematic variation from the series before developing an explanatory model.[10] On the other hand, the adjustments and transformations necessary to remove seasonality and autocorrelation may "overcorrect" and add their own systematic variation to the series. Not every crime series is seasonal or autocorrelated. Therefore, adjustments and transformations should not be performed mechanically on every series, but should be preceded by a description of the raw data.

Thus, pattern description of the raw data may not provide sufficient information for every application, but it does provide necessary information for any subsequent analysis, no matter how complex. As part of an initial data exploration, it serves as a preparatory step of a later detailed analysis. Complex analyses done without an initial description of the raw data may produce misleading results. The cleverest explanatory analysis is useless unless it is firmly rooted in a description of the patterns in the raw data. Without such an initial description, subsequent analysis risks error, and explanations and models risk misspecification.

If pattern description is too simple for some applications, it is unnecessarily complex for others. A linear spline pattern description is, essentially, a smoother. Some data are already so smooth that additional smoothing is superfluous. For example, population time series data at the local level are usually not actual measures of the population in each year or month, but interpolations between measures of one census and another. These interpolations are smooth by definition, and further smoothing is not necessary. Likewise, pattern description may be unnecessary for data that have been manipulated in any other way that produced a smooth series, or series that are too short (say, 10 observations) to show much variation. Raw data are unlikely to show such smoothness, but when they do, pattern description is not needed.

The chief limitation of time series pattern description is, paradoxically, also its chief advantage: simplicity. As long as pattern description is used only as its name implies, to describe the general pattern of a variable over time, simplicity is an advantage. However, users without a statistical background may find a pattern description so simple and so compelling that they are tempted to leap from descriptive conclusions to explanatory conclusions (for example, to forecast by extending the most recent line segment, or to assume that a turning point implies

---

[10]See the SAC publication, "How to Handle Seasonality," for a review. For another example, see "A Series Containing Seasonal Fluctuation," page 32, below. For an alternative method of handling autocorrelation, see Shine, 1981, 1982.

some intervention). If a picture is worth a thousand words, then a graph of a linear spline pattern description may be excessively verbose. It may appear to say more than it should. Therefore, users should be careful to use pattern description as a foundation for explanatory analysis, not as a substitute for it.

THE METHOD OF PATTERN DESCRIPTION

Time series pattern description by a line segment fit (linear spline regression) requires two steps. First, the analyst must find the best-fitting straight line, two segment line, three segment line, and so on, for the time series at hand. Second, the analyst must choose the overall best line segment fit from these alternatives.

SAC has compiled a package of computer programs that will find the best-fitting segmented line, given user criteria, and graph it. However, using this package is not a completely automatic process. As we have discussed above, pattern description has two criteria, simplicity and accuracy. While the package provides quantitative information as to the accuracy of a given segmented line, choosing an accurate fit that is also simple requires the user to make substantive decisions. The user must base the choice of best description not only on quantitative information but also on aspects of the practical situation, such as the questions that the audience is asking about the pattern of change over time in the variable.

The simplicity and accuracy of a pattern description depend upon each other. The most accurate description probably will not be simple, and the simplest description probably will be less accurate. For example, we could fit a 99 segment line to a 100 observation series, and produce an accurate, but certainly not a simple, pattern description. Choosing the best pattern description for a series requires the user to combine quantitative information about accuracy with qualitative information about the accuracy and simplicity the situation at hand requires. This section is a guide to the use of both kinds of information to find the best pattern description.

The Line Segment Fit Package

SAC uses two computer programs that calculate and graph segmented lines (linear splines). The Hudson/Fox program (Hudson,1966; Fox,1978; Block,1979) performs an exhaustive iterative search for the most accurate line segment fit, according to a least squares criterion, but it does this only for two segment lines. The Ertel/Fowlkes program (Ertel and Fowlkes,1975,1976) searches for the best line segment fit for any number of segments, also according to a least squares criterion, but the search is not exhaustive.

The Hudson/Fox program is easier to use than the Ertel/Fowlkes program, because it requires the user to make fewer decisions. It always finds the best-fitting two segment line, given one criterion: minimum length of a line segment. It has, however, a major disadvantage: it cannot calculate a linear spline regression that has more than two line segments.

The Ertel/Fowlkes program requires more participation by the user. It performs an abbreviated, not an exhaustive, search for the best-fitting segmented line. The user must choose the criteria for this abbreviated search, and then must interpret the results in light of those criteria. Thus, it is more difficult to use, but it can calculate linear spline regressions with more than two line segments.

Those who want a two segment line that they can produce almost automatically should use the Hudson/Fox program. Those who do not want to limit their pattern description to a two segment line should use the Ertel/Fowlkes program. Part II of this guide, the "Technical Manual," contains instructions for using each of these programs on the computer system of the Illinois Criminal Justice Information Authority. The following section contains a general description of what each program does, and guides the user in interpreting the results.

Hudson/Fox Two Segment Lines

The Hudson/Fox program finds the two segment line that best describes the data, by calculating every possible two-segment linear spline regression, and choosing the regression with the smallest sum of square residuals (SSR). As an example of Hudson/Fox results, figure 3 shows the best-fitting two segment line for the same series as figure 1 (page 5), felony commitments to the Illinois Department of Corrections.

The only user option required to produce a Hudson/Fox two segment line graph is to choose the minimum length for a line segment. Given that specification, the Hudson/Fox program computes every two-segment linear spline regression. For example, if the series is 100 observations long, and the user wants neither segment to be shorter than three observations, the program will calculate every two-segment linear spline regression with a turning point between observation 3 and observation 97.

The Hudson/Fox program then calculates the SSR for each of these regressions, chooses the fit with the smallest SSR, and plots it in a graph, such as figure 3. Information about the fit (the intercepts and slopes of the line, and the SSR) also appears on the graph.[11] The user may, instead of a plot of the single best fit, request intercept, slope, and SSR information for each two-segment spline regression.[12]

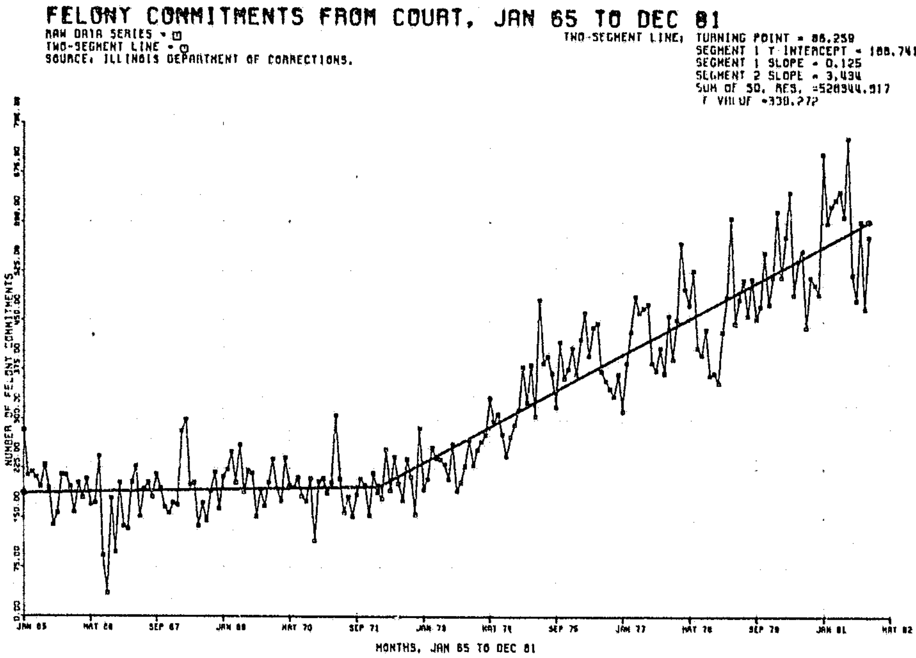
---

[11]An F value also appears on the graph. See the next section for an explanation.

[12]For examples and instructions for using this information, see Part II of this report, the "Technical Manual."

Figure 3

Example of a Hudson/Fox Pattern Description



Thus, the Hudson/Fox two segment line program produces a plot of the raw data, with the best-fitting two segment line superimposed on it. It will also plot the best-fitting straight line (ordinary least squares regression line). The user may choose a plot of the straight line only, the two segment line only, or both. The user has other plotting options, such as the x-axis and y-axis range of values, and the labels to be printed on the graph. These details are discussed in Part II of this guide, the "Technical Manual."

Ertel/Fowlkes n Segment Lines

A program that would examine every possible linear spline to find the best-fitting segmented line with two, three, four, five, and so on segments would be so large and unwieldy that it would be neither simple nor inexpensive to use (Tishler and Zang, 1981: 982). The Ertel/Fowlkes program provides an alternative to such an exhaustive search, and produces results that, for the two segment case, are usually the same as Hudson/Fox results. Like the Hudson/Fox program, the Ertel/Fowlkes program searches for the fit with the smallest SSR, but the search is abbreviated, beginning with an initial routine that reduces the number of calculations necessary. The Ertel/Fowlkes algorithm has two parts: an initial search for a starting partition, and a more exhaustive search for the best linear spline fit. The user must set criteria for each part.

The initial search for a starting partition fits an ordinary least squares regression to the first n observations in the series,[13] predicts from this the next (n + 1) observation, and calculates the standard error by which the actual n + 1 observation differs from the predicted n + 1. It then fits a regression to the first n + 1 observations, and continues, all the while cumulating the standard errors and counting the number of observations in this "run" of positive or negative standard errors. If it encounters a positive standard error in a run of negatives (or vice versa), it stops cumulating and begins again with a new run. This continues until both of two things happen: 1) the length of a run of positive or negative standard errors reaches a minimum number of observations, and 2) the cumulated standard error in the run reaches a minimum. When both cf these happen, the process stops, and the program records a partition that begins with the first observation in the series and ends with the initial observation in the run that met the two minima. The process then begins again, starting with that initial observation in the run. The eventual result, after this search has continued to the end of the series, is a partition of the series into segments. The second part of the Ertel/Fowlkes algorithm then begins with this starting partition, and searches for the best fit.

The user sets two criteria for the initial search: the minimum length of run and the minimum cumulated standard error. These two criteria, and the characteristics of the series, determine the starting partition that the initial search will find. A conservative choice, a long minimum run length for example, may result in a starting partition with only one segment (a straight line). With the same series, but a shorter minimum run length, the program may find a starting partition with two, three or more segments. The starting partition, in turn, determines the maximum number of segments that the second step of the program will find in any segmented line. With a starting partition of three segments, for example, the second stage of the program will search for the best-fitting three segment line, two segment line, and straight line, but it will not search for a four segment line.[14]

The second step of the Ertel/Fowlkes program begins with the starting partition, and conducts an iterative search for the best-fitting spline, using a least squares criterion. After it fits a linear spline to the starting partition and calculates

---

[13]The program sets n equal to the user-specified minimum length of run.

[14]See Part II (the "Technical Manual") for instructions, suggestions, and examples for choosing the minimum run length.

the associated SSR, it begins to move observations, one observation at a time, from one segment to another, and calculate each linear spline regression and each SSR. It continues to move observations as long as the SSR continues to decrease, maintaining the minimum segment size set by the user. By this iterative search, the program finds the best-fitting spline that has the same number of segments as the starting partition.[15] It next begins to search for the best line segment fit with one fewer segment. It combines segments one and two, and then searches for the smallest SSR by moving observations from segment to segment, as before. When it can no longer reduce the SSR, it goes back to the original line segment fit, combines segments two and three, and begins the search again. The final result will be line segment fits for whatever number of segments the initial search found, and for each fewer number, including a one segment (straight line) fit.

For example, figures 4a to 4f show the entire Ertel/Fowlkes graphic output for the felony commitment series. For this example, we set a minimum run length of six observations for the initial search, and a minimum segment length of 12 observations for the final line segment fit. The initial search found a starting partition with five segments. From this partition, the second step of the program searched for the best-fitting five segment line, and found the line graphed in figure 4a. The program then combined these segments, two at a time, and searched for the best-fitting four segment line (figure 4b), the best-fitting three segment line (figure 4c), and the best-fitting two segment line (figure 4d).[16] Finally, the program calculated and graphed a straight least squares regression line (figure 4e). (See the next section for a discussion of figure 4f, the final graph produced by the Ertel/Fowlkes program.)

---

[15]Although the number of segments in this line segment fit will be the same as the number of segments in the starting partition, the distribution of observations among the segments may be very different.

[16]Note that the two segment line found through the abbreviated Ertel/Fowlkes search is essentially the same as the two segment line found through the exhaustive Hudson/Fox search (figure 3). The slight difference in slopes and intercepts is due to slightly different methods of calculation. Ertel/Fowlkes calculates spline regressions so that the turning points are halfway between two observations (86.50 in this case). Hudson/Fox calculates an exact turning point, either exactly at an observation or the exact point (to three decimal places) between observations (86.259 in this case). Such a difference affects the slopes and intercepts slightly, but not enough to have any effect in most practical decision situations. Still, before making a major decision based on an Ertel/Fowlkes two segment description, it is a good idea to compare it to the exhaustive Hudson/Fox two segment line.

Example of a Complete Ertel/Fowlkes Graphics Output

Figure 4a
(Five Segment Line)

ILLINOIS FELONY COMMITMENTS FROM COURT, 1965-1981

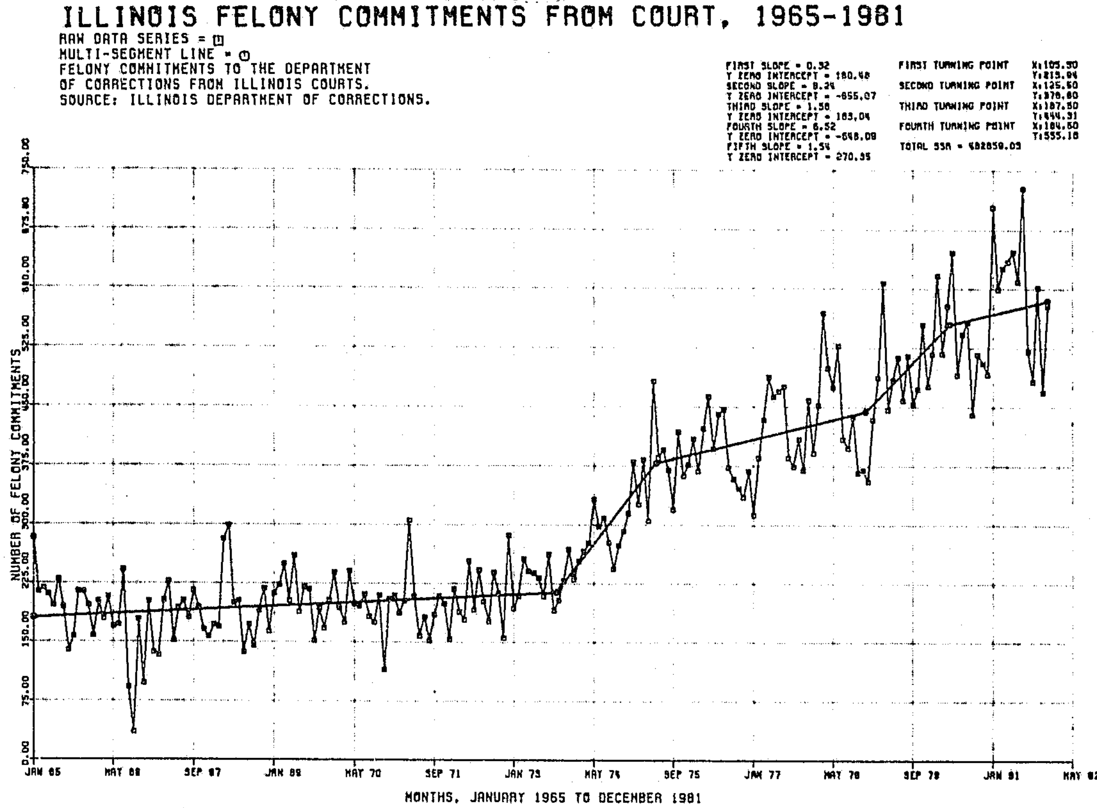
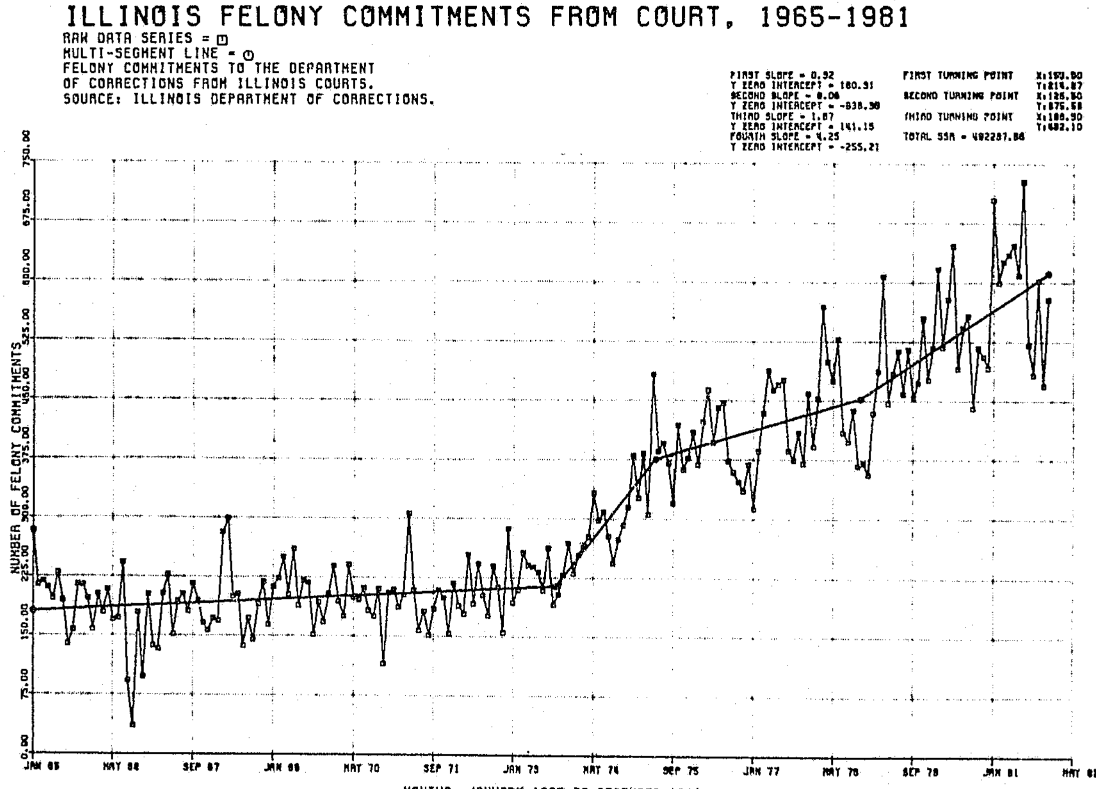


Figure 4b
(Four Segment Line)

ILLINOIS FELONY COMMITMENTS FROM COURT, 1965-1981



Example of a Complete Ertel/Fowlkes Graphics Output (Cont.)

Figure 4c
(Three Segment Line)

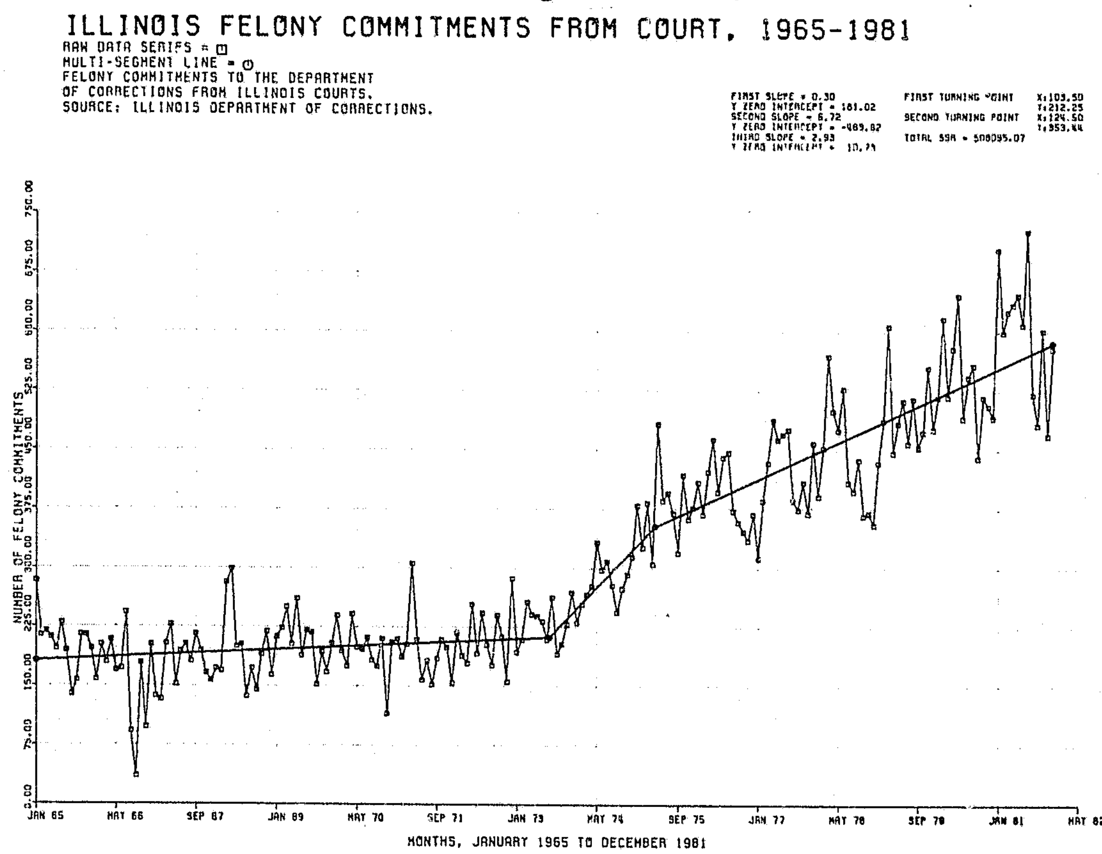
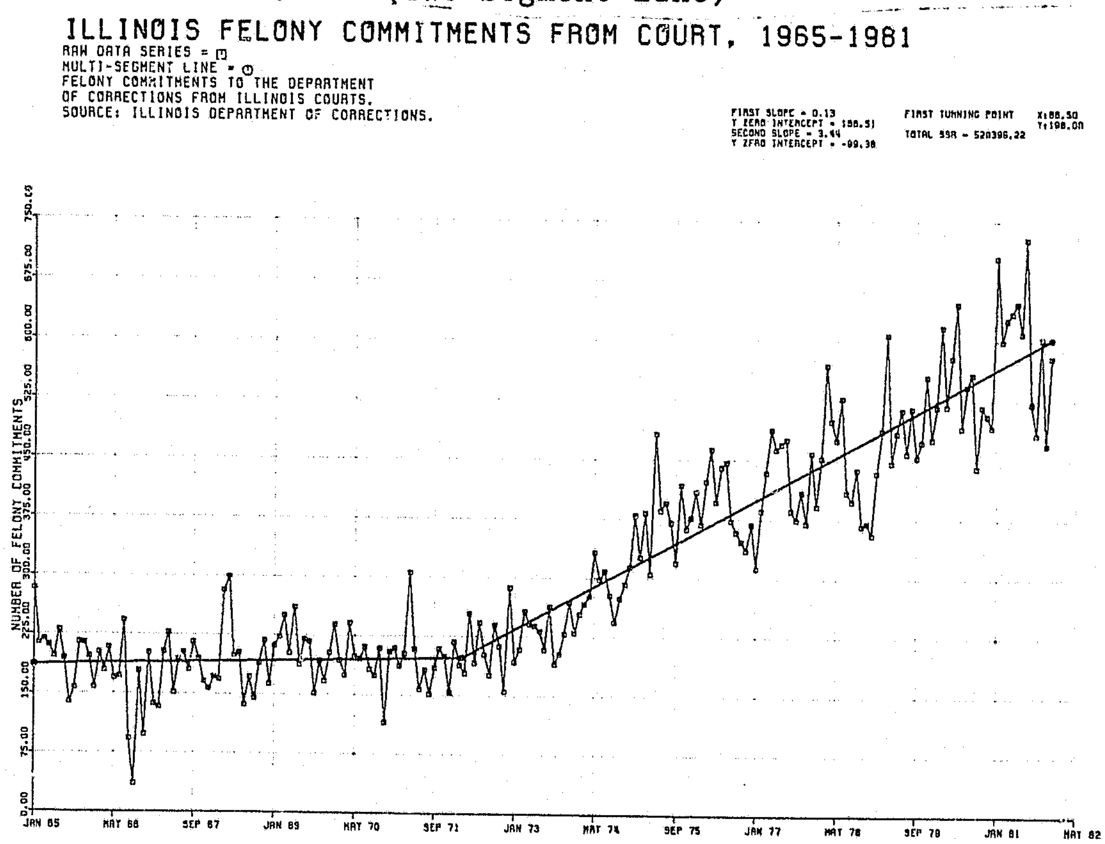


Figure 4d
(Two Segment Line)



Example of a Complete Ertel/Fowlkes Graphics Output (Cont.)

Figure 4e
(One Segment Line)

ILLINOIS FELONY COMMITMENTS FROM COURT, 1965-1981

RAW DATA SERIES = ◻
MULTI-SEGMENT LINE = ⊙
FELONY COMMITMENTS TO THE DEPARTMENT
OF CORRECTIONS FROM ILLINOIS COURTS.
SOURCE: ILLINOIS DEPARTMENT OF CORRECTIONS.

FIRST SLOPE = 2.17

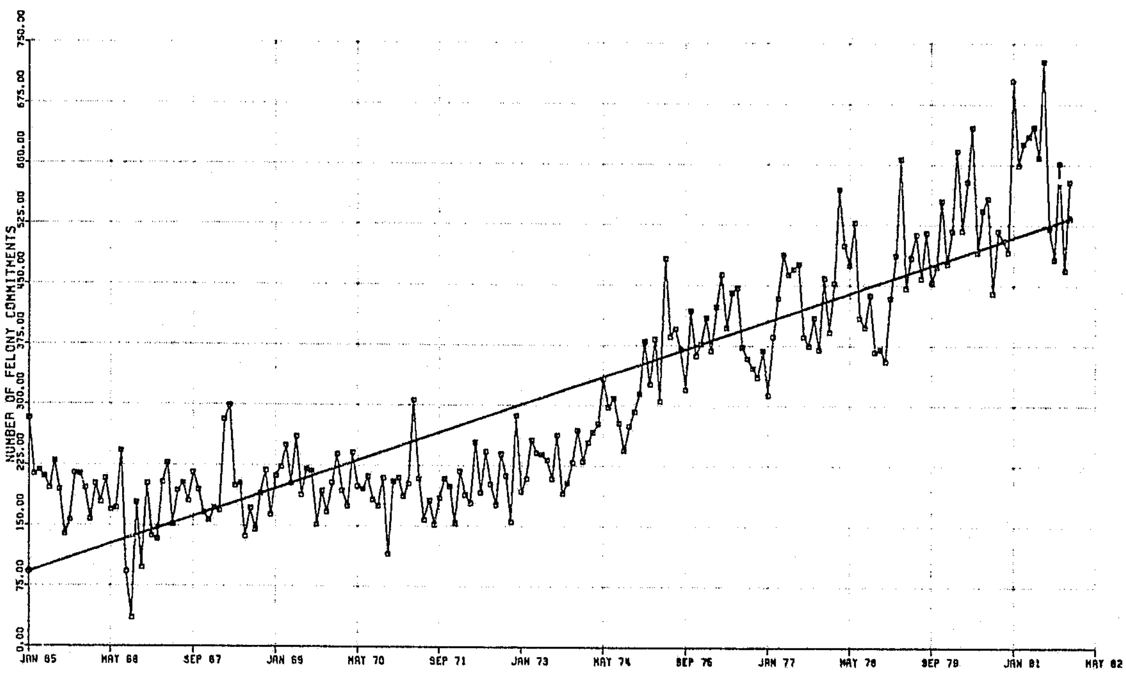
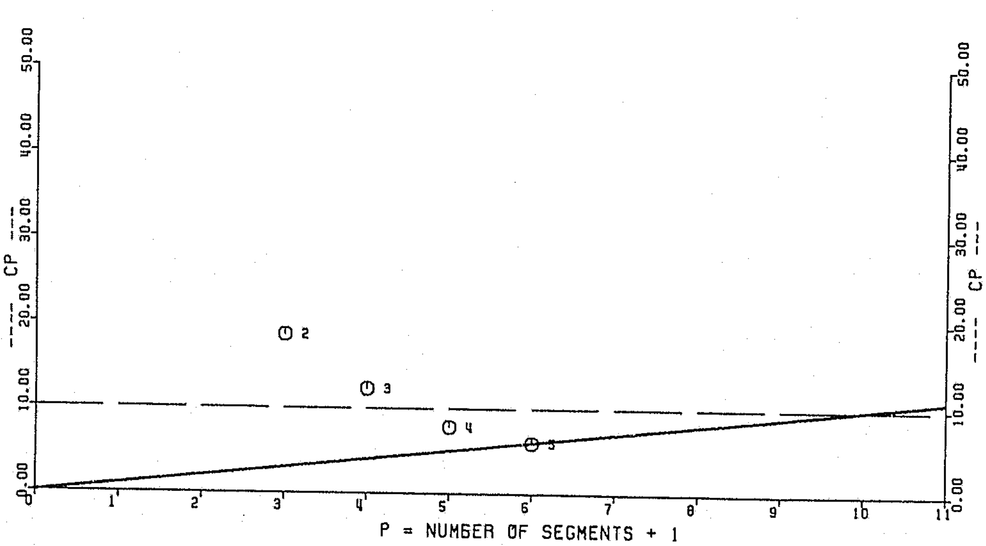


Figure 4f
(C_p Plot)

ILLINOIS FELONY COMMITMENTS FROM COURT, 1965-1981

SUMMARY: CP BY P
NUMBER OF OBSERVATIONS:204
MINIMUM PTS./SEGMENT :12
MIN STARTING PARTITION:6

NUMBER OF SEGMENTS IN INITIAL SEARCH = 5
CP. 1 SEGMENT FIT = 200.4 * OFF CHART
CP. 2 SEGMENT FIT = 18.7
CP. 3 SEGMENT FIT = 12.3
CP. 4 SEGMENT FIT = 7.8
CP. 5 SEGMENT FIT = 6.0



Thus, the Ertel/Fowlkes program produces a package of alternative line segment fit pattern descriptions. The number of graphs in the package depends upon the number of line segments found in the initial search for a starting partition, which depends upon the user's criteria. As with the Hudson/Fox program, the user has options as to the plots themselves--the labels, the ranges of the x-axis and y-axis, and so on. Part II, the "Technical Manual," discusses all of these options for obtaining Ertel/Fowlkes results. In the remaining sections of this guide, we discuss how to interpret results once they have been obtained.

Choosing the Best Pattern Description

Both the Hudson/Fox and the Ertel/Fowlkes programs produce alternative pattern descriptions of a time series. The Hudson/Fox program produces graphs of the best-fitting two segment line, and (optionally) the best-fitting straight line.[17] The Ertel/Fowlkes program searches for the best line segment fit with the greatest number of segments, given the characteristics of the series and the user's criteria, and then searches for the best fit for each lesser number of segments. Each of these line segment fits is an alternative description of the pattern of change over time in the series. The question to be asked now is this: which of these alternative pattern descriptions is accurate enough, yet simple enough, to use in the situation at hand?

The user must weigh the degree of accuracy produced by a particular segmented line pattern description against both the degree of accuracy required to answer the questions at hand and the degree of simplicity the audience or the practical situation requires. Thus, the choice of the "best" description depends not only upon the degree of precision and accuracy of alternative line segment fits, but also on a variety of considerations relevant to the particular application. The technique of time series pattern description is not just an exercise in interpreting quantitative information. Pattern description must also be qualitative.

This section describes some quantitative and qualitative guidelines to choosing the best overall pattern description. Although the choice of the best pattern description is necessarily subjective, these guidelines provide a framework for that subjective decision.

---

[17]Even if the user does not choose to obtain a straight line graph from the Hudson/Fox program, the program calculates (internally) a straight line (ordinary least squares) regression. The complete regression statistics, if requested, appear on the non-graphic output, and the regression information is used in the calculation of the F value.

In addition, regardless of which pattern the user ultimately chooses, the comparison of one pattern to another can be instructive. Comparing alternative pattern descriptions draws the user's eye not only to discrepancies from each pattern, such as extremes, cycles, and changes in variation, but also to subtle differences between more and less complex patterns.

Indicators of Accuracy

Poirier (1976:107-144) surveyed estimation methods for the most accurate spline fit and concluded that, unless we begin with a finite set of possible turning points, there is no way to find the best-fitting of all possible segmented lines.[18] He recommends that we consider using conditional probability (Bayesian) methods, which begin with some prior information as to the location of points of structural change.[19] But where do we get this prior information? An initial pattern description, coupled with familiarity with the characteristics of the data set, can provide the background information necessary for explanatory analysis, including estimation and hypothesis testing. However, at the initial descriptive stage of analysis, exact statements of accuracy are not possible.[20]

The two indicators of accuracy discussed in this section are not, therefore, exact estimates of the best of all possible line segment descriptions, and cannot be used to test hypotheses or to establish confidence intervals for turning points. One of the indicators provides a graphic summary of the relative amount of accuracy in the alternative segmented lines produced by the Ertel/Fowlkes program, and the other indicator compares two alternative fits, such as the Hudson/Fox straight line and two segment line. Neither indicates the exact amount of accuracy, but either can be used as a general, exploratory indicator of relative accuracy, in conjunction with the indicators of simplicity (below) to arrive at a subjective decision as to the best pattern description for a particular situation.

---

[18]"Classical attempts to both estimate the unknown switch point (or knot) and test for parameter change across regimes (or segments) are almost surely to break down since under the null hypothesis of no parameter change, the unknown switch point or knot is not identified" (Poirier,1976:142).

[19]According to Wahba (1978), spline smoothing is equivalent to Bayesian estimation with a partially improper prior.

[20]Feder (1975a,1975b) discusses the problem of inferring that two adjacent segments are identical (that the segmented line contains one fewer segment). He shows (1975b:84) that, although chi square results are applicable under "suitable identifiability conditions," if there actually are fewer segments than in the model, "then the least squares estimates are not asymptotically normal and the log likelihood ratio statistic is not asymptotically X^2."

- C_p Plot

A C_p plot, such as figure 4f, graphically shows the amount of bias versus the amount of random error in a number of alternative regression equations. In the case of pattern description, the alternative equations are various line segment fits. Developed by Mallows (1964; also see Gorman and Toman, 1966; Mallows, 1973, 1980; Kennard, 1971; Daniel and Wood, 1980), the C_p plot differentiates between two components of error: bias and random error.[21] Bias is the difference between the expected values of the true (unknown) equation and the expected values of the equation being used to fit the data. Random error is the random variation around the expected value of the fitted equation. We may decrease bias by increasing the number of terms in the equation, but the random error may then increase.

C_p is a relative, not an absolute, measure of accuracy. It indicates the accuracy of a certain line segment fit relative to a group of alternative fits. If we have a number of alternative equations, we can plot the C_p against the p for each equation. P is equal to one plus the number of segments, and C_p is defined below.[22] Those line segment fit equations with little or no bias will cluster around the line $C_p = p$. Equations with substantial bias will plot above the line.

For example, in figure 4f, which is the C_p plot for the five alternative line segment fits for felon commitments (figures 4a-4e), the distance from each plotted C_p value and the line $C_p = p$ indicates the amount of bias in the line segment fit, and the distance from the line $C_p = p$ to the x-axis indicates the amount of random error. Thus, for the two segment fit, the total error (C_p) is 18.7, consisting of random error (3.0) plus bias (15.7). For the five segment fit, the total error is 6.0, the random error is 6.0, and the bias is negligible, or about zero. For the one segment fit, C_p is completely off the chart (200.4), and since the amount of random error is low (2.0), the amount of bias is very high. The two, three, and four segment fit C_p's are closer to the estimated zero bias line than the one segment C_p, but the five segment C_p is actually on the line. The five segment fit, therefore, has more random error but less bias than the four segment fit. Although the four segment fit has more bias, the user may decide to accept this degree of bias in exchange for a simple pattern description. The C_p plot graphically illustrates the gains and losses involved in making this choice. Thus, the C_p plot assists the user in making what is, essentially, a subjective decision.

---

[21]The C_p plot is similar in concept to the scree plot used in factor analysis. See Cattell (1966) or Harman (1976).

[22]The number of possible line segments is, of course, limited by N. In practice, SAC limits the number of line segments to a maximum of 10, regardless of N. See Part II, the "Technical Manual."

C_p is defined in the following way:

$$C_p = \frac{SSR_p}{s^2} - (N - 2p)$$

Where: SSR_p is the sum of square residuals from a line segment fit with (p - 1) line segments;
s^2 is an unbiased estimate of the variance;
N is the number of observations;
p is the number of line segments plus one.

The calculation of C_p depends upon an unbiased estimate of the variance (s^2).[23] As an estimate of variance, Mallows uses the variance of the best-fitting regression of the alternative regressions being compared. Therefore, the C_p for a certain line segment fit might change, depending on the estimate of variance, which depends on the line segment fit with the smallest SSR of the line segment fits being compared, which depends on the starting partition. Given different user criteria, the initial search routine of the Ertel/Fowlkes program will find different starting partitions.[24] Thus, a starting partition of five segments will result in five alternative line segment fits: a five segment fit, a four segment fit, a three segment fit, a two segment fit, and a straight line. The estimate of variance used to calculate the C_p for each of these five will be the variance of the best-fitting regression of the group.[25] Now, another Ertel/Fowlkes run with the same series, but different criteria for the initial search, might result in a starting partition with four segments, and thus a four segment fit, a three segment fit, and so on. These two separate runs may produce identical three segment fits, with exactly the same line segments and exactly the same SSR's, but with different C_p's. One C_p is being compared to a set of alternative fits that includes a five segment fit (with a low SSR), and the other is being compared to a different set of alternatives. Therefore, use C_p as a relative measure of accuracy.

The next section of this guide provides a number of examples of the use of the C_p plot as an indicator of relative accuracy. Together with the other guidelines to accuracy and simplicity, these examples will help the user choose the best segmented line pattern description for a given situation.

---

[23]See Wahba (1977) for a discussion of the situation when the variance is not known.

[24]Part II of this guide, the "Technical Manual," contains suggestions and examples for choosing criteria for the initial search routine.

[25]This will usually, but not always, be the line segment fit with the greatest number of segments. See "Fewer Segments may be More Accurate," page 44 below.

• The F-Test

Because the Ertel/Fowlkes program produces a number of alternative line segment fits, the C_p plot, which is a graphic means of comparing numerous alternative fits, is well suited to it. However, it is not particularly suited to choosing between only two alternative fits, such as are produced by the Hudson/Fox program.[26] Cox (1971) and McGee and Carleton (1970; also see Chow, 1960:602) suggest an F-test as an indicator of the degree of improvement in the least squares fit with the addition of a line segment.[27] We have found an F to be very useful as a general indicator of the accuracy of the best-fitting two segment line relative to the best-fitting straight line. It should be used, however, only as a rough, exploratory indicator, and not an exact statistic.[28]

It is defined in the following way:

$$F = \frac{(SSR_{s-1} - SSR_s)/2}{SSR_s/(N - 2s)}$$

Where: SSR_s is the sum of square residuals for a line segment fit with s segments;

SSR_{s-1} is the sum of square residuals for a line segment fit with one fewer segment than SSR_s;

s is the number of line segments in SSR_s.

2 and (N - 2s) are the degrees of freedom.

For example, the two segment Hudson/Fox fit for felony commitments (figure 3, page 15) has a very high F value (339). With 2 and 200 degrees of freedom, and given the usual assumptions for F tests, the probability that this two segment description is not really more accurate than the best straight line description is less than .001 (one in a thousand). However, because at least one of these assumptions, independence of observations, does not usually hold for time series, the F value should be interpreted in a descriptive, exploratory way. Rather than saying that the

---

[26]Although we recommend that the C_p plot be used with Ertel/Fowlkes results, and the F be used with Hudson/Fox results, a C_p can be calculated for a Hudson/Fox two segment line, and an F value can be calculated for any of the n segment lines produced by Ertel/Fowlkes. The equations given here for C_p and F can be generally applied to any number of segments.

[27]We are also grateful for the suggestions of an anonymous reviewer.

[28]The interpretation of the F value in pattern description is analogous to the interpretation of the F value as a general indicator of the presence of stable seasonality in the Census X-11 program.

pattern of felony commitments suddenly changed direction in January 1972, say that the number of felony commitments per month stayed fairly steady (about 190 or 200) in the late 1960s, and then increased rapidly in the 1970s.

The F value appears on each Hudson/Fox plot, with the other information about the fit (turning point, intercept, slopes, and SSR). For examples of the use of the F in conjunction with other guidelines to choose the best pattern description for the situation at hand, see the following section.

Indicators of Simplicity

The amount of accuracy versus simplicity a particular situation requires is very subjective and difficult to quantify. The analyst can, however, ask a standard set of questions about each pattern description problem. These questions--the maximum number of segments, the minimum length of any segment, and the minimum amount of change between one line segment and the next that could possibly make a difference to the audience and to the question or decision under discussion--are discussed in this section.

The answers to these questions will determine the user criteria for the line segment fit programs. They will also provide guidelines for weighing accuracy against simplicity in the choice of a single best description from among the alternative line segment descriptions produced by the programs.

- Length and Number of Segments

The fewer the number of line segments, and the longer the length of each line segment, the simpler the pattern description. The user should ask: what is the shortest time interval that would affect the situation at hand? What is the largest number of changes in the pattern of the series that would affect the situation at hand? As a general rule, choose the simplest possible pattern description--the segmented line containing the longest segments and the fewest segments--that will still answer the question at hand.

The user must make two decisions: the shortest acceptable length for any line segment, and the greatest acceptable number of line segments. These two decisions are interdependent. The number of segments can be no greater than the number of observations in the series divided by the number of observations in the shortest acceptable line segment. Suppose we have a series that is 144 observations long (a 12-year monthly series), and we say that we will not permit any single segment to be shorter than 24 months. In that case, the greatest possible number of line segments will be six. Thus, although minimum line segment length is a user option with both line segment fit programs, and number of segments is not an option, choosing a minimum line segment length implies a choice of maximum number of line segments.

Different choices of minimum line segment length, when translated into user options for the Ertel/Fowlkes and the Hudson/Fox programs, can produce very different pattern descriptions. For example, five-month and 12-month options for the same series, Chicago homicides without a gun, produced different line segment fits (figure 5). When allowed a minimum segment length of five months, the Ertel/Fowlkes program found an eight segment line (figure 5a). When the minimum segment length was 12, the program found a three segment line (figure 5b). The eight segment fit may be more accurate, but does such a complex description really matter in the situation at hand? Does describing such a brief change as the five-month-long dip in early 1975 make any practical difference, or is the audience interested only in changes that occurred at least 12 (or more) months apart? Only the user, who knows the needs of the practical situation, can determine this.

- Difference between Neighboring Segments

A given segmented line may contain adjacent segments that are really very similar to each other. They may both increase (or decrease), but at slightly different rates. Should segments such as these be combined into one long segment, producing a segmented line with one fewer segment in it, or is the small amount of change from one segment to the next important in answering the questions at hand? A segmented line with more segments will often be more accurate, and Feder (1975a:68; also see Poirier, 1976:131) suggests a complex "rule of thumb" for inferring that two adjacent segments are identical.[29] However, the decision must depend not only on accuracy, but also on the degree of accuracy and simplicity the situation requires. Does a slight difference in slope between one line segment and the next affect the practical situation?

For example, the alternative line segment descriptions of felony commitments (figures 4a-4d above) differ in the degree of detail they show in the latter part of the series, from the early 1970s through 1981. All the alternative patterns (except the straight line) show a long segment, increasing slightly, from 1965 through 1971 (or slightly longer). They all show a rapid increase in the 1970s. The only difference is the amount of detailed change shown within this rapid increase. Each alternative contains progressively more detail, until the five segment fit (figure 4a) describes a rapid increase in 1974, a slight increase from 1975 to 1978, another rapid increase in 1979, and a slightly less rapid increase in 1980 and 1981. The four segment alternative description (figure 4b) is similar to the five segment line in all respects, except that it draws a straight line from 1979 to 1981.

---

[29]See note 20, page 22 above.

Different Minimum Segment Lengths for the Same Series

Figure 5a
(Eight Segment Line)

CHICAGO HOMICIDES WITHOUT A GUN, 1965-1978

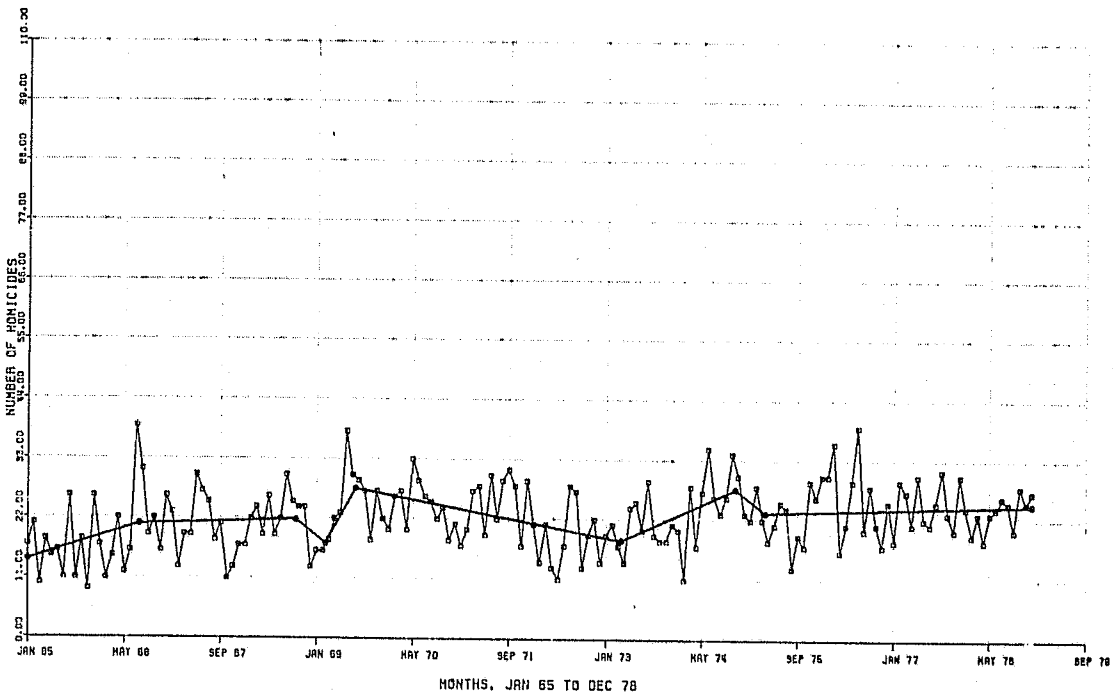
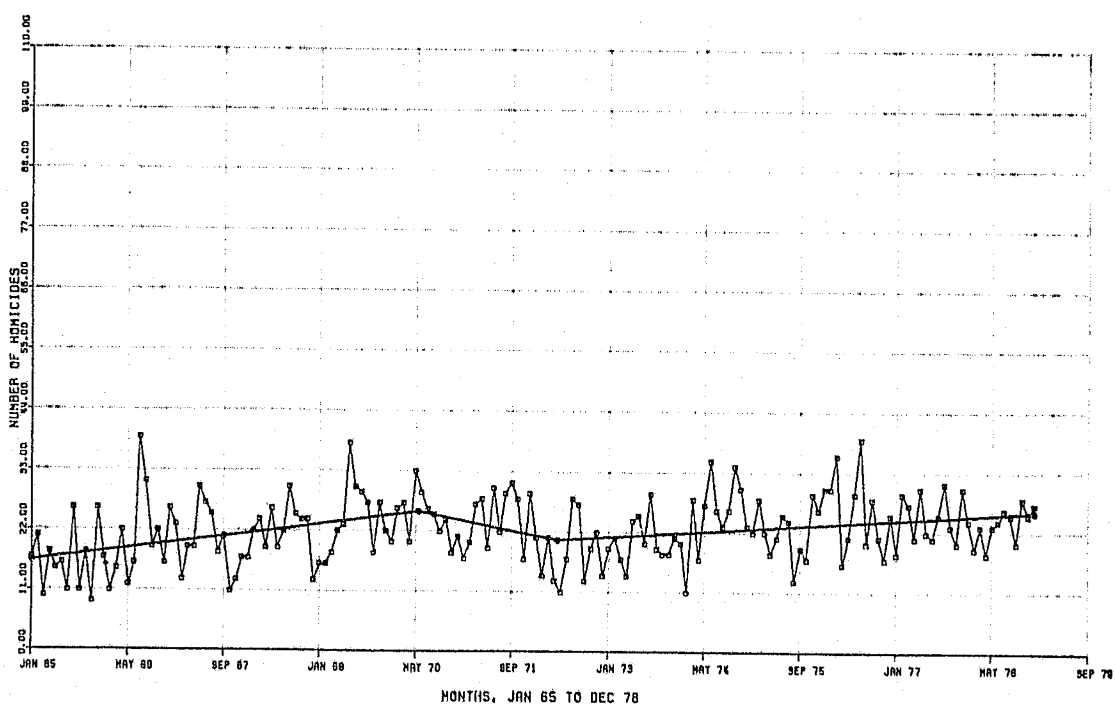


Figure 5b
(Three Segment Line)

CHICAGO HOMICIDES WITHOUT A GUN, 1965-1978



What degree of precision does the situation require? Do you need to know whether or not the rate of increase changed between 1974 and 1981, or do you need to know only whether or not there was a general increase during those years? If additional segments are not important to a practical decision, there is no reason to choose a complex description over a simpler description, even if the complex fit is more accurate.

EXAMPLES OF PATTERN DESCRIPTION

Time series pattern description is as much an art as it is a science. Although the pattern description computer package, with its graphic descriptions and C_p plots, provides much information regarding the accuracy and simplicity of alternative patterns, this information cannot be interpreted mechanically. Every practical situation is slightly different.

Like all arts, time series pattern description requires practice. The more familiarity the user has had with a variety of practical decision situations and a variety of time series, the easier pattern description will be. In this section, we give users the benefit of some of our experience in describing patterns over time of criminal justice data. The examples below are a selection of pattern descriptions SAC has done during the development of the method. Most of the examples are the result of requests from criminal justice personnel and other SAC users for answers to practical questions involving time series pattern description.

A Series Containing Seasonal Fluctuation

The pattern of change in the number of Index burglary offenses known to the police in Illinois during the 10-year period from 1972 through 1981 is best described as a three segment line (figure 6a). However, the best description of the same series adjusted for seasonality (figure 6b) is somewhat different.

The best description of the raw data, the number of Index burglaries occurring in Illinois, is a three segment line that increases rapidly from 1972 to 1974, decreases slightly in 1976, and gradually increases over the next five years. In 1972, there were about 7,000 burglaries known to the police in a typical month, but in 1981, there were almost 12,000.

Burglary in Illinois fluctuates slightly with the seasons.[30] There are usually more burglaries in August than in other months. This seasonal fluctuation is much more clearly present in the later years than in the earlier years.

Although the degree of seasonal fluctuation is slight, it does obscure one interesting aspect of the pattern of change over time in Illinois burglary. The best description of the seasonally adjusted series is a six segment line. The difference between the two descriptions may be important for some decisions. Although the description of the raw data shows that burglary increased in 1980 and 1981, the description of the adjusted series shows that, controlling for seasonal fluctuation, burglary decreased in those years.

In general, removing seasonal fluctuation removes variation due to a known cause (the seasons). With this variation removed, other patterns in the series will be easier to detect. However, the process of removing seasonal fluctuation is not completely objective. It is a complex transformation that may add systematic error to the series. Therefore, we recommend two things: 1) Do not assume that every monthly series is seasonal. Begin with a question, not an answer. Ask yourself, "Is this series seasonal?", and set some objective criteria to determine your answer. 2) Never lose sight of the raw data. Examine the pattern over time of the raw data before examining the pattern of the seasonally adjusted data.[31]

---

[30] For details of this analysis of seasonality, contact the authors.

[31] For an introduction to using the most common seasonal analysis methods, see the SAC publication, "How to Handle Seasonality."

Figure 6a

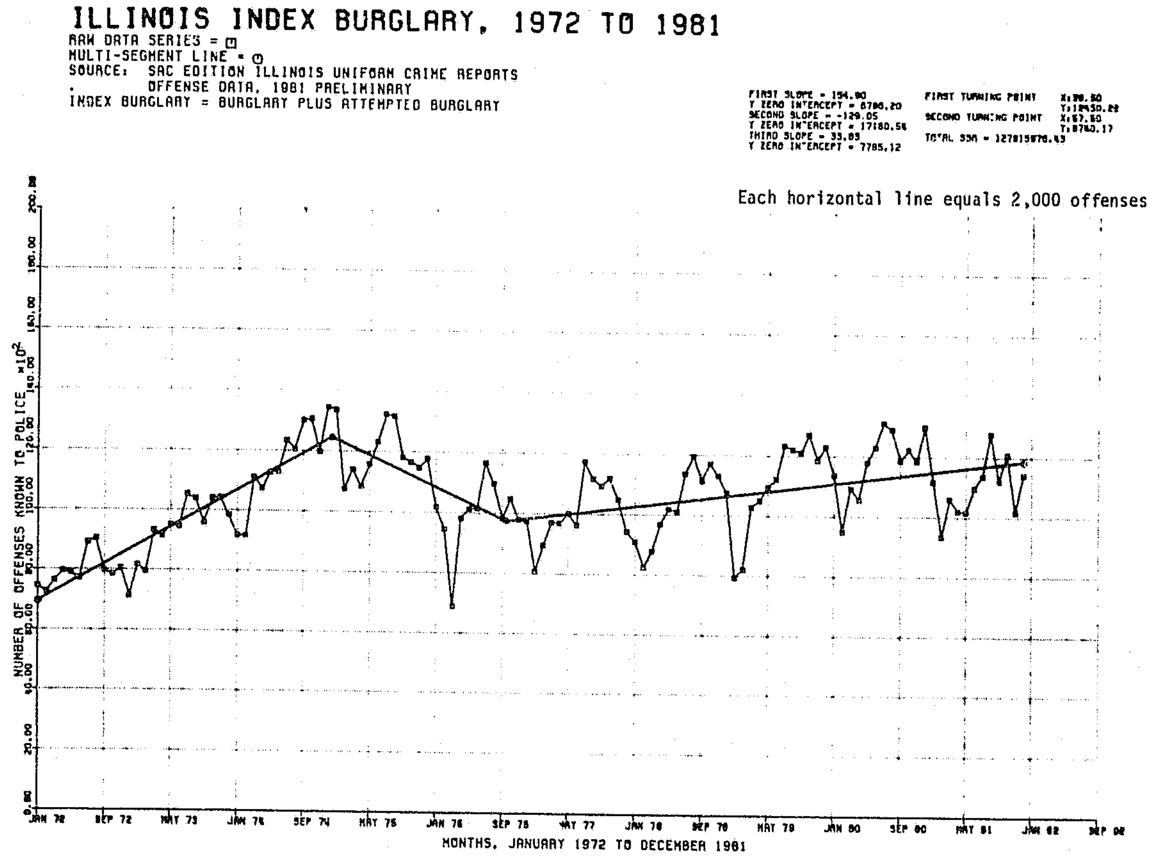
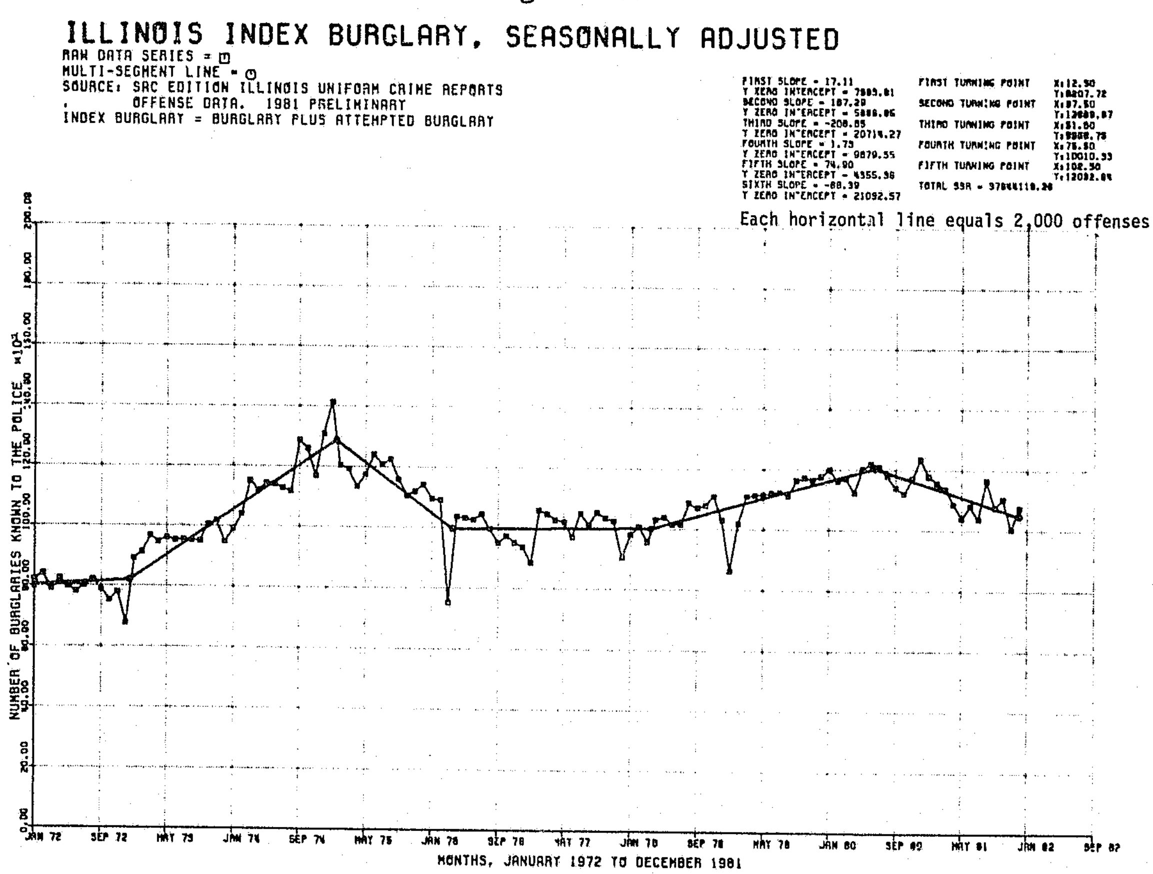


Figure 6b



Initial Search versus Final Search Criteria

The pattern description of the number of Index larceny/theft offenses known to the police per month in Quincy, Illinois, from 1972 to 1981 (figure 7) illustrates the different uses of the criterion for the initial search (run length) and the criterion for the final fit (shortest segment length).

We often find that it is a good idea to use a conservative (short) initial search criterion, even though we know that we will not want such short segments in the final description. The object of the initial search is to find a beginning partition for the second stage, the exhaustive search. If the criterion is too long, the initial search may find a partition with only one, two, or three segments. Even though you may eventually decide that a fit with one, two, or three segments is the best description, you want to base this decision on a comparison of a number of alternatives. If you start with a short initial search criterion, you are more likely to get these alternatives.

For example, in the larceny/theft series, we decided that we did not want a pattern description containing any segment with fewer than 12 observations. However, an initial search criterion of 12 found a starting partition with only two segments. On the other hand, an initial search criterion of eight results in a partition with five segments. Figure 7 shows the results of the exhaustive search that began with this five segment partition.

The best five segment pattern description contains a segment with only eight observations (figure 7a). This short segment appears also in the best four segment description (figure 7b). However, the best three segment fit does not contain any segment shorter than our criterion of 12 observations. It is less accurate, according to the comparative SSR's and the position on the C_p plot (figure 7f), than the four or five segment description, but it is more accurate than the two segment description, and it meets the criterion for minimum segment length. Unless we have some substantive reason to change that criterion, the three segment description (figure 7c) seems to be the best.

Different Initial and Final Search Criteria

Figure 7a
(Five Segment Line)

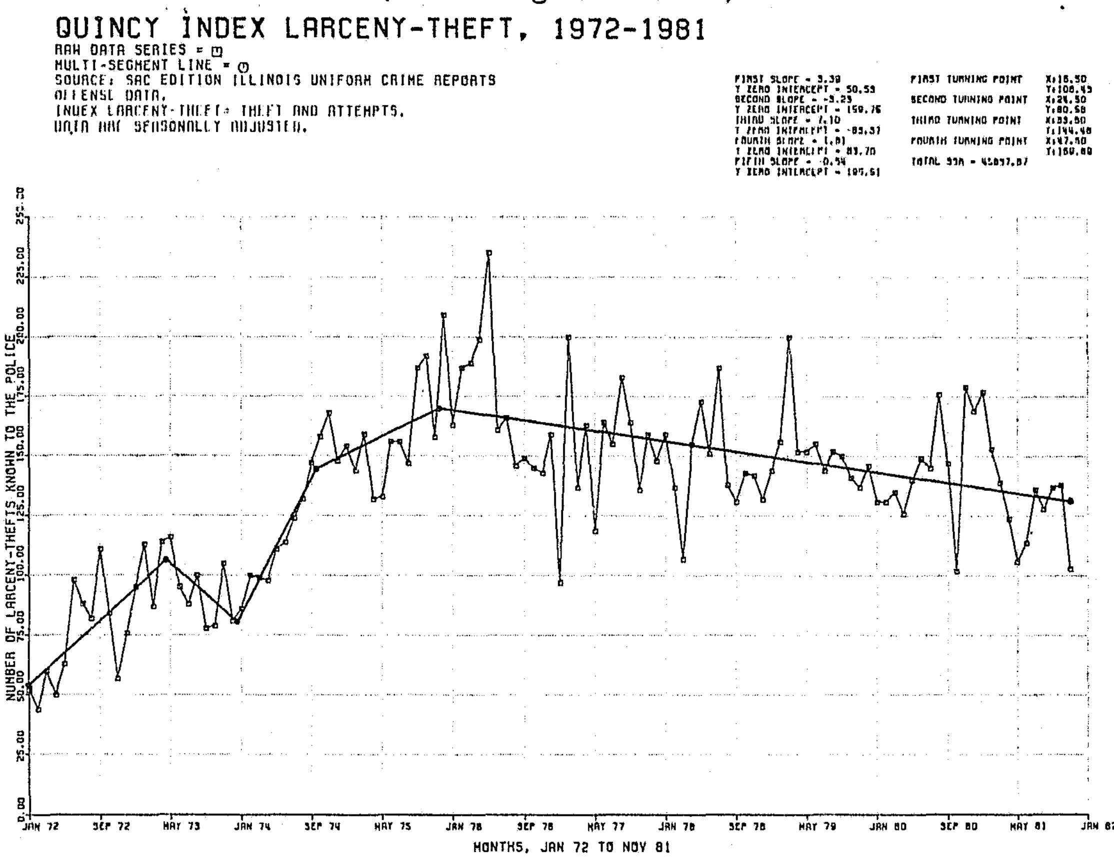
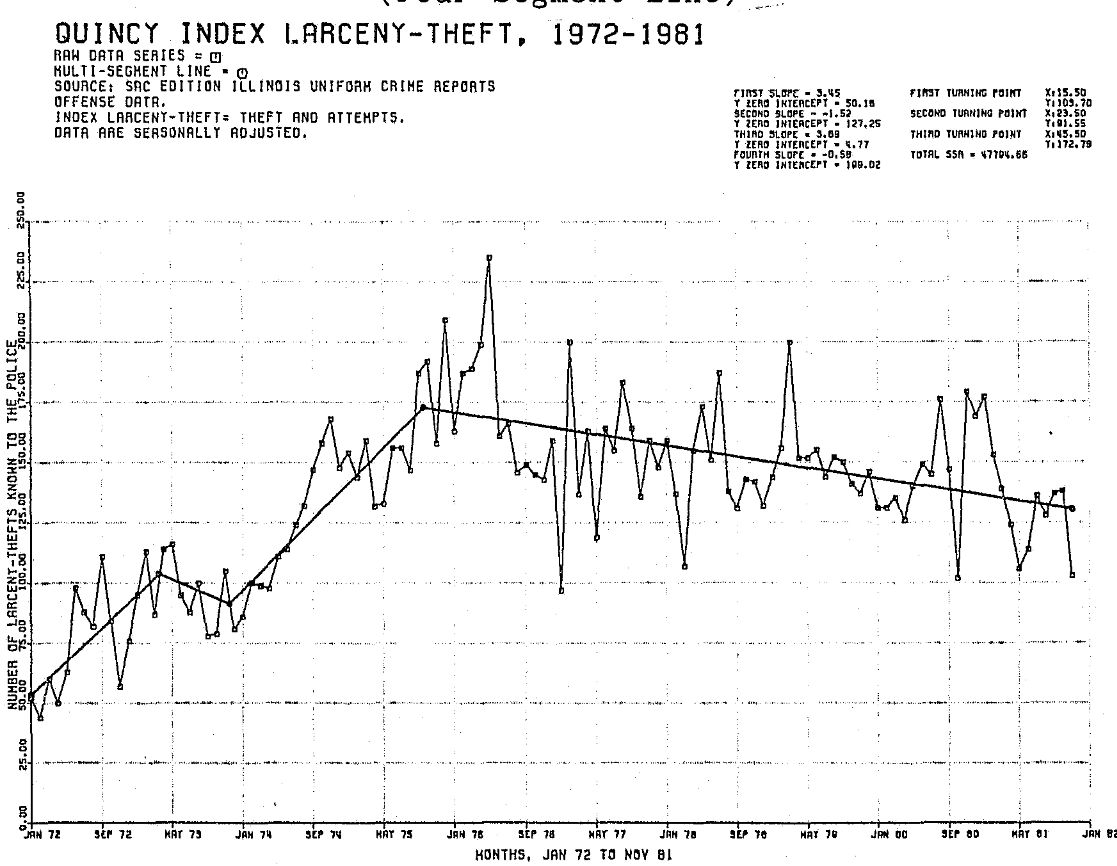


Figure 7b
(Four Segment Line)



Different Initial and Final Search Criteria (Cont.)

Figure 7c
(Three Segment Line)

QUINCY INDEX LARCENY-THEFT, 1972-1981

RAW DATA SERIES = ⊡
MULTI-SEGMENT LINE = ⊙
SOURCE: SAC EDITION ILLINOIS UNIFORM CRIME REPORTS
OFFENSE DATA.
INDEX LARCENY-THEFT= THEFT AND ATTEMPTS.
DATA ARE SEASONALLY ADJUSTED.

FIRST SLOPE = 1.56
Y ZERO INTERCEPT = 64.07
SECOND SLOPE = 3.40
Y ZERO INTERCEPT = 17.16
THIRD SLOPE = -0.58
Y ZERO INTERCEPT = 187.55

FIRST TURNING POINT X:25.50 Y:103.95
SECOND TURNING POINT X:45.50 Y:172.03
TOTAL SSR = 48736.37

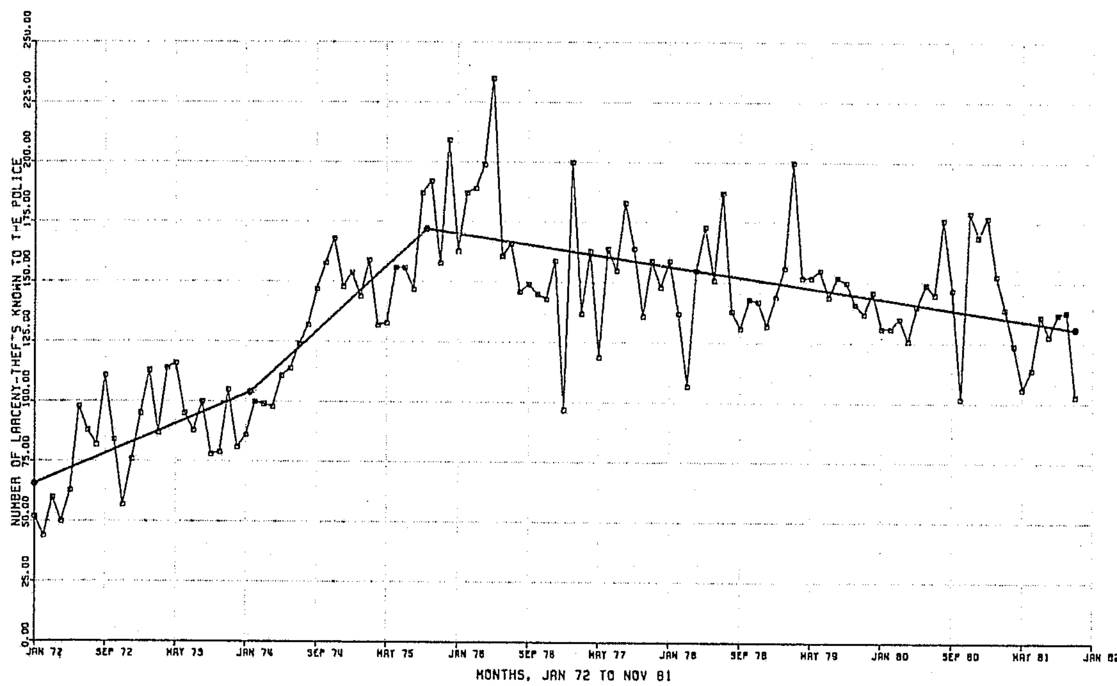


Figure 7d
(Two Segment Line)

QUINCY INDEX LARCENY-THEFT, 1972-1981

RAW DATA SERIES = ⊡
MULTI-SEGMENT LINE = ⊙
SOURCE: SAC EDITION ILLINOIS UNIFORM CRIME REPORTS
OFFENSE DATA.
INDEX LARCENY-THEFT= THEFT AND ATTEMPTS.
DATA ARE SEASONALLY ADJUSTED.

FIRST SLOPE = 2.45
Y ZERO INTERCEPT = 52.66
SECOND SLOPE = -0.59
Y ZERO INTERCEPT = 194.10

FIRST TURNING POINT X:47.50 Y:169.21
TOTAL SSR = 51661.59

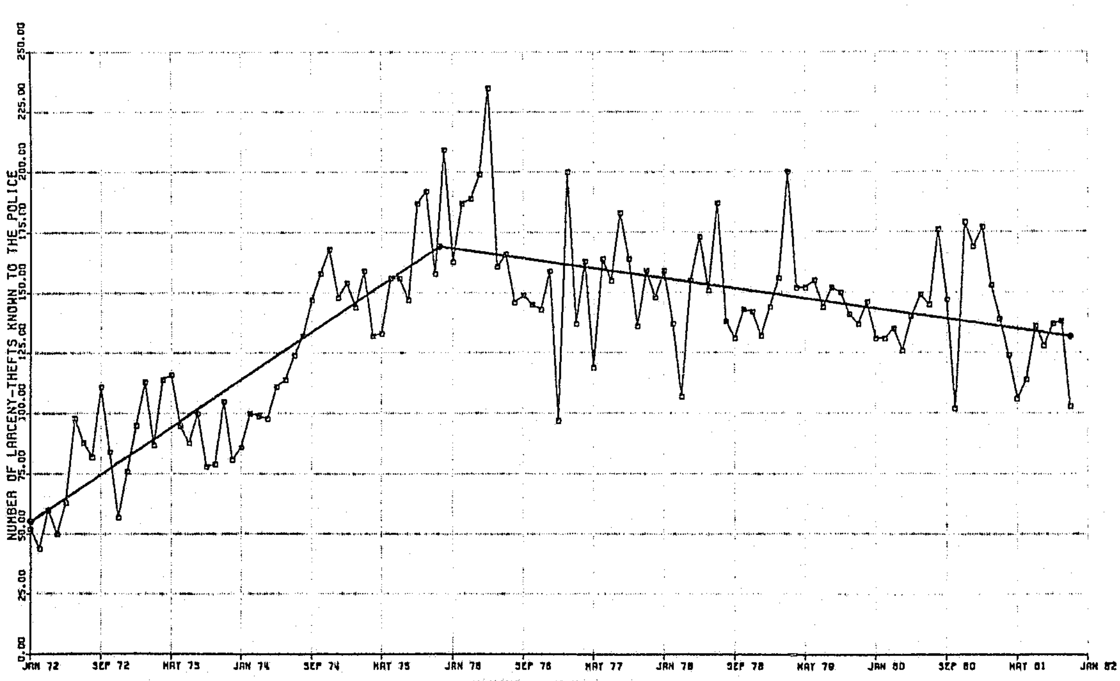


Figure 7e
(One Segment Line)

QUINCY INDEX LARCENY-THEFT, 1972-1981

RAW DATA SERIES = □
MULTI-SEGMENT LINE = ○
SOURCE: SAC EDITION ILLINOIS UNIFORM CRIME REPORTS
OFFENSE DATA.
INDEX LARCENY-THEFT- THEFT AND ATTEMPTS.
DATA ARE SEASONALLY ADJUSTED.

FIRST SLOPE = 0.50
Y ZERO INTERCEPT = 104.00
TOTAL SSR = 119684.74

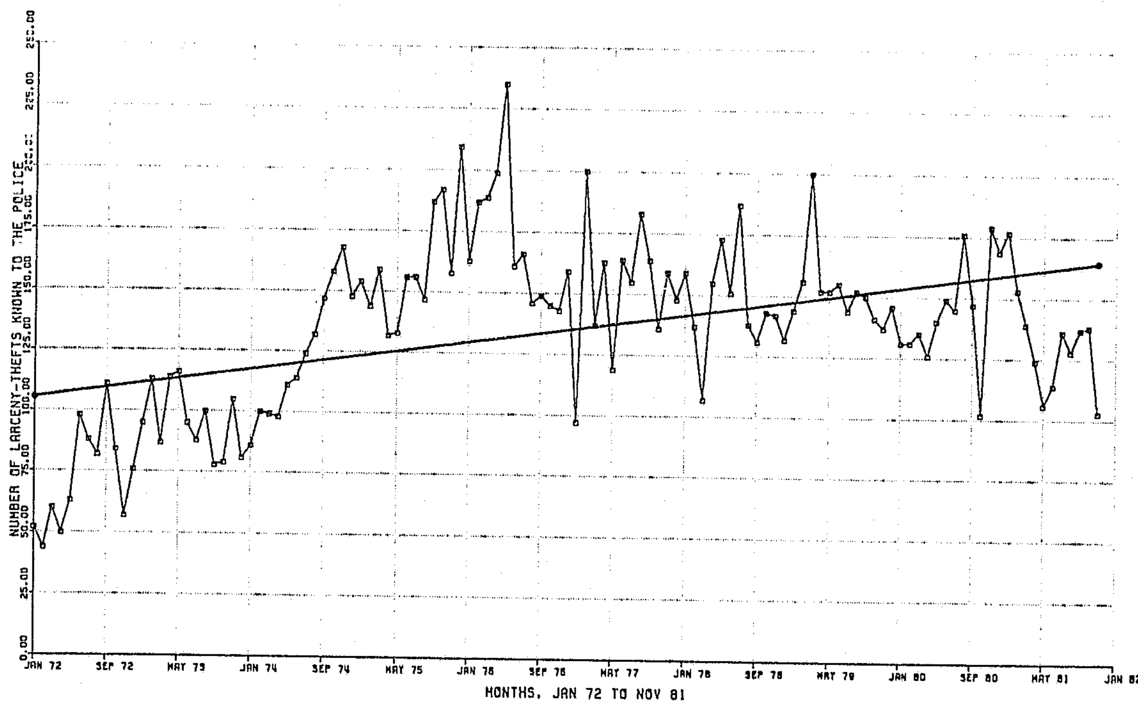
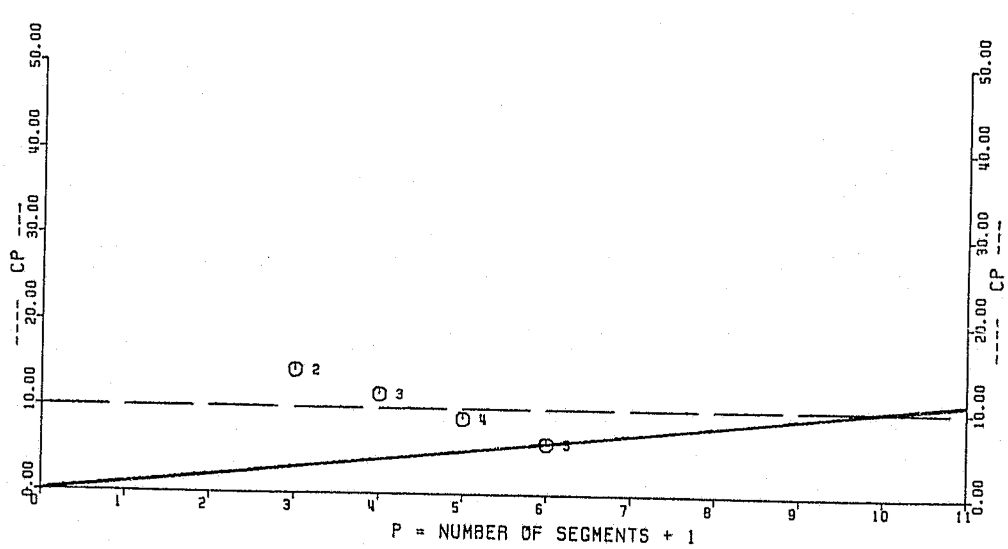


Figure 7f
(C_p Plot)

QUINCY INDEX LARCENY-THEFT, 1972-1981

SUMMARY: CP BY P

NUMBER OF SEGMENTS IN INITIAL SEARCH = 5
CP. 1 SEGMENT FIT = 180.0 * OFF CHART
CP. 2 SEGMENT FIT = 14.4
CP. 3 SEGMENT FIT = 11.6
CP. 4 SEGMENT FIT = 8.8
CP. 5 SEGMENT FIT = 6.0



Comparing Patterns of Two Series

A common use for time series pattern description involves comparing the pattern in one series to the pattern in another. In response to such a request, we compared the pattern of change in the number of Index burglary offenses known to the police from 1972 to 1980 in Champaign and Danville, Illinois, which are neighboring cities.

While Champaign generally had more burglaries than Danville, it was the different patterns of change over the nine years that proved to be most interesting. The number of burglaries in Champaign (figure 8a) fluctuated around 80 a month during the early years, from 1972 through early 1977, but then increased rapidly. By 1981, the number of burglaries in a typical month was approaching 140. In comparison, burglaries in Danville (figure 8b) increased in the early years, from 1972 through 1974, but then decreased. Instead of increasing rapidly after mid-1977, Danville burglaries increased only slightly.

In comparing the patterns of two series, be sure that the scales of the two graphs are the same. Otherwise, what seems to be a "rapid" increase on one scale may seem to be a "slight" increase on the other scale. Consider whether you want to use raw numbers, or rates per capita. If you want to compare crime in a single year in two places, it is better to use rates, but if you want to compare patterns over time, it may be better to use raw numbers. (See the section, "Rates versus Raw Numbers," page 50 below.) In this case, the state's attorney wanted to know the pattern over time of the number of burglaries. The patterns of the rates would not have answered his question.

Figure 8a

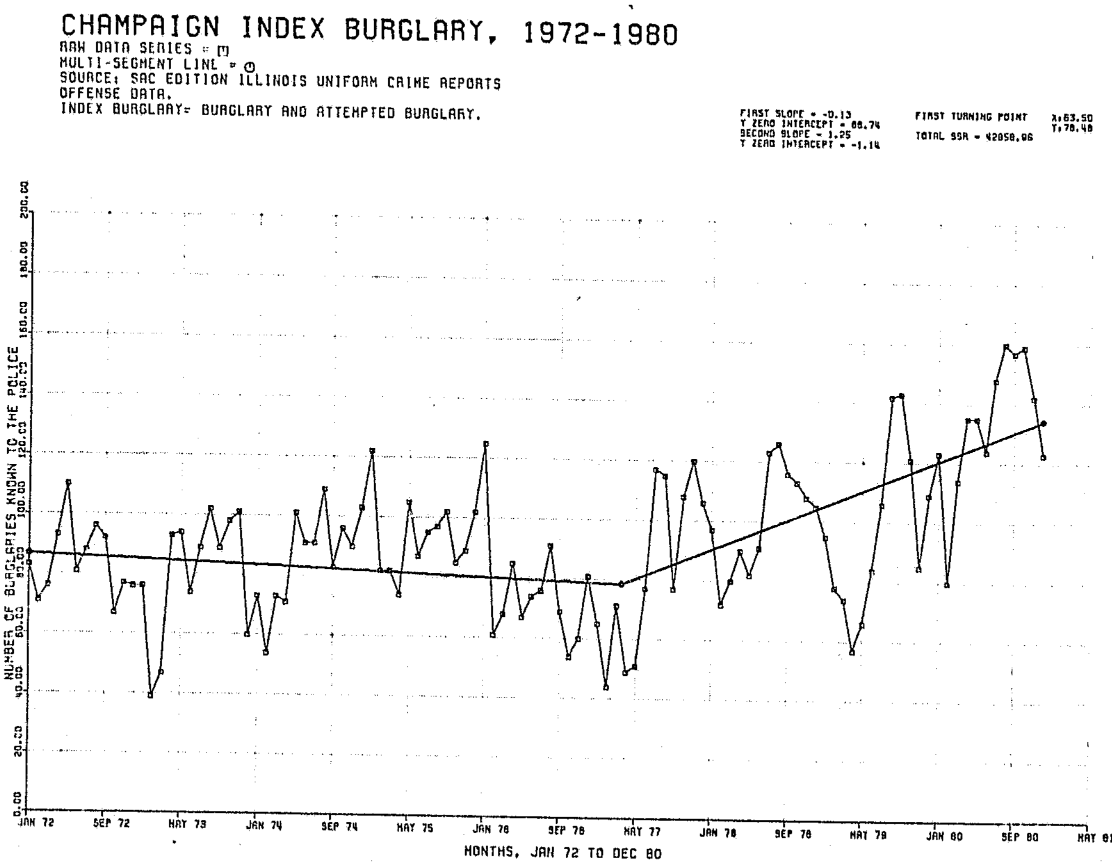
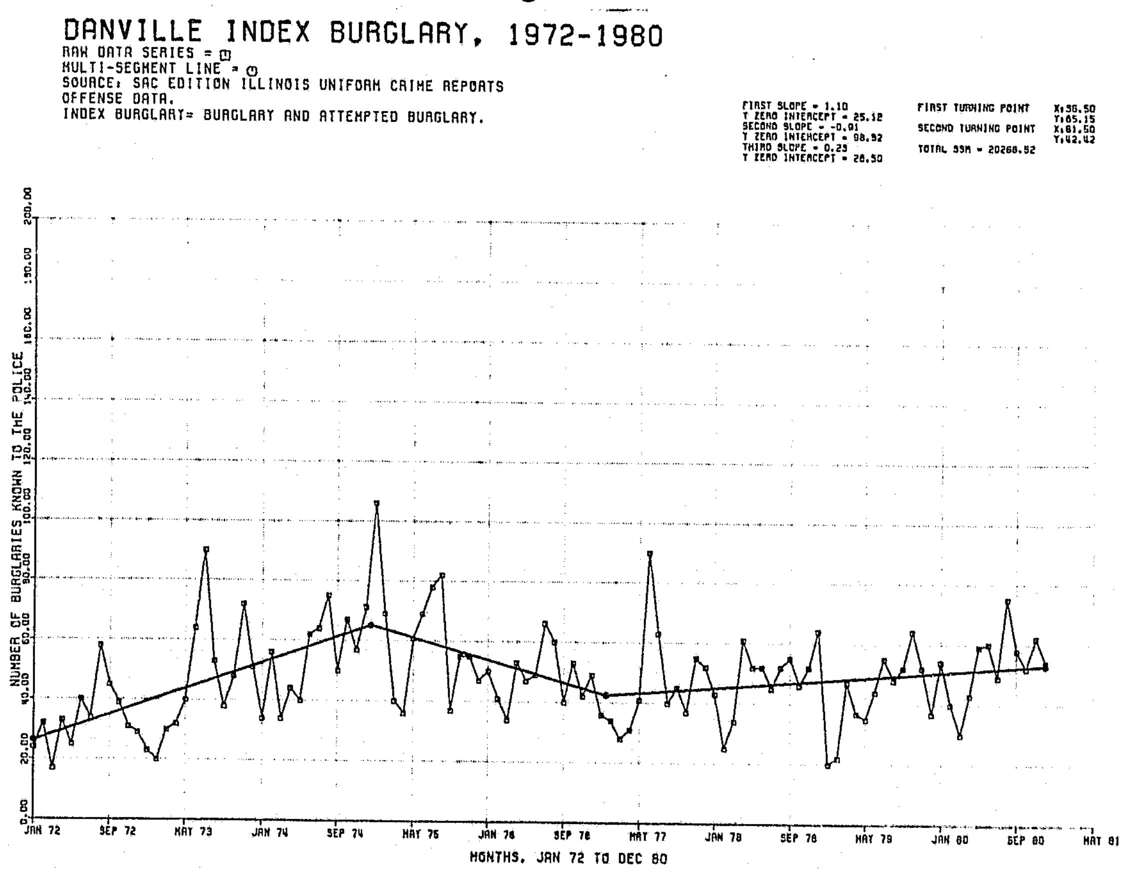


Figure 8b



Time Series Specification

Time series specification divides a time series into components, and compares the patterns of the components to the pattern of the whole. If the pattern of one component is similar to the pattern of the whole, and the pattern of the other component is different, then the pattern of the whole has been specified. That is, the pattern of change over time in that component accounts for the pattern of change over time in the whole. Time series specification is similar to cross-sectional specification (Selvin,1972; Kendall and Lazarsfeld,1950,1955; Hyman,1955) in that it defines the conditions under which a phenomenon occurs. In cross-sectional specification, that phenomenon is a certain association between two variables. In time series specification, that phenomenon is a pattern of change over time.

For example, the best description of the pattern of Chicago homicides from 1965 through 1976 is a three segment line that increases rapidly from 1965 to 1970, levels off from 1970 to 1974, and decreases rapidly in 1975 and 1976.[32] One possible explanation for this pattern is that it was caused by a similar pattern of change in the number of homicides committed by young people. If this demographic explanation is true, then the pattern of homicides committed by young people should be similar to the pattern of total homicides, and the pattern of homicides committed by other age groups should be different. However, this is not the case. The patterns of homicides by youthful offenders and homicides by other age offenders are more similar to each other than they are to the pattern of the total homicide series. Therefore, change in the number of homicides attributed to youthful offenders cannot explain change in the number of total homicides. The pattern of homicides is not specified by age of the offender.

Having rejected offender's age, we considered other possible specification variables, including race and sex of victim and offender, victim-offender relationship, precipitating event, and weapon. Only weapon specified the homicide pattern (figure 9). Homicides with a gun followed almost exactly the same pattern between 1965 and 1976 as total homicides. Homicides without a gun followed a completely different pattern. The pattern of change over time of Chicago homicides is a conditional one: it occurs only in homicides with a gun. Specification, of course, does not explain the cause of change in Chicago homicides, but it narrows the search for an explanation. We now know that, to explain the pattern of change in Chicago homicides, we must first explain the pattern of change in homicides with a gun.

---

[32]For details of this analysis, see the SAC publication, "Patterns of Change in Chicago Homicide: The Twenties, the Sixties, and the Seventies."

Example of Time Series Specification

Figure 9a

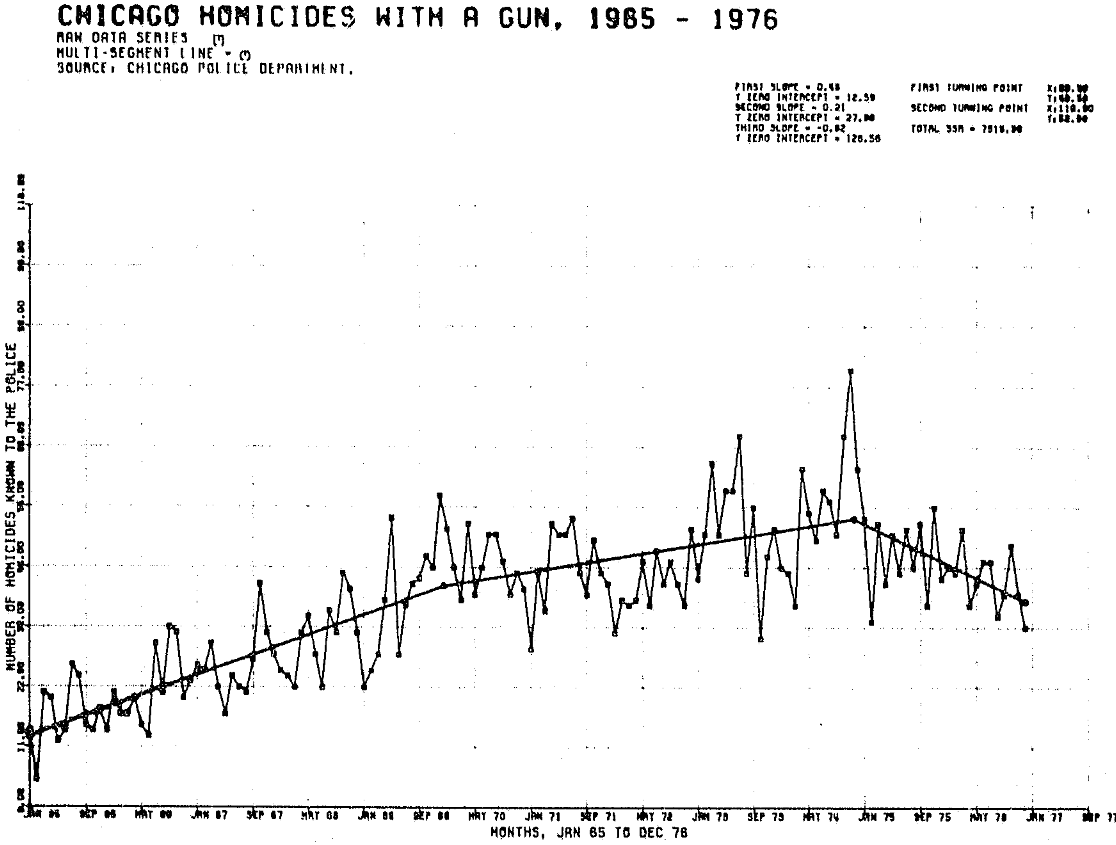
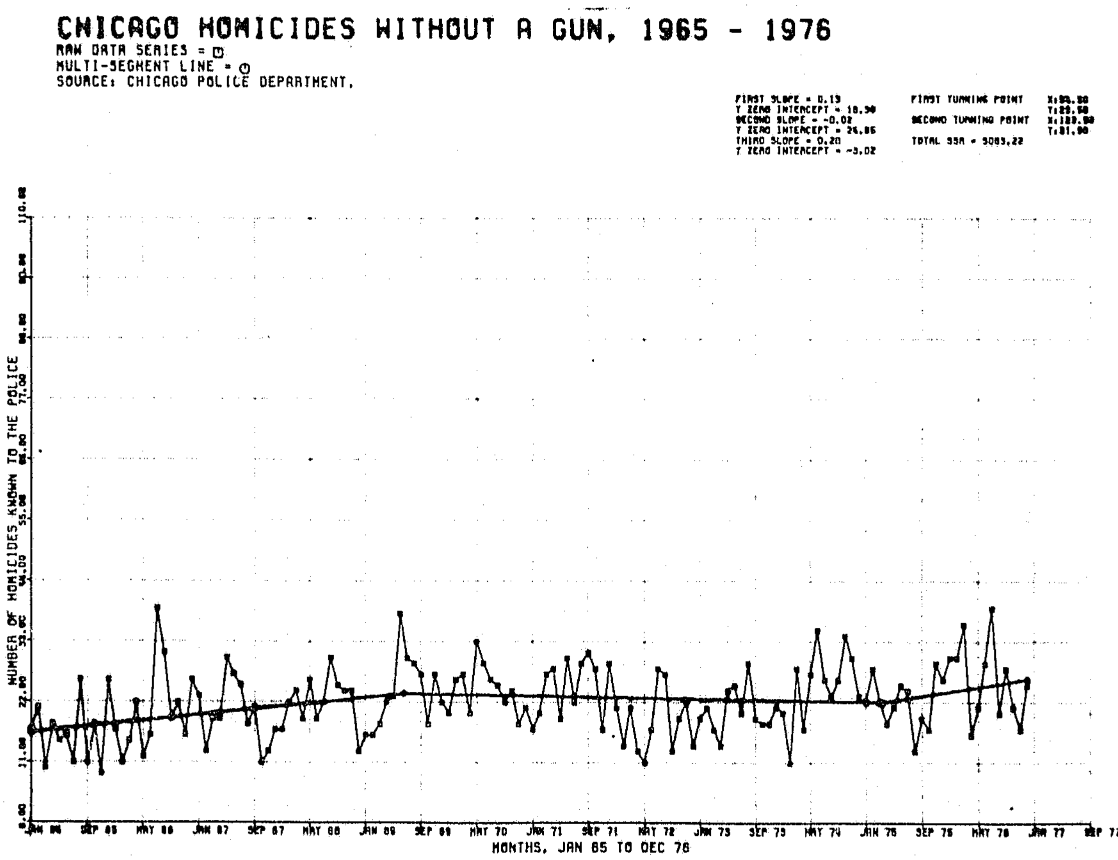


Figure 9b



Simplicity versus Accuracy

A good example of the choice of simplicity versus accuracy in a pattern description occurs with the series of homicide rates from 1940 to 1977 in the East South Central states of the United States. Figure 10a, a seven segment fit, is a more accurate and more detailed description than figure 10b, a two segment fit. The seven segment fit follows the dip during World War II closely, and also shows a decrease in the final years, but the two segment fit shows none of this detail. The greater accuracy of the seven segment fit is indicated by its smaller SSR (6.86 versus 37.26 for the two segment fit), and by its C_p. The C_p for the seven segment fit is 8.0, but for the two segment fit, it is 130.8. If accuracy were the only consideration, then, clearly, the seven segment description would be the better choice.

However, accuracy is usually not the only consideration. The two segment description has a clear advantage in simplicity. Although it does not follow every dip of the raw data, the two segment fit gives the reader a simple description of the overall pattern of change over time. The homicide rate generally decreased in the 1940s and 1950s, and increased in the 1960s and 1970s. If this rough pattern description is all that the reader needs, the seven segment fit would be superfluous. The seven segment fit is so detailed that it can hardly be called a simple description. It is not really simpler than the raw data, for practical purposes.

Figure 10a
(Seven Segment Line)

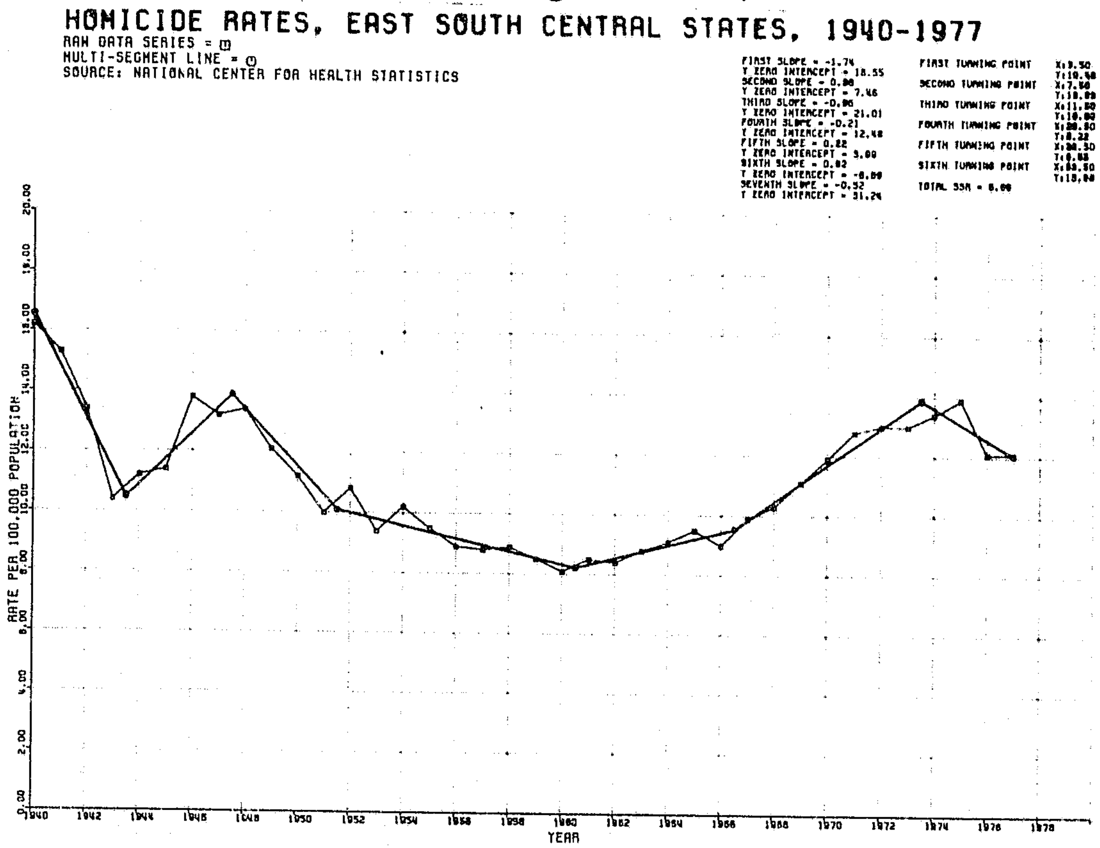
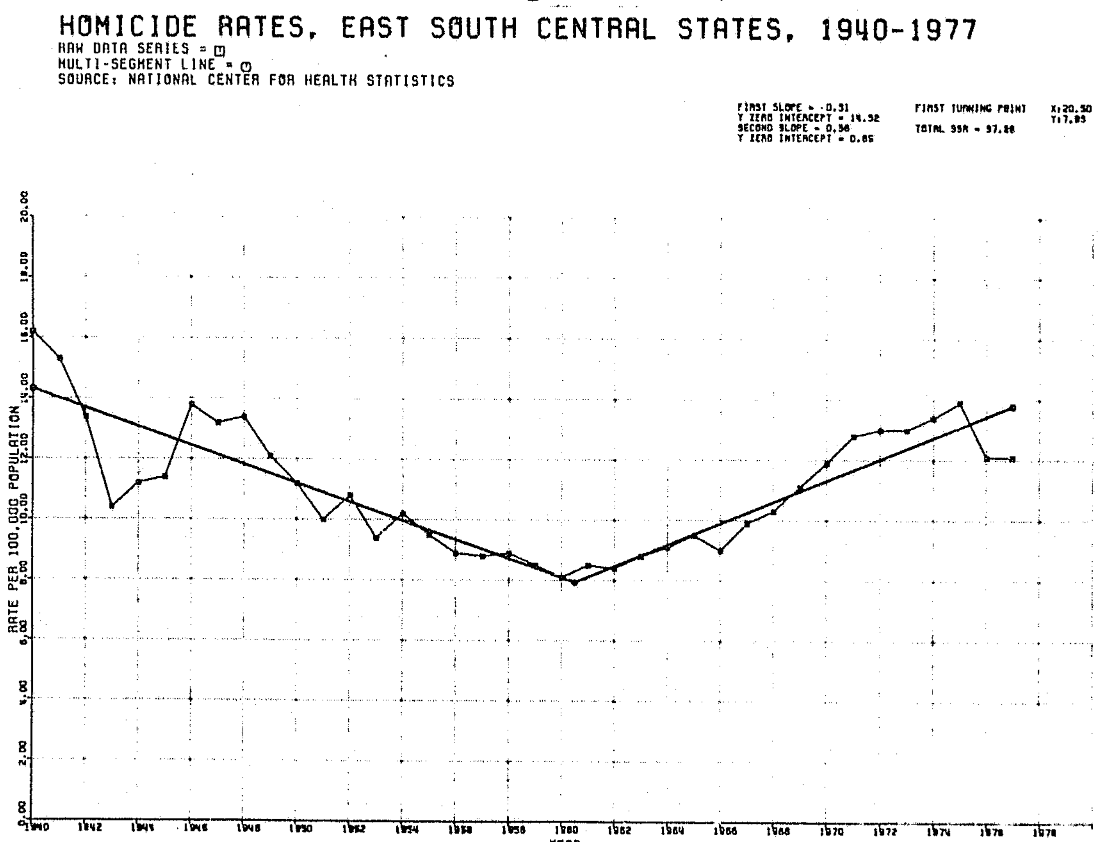


Figure 10b
(Two Segment Line)



Fewer Segments may be More Accurate

It is not always true that a given line segment pattern description with more segments will be more accurate than an alternative pattern description with fewer segments. Here we have two examples.

The best pattern description for the homicide rate from 1950 to 1977 in United States metropolitan counties, given the criterion that no segment will be shorter than four years, is a four segment line (figure 11a). The five segment alternative (figure 11b) has a higher total SSR. Similarly, a three segment fit describes Canadian homicides attributed to a stranger better than a four segment fit (figures 12a and 12b).

The C_p plots for the two sets of alternative descriptions (figures 13a and 13b) illustrate this. The four segment description for metropolitan homicide is more accurate than the three or five segment description, and the three segment description for Canadian homicides by a stranger contains both less random error and less bias than the four segment description.

Figure 11a
(Four Segment Line)

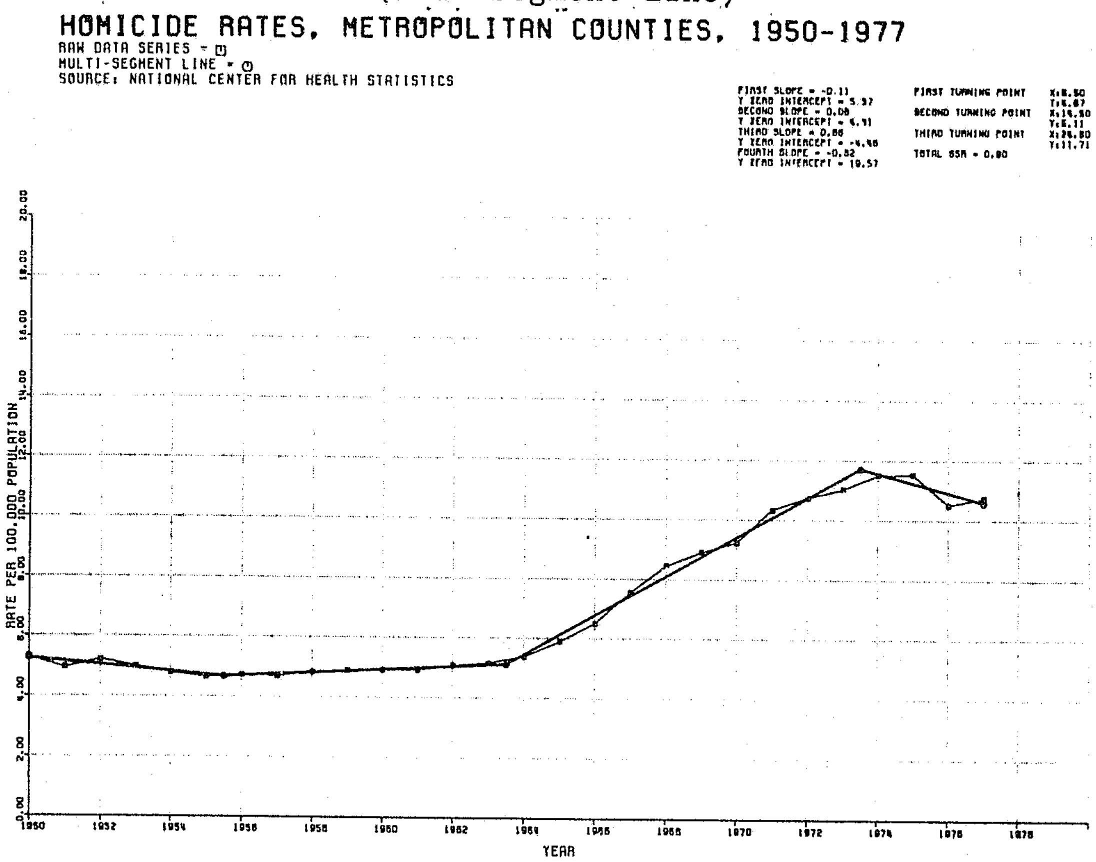
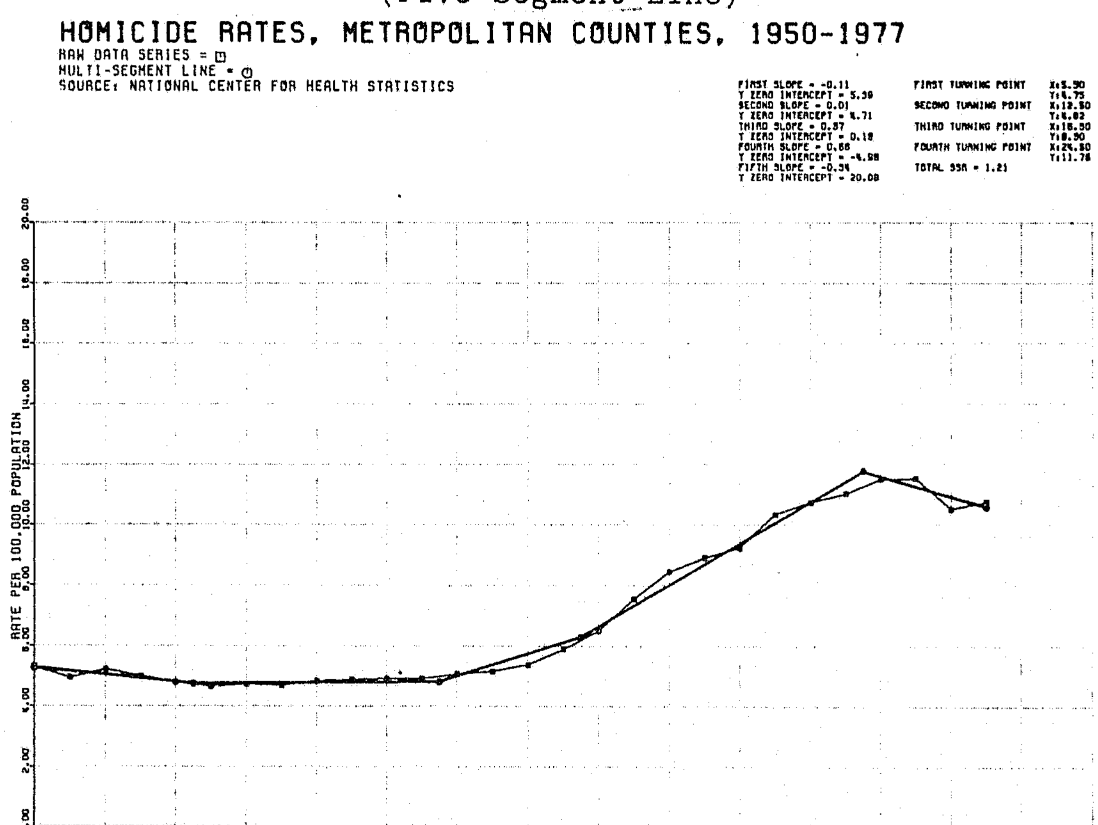


Figure 11b
(Five Segment Line)



Canadian Homicides Attributed to a Stranger

Figure 12a
(Three Segment Line)

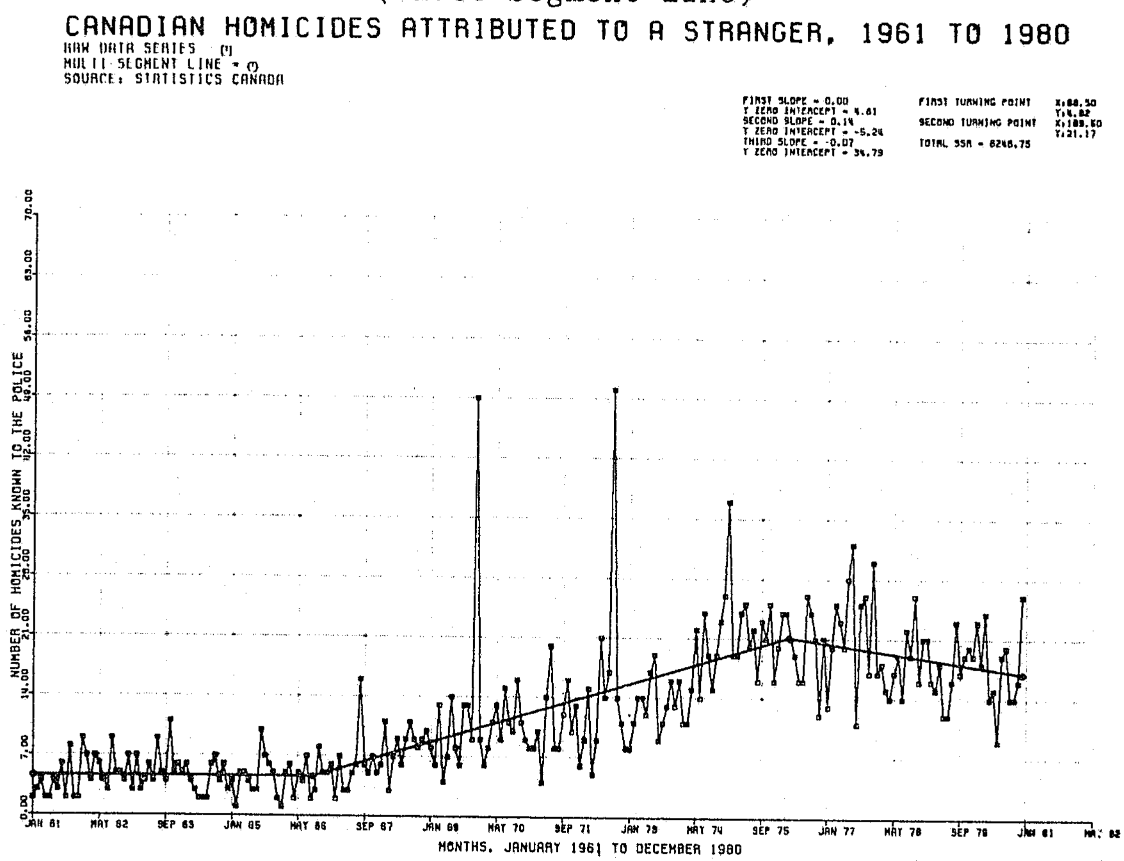
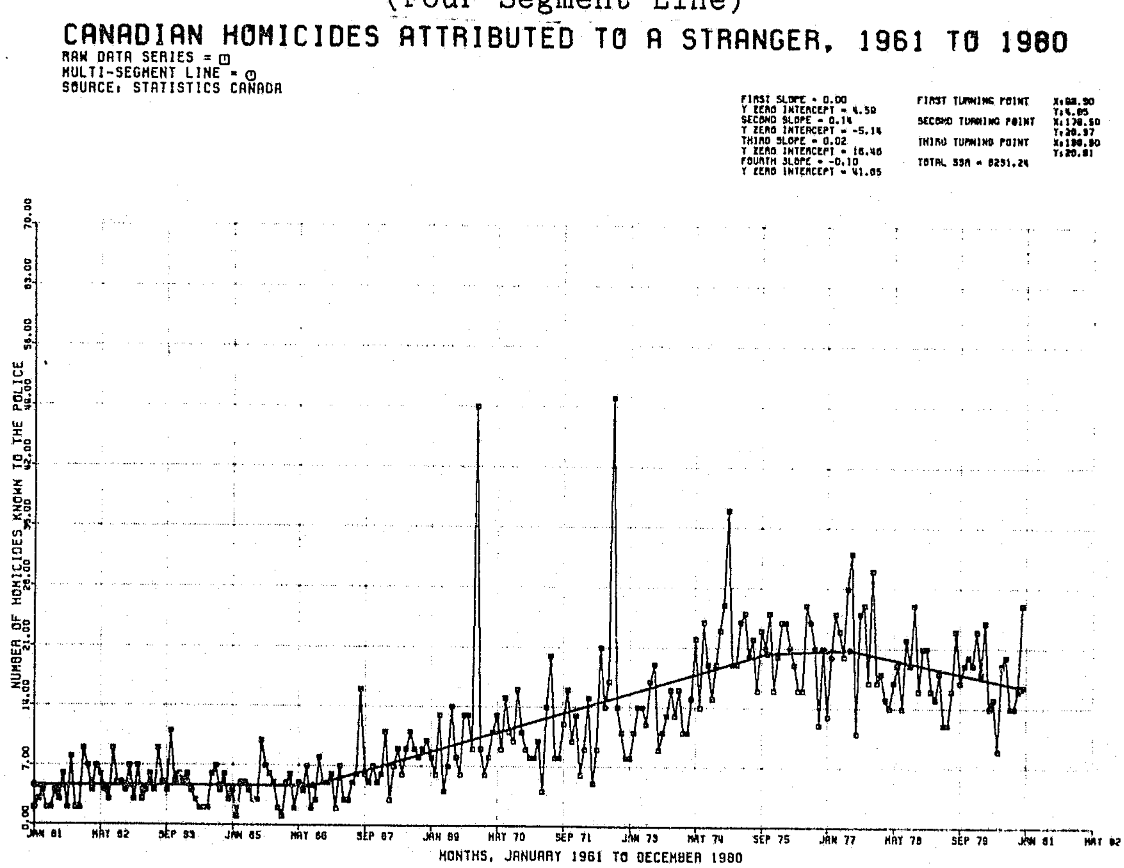


Figure 12b
(Four Segment Line)



C_p Plots for United States Metropolitan County Homicides and Canadian Homicides Attributed to a Stranger

Figure 13a

HOMICIDE RATES, METROPOLITAN COUNTIES, 1950-1977

SUMMARY: CP BY P

NUMBER OF SEGMENTS IN INITIAL SEARCH = 5
CP. 1 SEGMENT FIT = 901.8 * OFF CHART
CP. 2 SEGMENT FIT = 200.7 * OFF CHART
CP. 3 SEGMENT FIT = 14.2
CP. 4 SEGMENT FIT = 5.0
CP. 5 SEGMENT FIT = 15.1

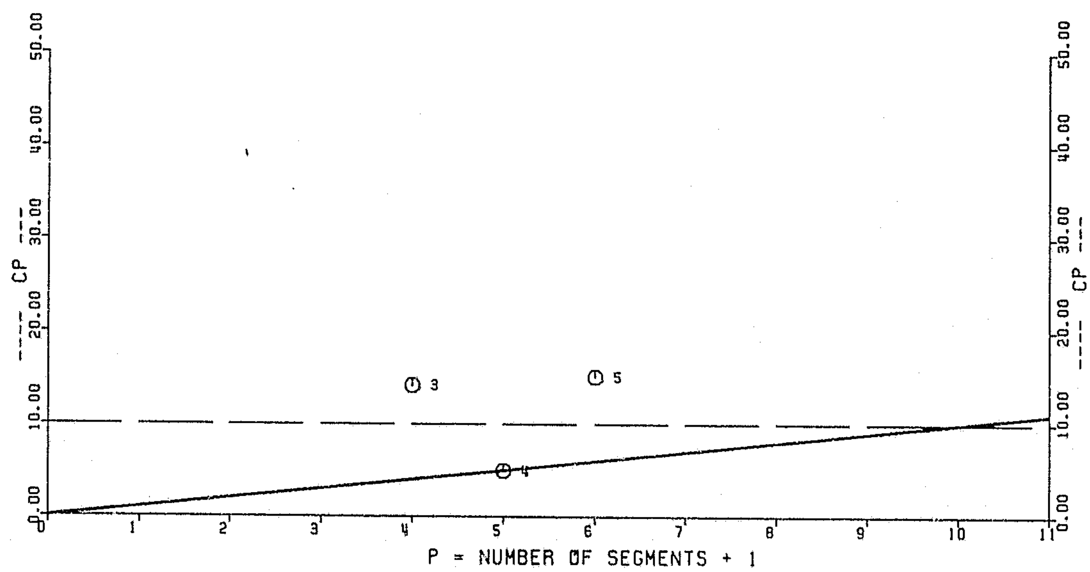
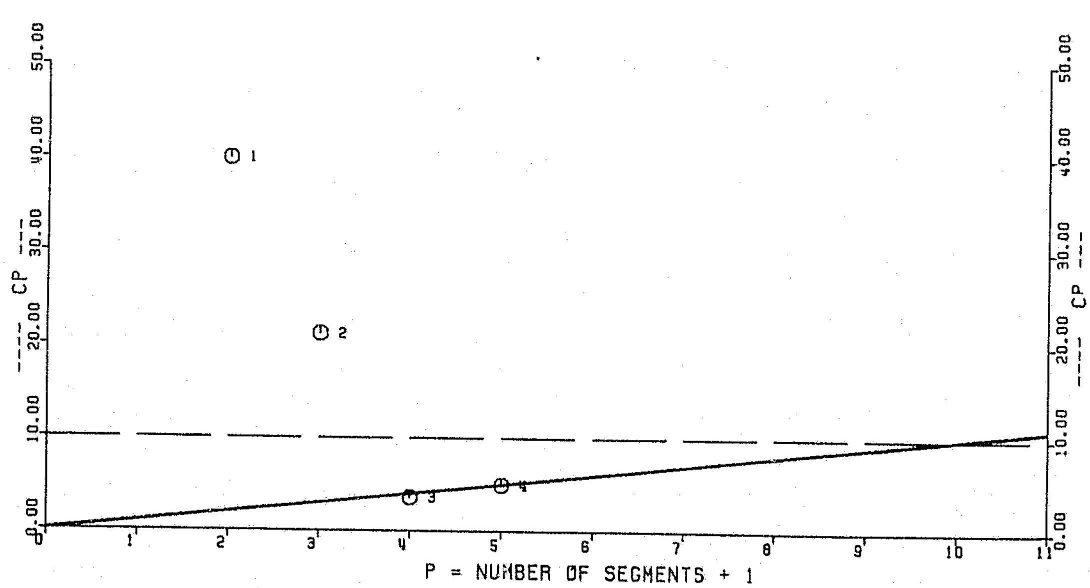


Figure 13b

CANADIAN HOMICIDES ATTRIBUTED TO A STRANGER, 1961 TO 1980

SUMMARY: CP BY P

NUMBER OF SEGMENTS IN INITIAL SEARCH = 4
CP. 1 SEGMENT FIT = 40.0
CP. 2 SEGMENT FIT = 21.2
CP. 3 SEGMENT FIT = 3.6
CP. 4 SEGMENT FIT = 5.0



Short Line Segment

Although a pattern description containing short line segments is more complex than alternative descriptions with longer segments, allowing for a short segment in the description may be important in some practical situations.

For example, figures 14a and 14b are the best six segment and five segment descriptions for the number of people released each month from the Illinois Department of Corrections (IDOC). The main difference between the two is the short fifth segment, only 11 months long, in figure 14a.[33] If such a short line segment contributed no interesting information, it would be better to choose the simpler alternative with fewer and longer segments (figure 14b). In this instance, however, the IDOC was interested in a pattern description that would include changes, if any, that occurred less than a year apart, because it had an "Early Release" program that temporarily stopped during those months.

---

[33]As a general rule, data that may be seasonal, such as this monthly series, should be analyzed for the presence of seasonality before using a pattern description with a segment less than a year long. According to our analysis, using the U.S. Bureau of the Census X-11 program, IDOC releases are not seasonal. See Miller (1983).

People Released, Illinois Department of Corrections

Figure 14a
(Six Segment Line)

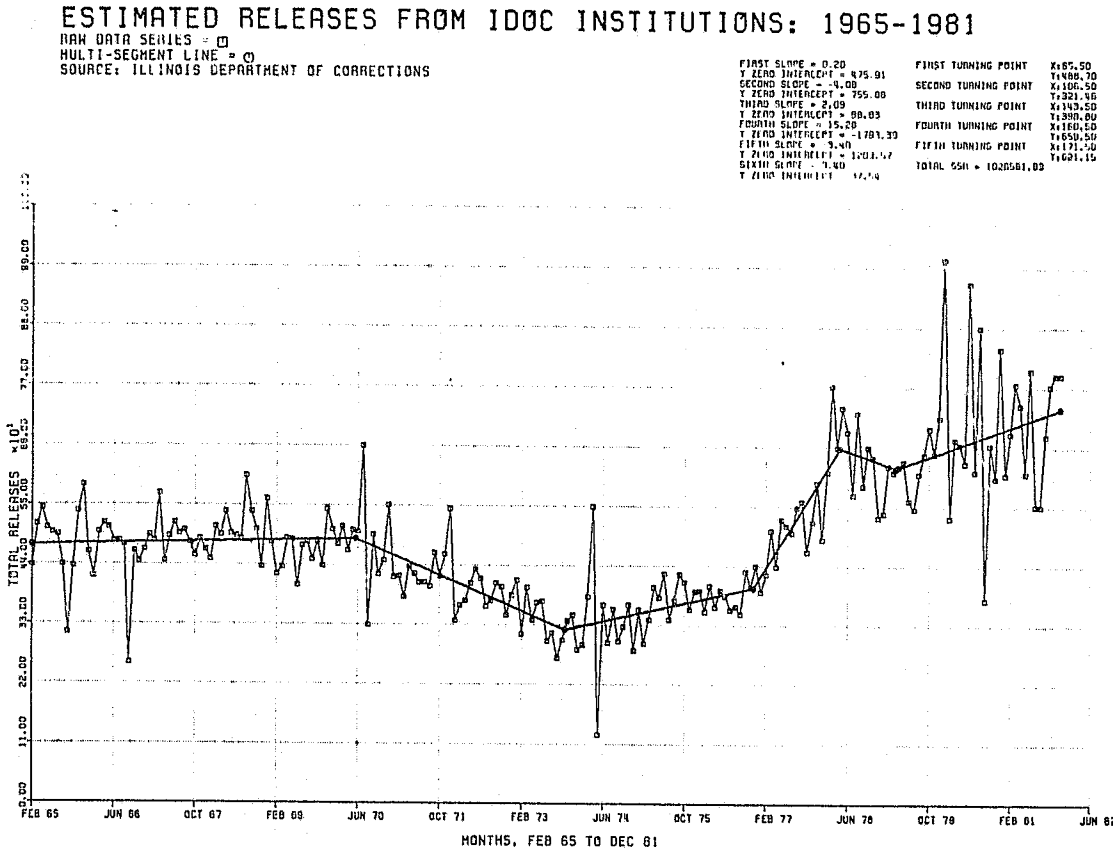
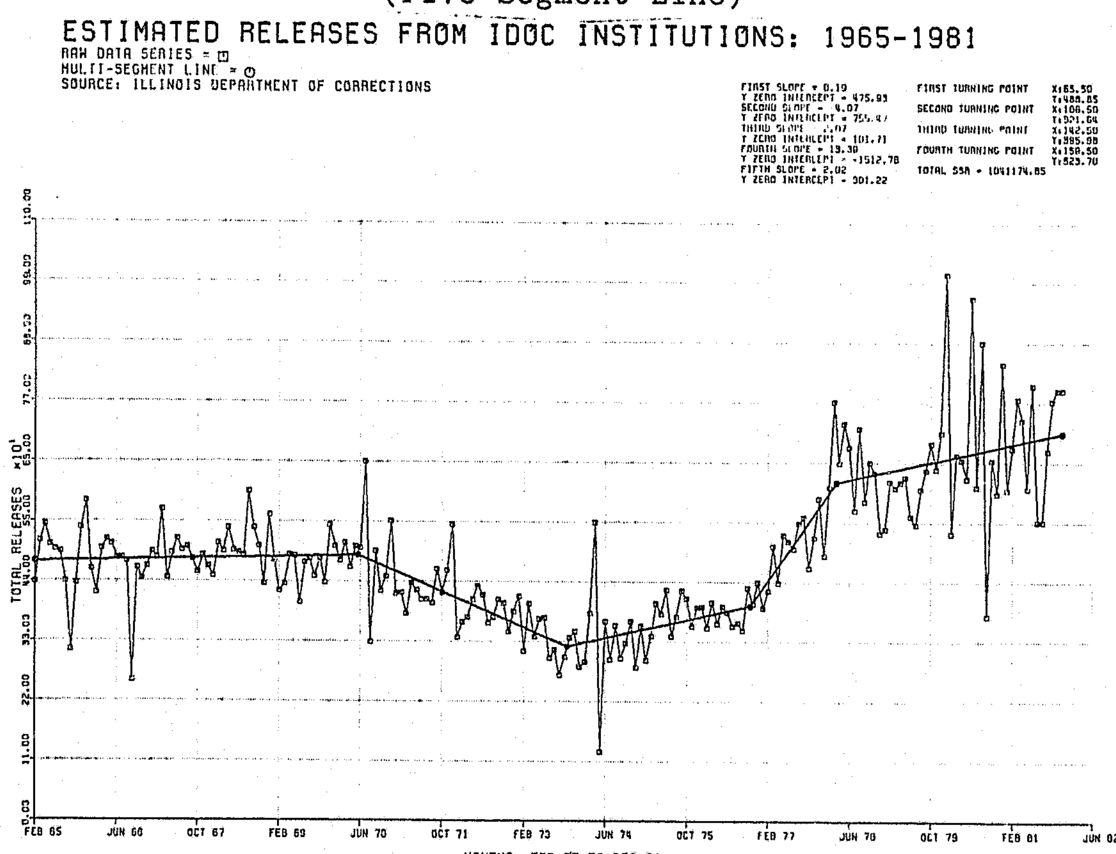


Figure 14b
(Five Segment Line)



Rates versus Raw Numbers

A rate is a complex number constructed by dividing the number of occurrences by some total figure that represents the opportunity for those occurrences to take place. A rate standardizes a figure against a base of comparison. For example, the number of reported robberies per capita (the robbery rate) standardizes the number of reported crimes against the size of the population. If we want to compare two cities, one more populous than the other, the robbery rate would provide a fairer basis for comparison than the number of robberies. The most common crime rate has the number of reported crimes in the numerator and the population in the denominator, but there are countless variations. The number of reported auto thefts could be standardized against the number of registered autos. The familiar "unemployment rate" is the number of unemployed people standardized against the number in the labor force.

Analysts use rates almost automatically to compare one place to another or one population to another. However, a rate can be misleading in comparing the same place and population over time, because it is affected by changes in both the numerator and the denominator. We have no way of knowing whether the pattern of change over time in the rate is due to change in the numerator, change in the denominator, or both. If the crime rate increases, is the increase due to more crimes or to fewer people in the population?[34] If the unemployment rate decreases, is the decrease due to fewer unemployed or to more people dropping out of the labor force? The only way to answer these questions is to examine the patterns over time of the raw data, not only the pattern of the rate. Therefore, in general, pattern descriptions of rates should be done only after pattern descriptions of the variables that make up the rate.

The patterns of arrests and arrest rates in two census tracts of Racine, Wisconsin, illustrate this. The arrest rate in Tract 1 increased rapidly in the late 1960s, while the arrest rate in Tract 6 hardly changed (figure 15a). If this were our only information, we might be tempted to conclude (erroneously) that Tract 1 arrests increased and Tract 6 arrests were stable in the 1970s. Actually, the opposite is closer to the truth, as figure 15b shows. The pattern of the arrest rate reflects not only the pattern of arrests, but also change over time in population within each tract.

---

[34]In addition, population data are often available only in 10-year increments, while crime data are often available in monthly or yearly increments. To compute yearly (or monthly) rates, we must interpolate population figures. The variation over time of the interpolated estimate is artificially smooth within census periods, with artificially abrupt changes every 10 years. This pattern is reflected in the pattern of the rate.

Patterns of Rates versus Patterns of Raw Numbers: Arrests and Arrest Rates in Racine, Wisconsin

Figure 15a

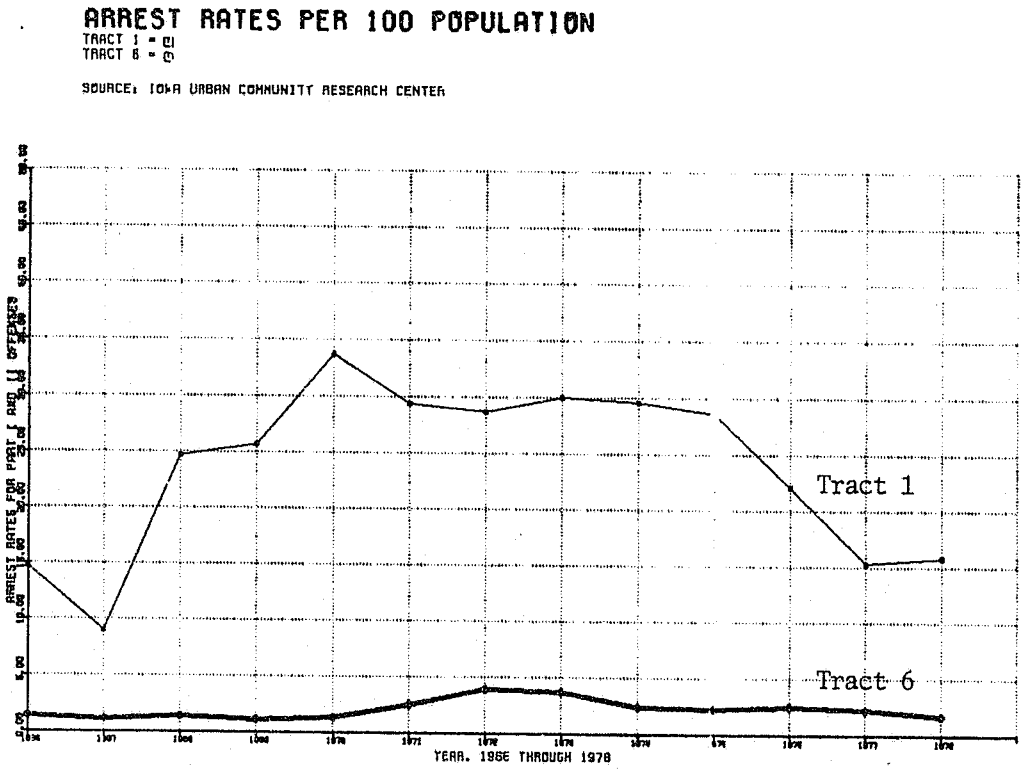
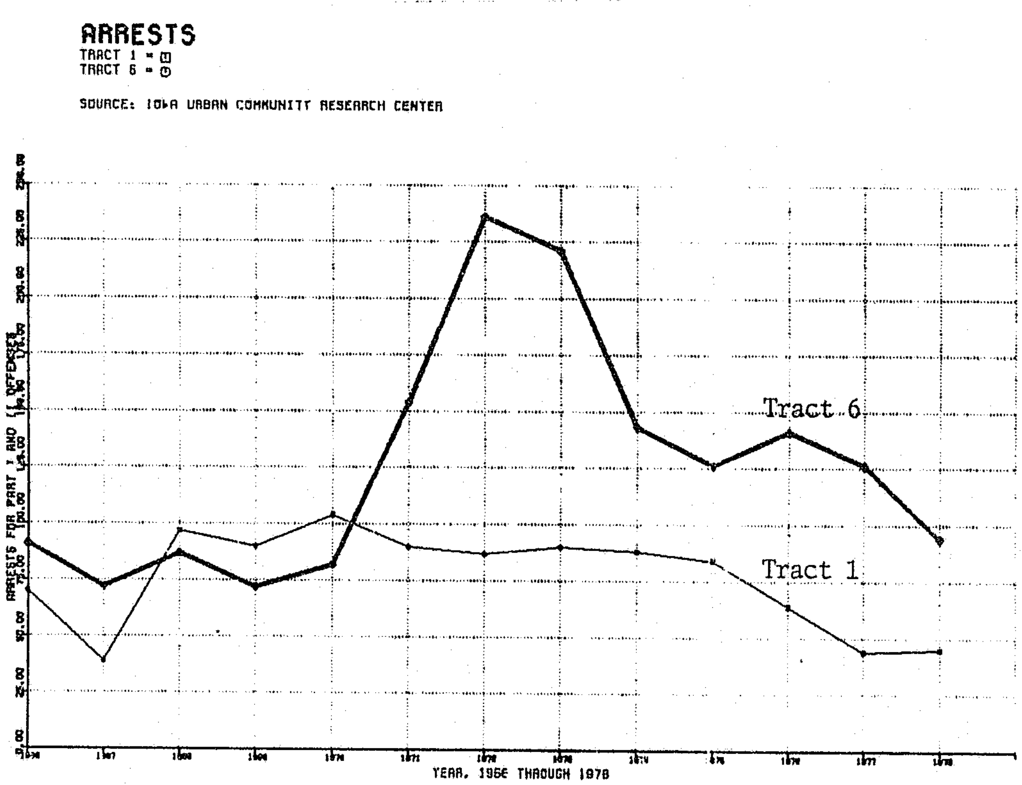


Figure 15b



Series with Extreme Values

As discussed above (see "Appropriate Applications," page 7), the decision as to whether or not to eliminate an extreme value from a series should be based, in part, on the validity of the extreme as a representation of what the series is measuring. Does the extreme represent some unusual event that may never happen again, or is it simply an unusually high or low occurrence of a continuously defined phenomenon? Here, we offer an example of each situation.

The number of firearms reported stolen per month in Illinois excluding Chicago was never more than 700 between 1969 and 1981, except in October 1977, when there were 2500 (figure 16). Upon investigation, we discovered that this extreme was caused by a change in reporting practices. Before October 1977, Cook County suburbs reported their stolen firearms through the Chicago Police Department, and not directly to the Illinois Department of Law Enforcement (I-DLE). Therefore, in excluding Chicago data from this series, we had also excluded Chicago suburban data from the early years. In addition, when the suburbs began to report directly to the I-DLE, all their outstanding (uncleared) lost and

Figure 16

An Extreme that Revealed a Data Definition Problem

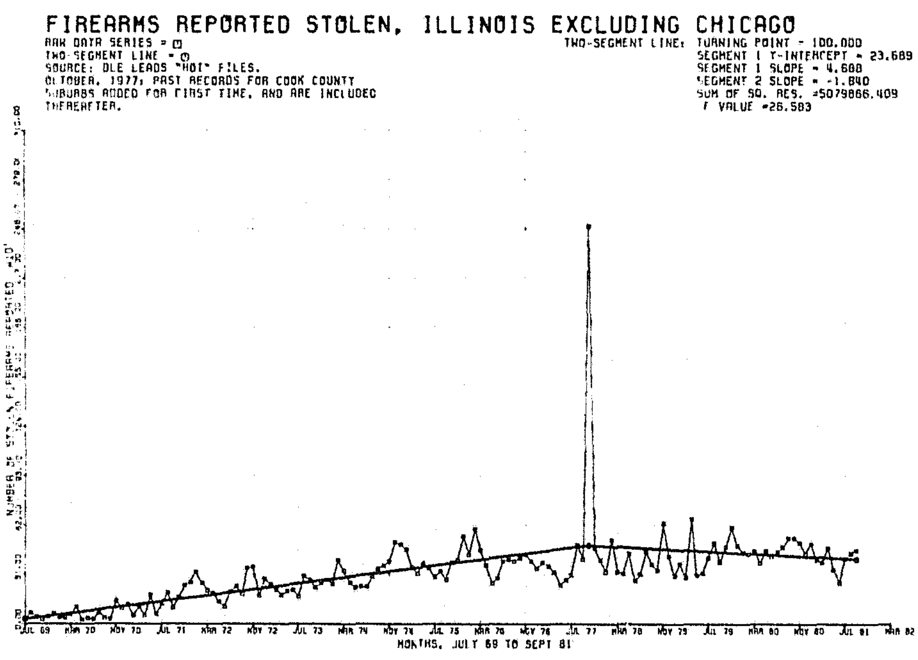
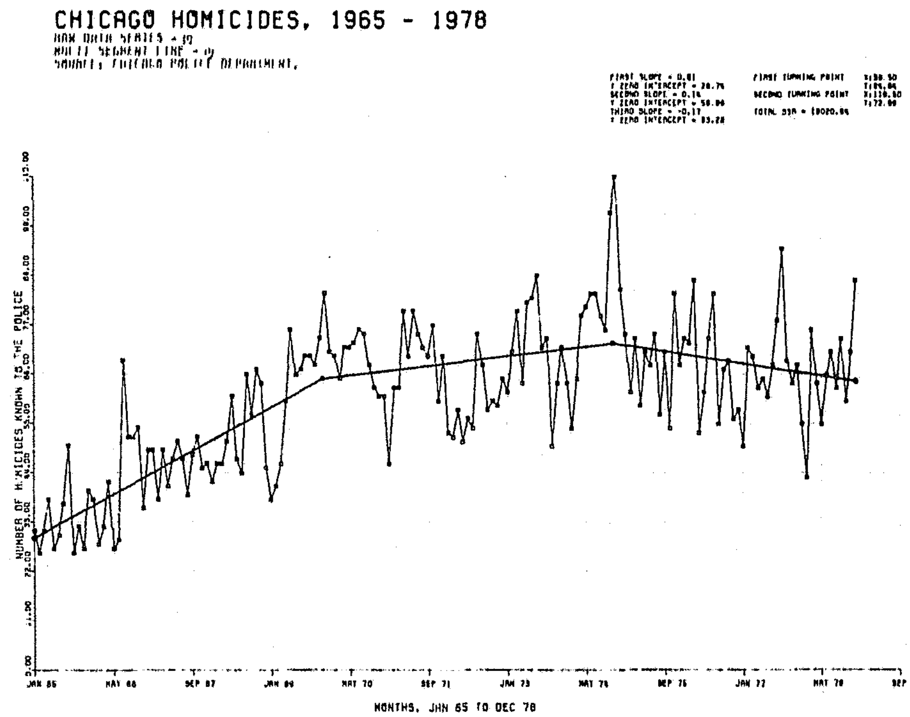


Figure 17

An Extreme that Actually Occurred



stolen firearm records were added to the file in a single month. This produced the extreme value. To produce a series in which the definition of stolen firearms is the same throughout, Cook County suburbs should be separated from the rest of "downstate" Illinois, and the October 1977 extreme should be eliminated.

In our analysis of the pattern of Chicago homicides from 1965 to 1976, we noticed that there were 110 homicides in November 1974, and 102 in October 1974, which were well above the usual number (figure 17). Should we have corrected the series for these extreme months? We considered doing that, but first examined the series carefully to determine whether there was some unusual event, a mass murder perhaps, that could account for the extremes. We searched the Chicago Tribune for each day of those months, asked Chicago police officers who were "on the streets" at the time if they remembered anything unusual, and compared the extreme months to all other months on a number of variables. We found that the homicides committed in the two extreme months were typical of homicides committed in all other months. There were just more of them. We decided, therefore, to keep the extreme months in the analysis. We could not justify eliminating them or weighing them less.

Long and Short Series

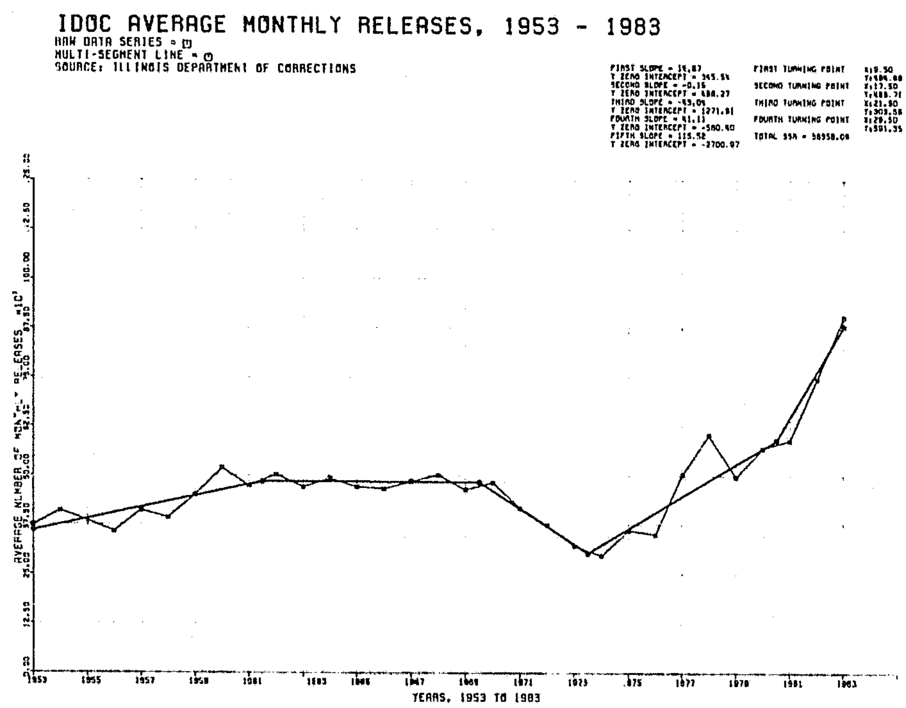
Long Series

Although the collection of many criminal justice data series is recent, compared to the collection of many economic data series, a few crime series are very long. In such a series, a simple description can be difficult to achieve. A 50-year-long monthly series, for example, may have 50 alternative pattern descriptions, too many for quick comprehension.

One way to deal with the inherent complexity of a long series is to split the series into parts, and analyze each part separately. However, this solution creates additional problems. What objective criteria do we use to divide the series into parts? Having divided the series, and described each part separately, how can we connect the descriptions into a unified whole?

Figure 18

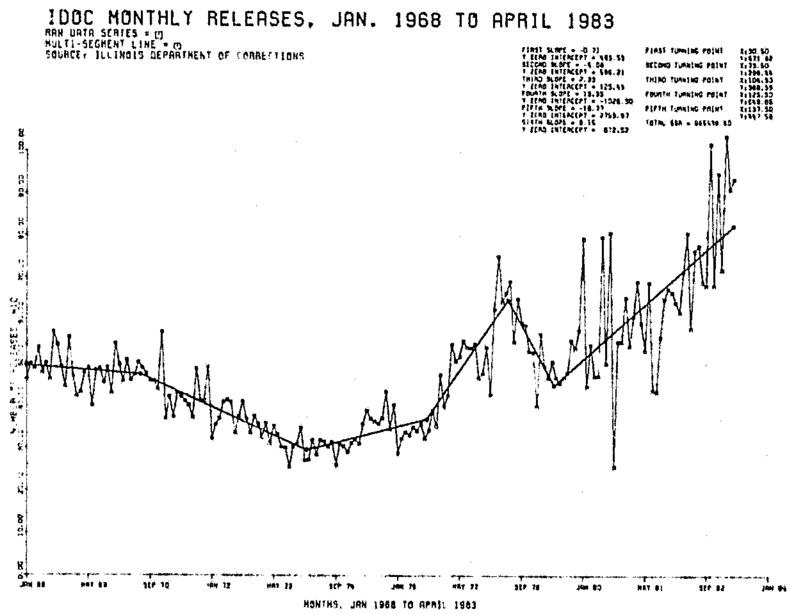
Yearly Patterns in a Long Series



Another solution, in the case of a monthly series, is to begin the analysis with a description of yearly aggregate data. For example, the description of the pattern of the number of prison releases per month in Illinois, averaged yearly (figure 18), is relatively stable from 1953 to 1968, but declines from 1968 to 1974 and increases from 1974 to 1983. This yearly pattern description may provide all the information a particular situation requires. If more detail is necessary, the yearly pattern can be used as one criterion for dividing the series into parts. Since the period after 1968 shows the greatest amount of fluctuation, and since the more recent period was more interesting to IDOC policy makers, we divided the series in 1968 for monthly analysis (figure 19). The pattern description of the monthly series is similar to the pattern description of the yearly series for the 1968 to 1983 period, but the monthly description contains more detail.

Figure 19

Patterns in a Monthly Series



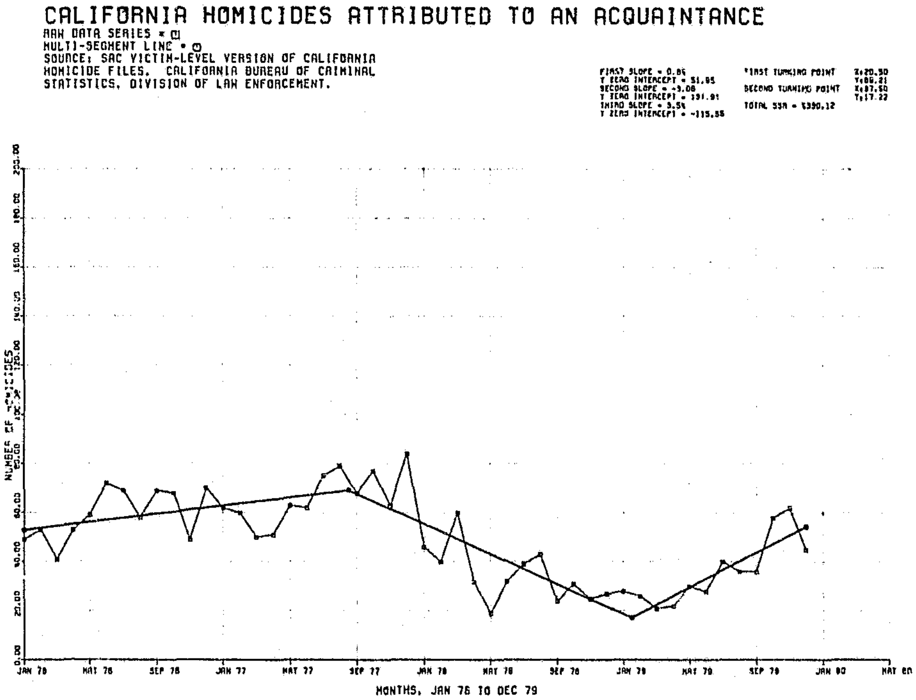
Short Series

The fewer the number of observations in a series, the fewer the possible number of segments in a pattern description, given a minimum line segment length. Some series are so short that very few alternative pattern descriptions, perhaps only a straight line and a two segment line, are possible.

For example, the number of homicides between acquaintances in California per month from 1976 to 1979 (figure 20) contains only 48 observations. If we limit the length of any line segment to a minimum of 12 months, we are unlikely to find as many as four segments.

Figure 20

A Short Series



Presence of a Possible Discontinuity

The reader of a time series graph generally assumes that the graph shows the number of occurrences of some phenomenon over time. The definition of the phenomenon remains the same; only the number of occurrences changes. The very form of a line graph, with connected dots marching steadily across the page, implies that this is true. However, it is possible that it is not true. Occasionally, the definition of the phenomenon changes in the middle of the series. When such a change in definition occurs, you do not really have one time series; you have two: one before the change and one after. Because a segmented line implies continuity of definition, it is not the appropriate pattern description for such a discontinuous series.[35]

How can you tell if a series contains a discontinuous, instantaneous change in its definition, or whether it is really a continuous series containing a short but rapidly changing segment? There are explanatory time series analysis methods that

---

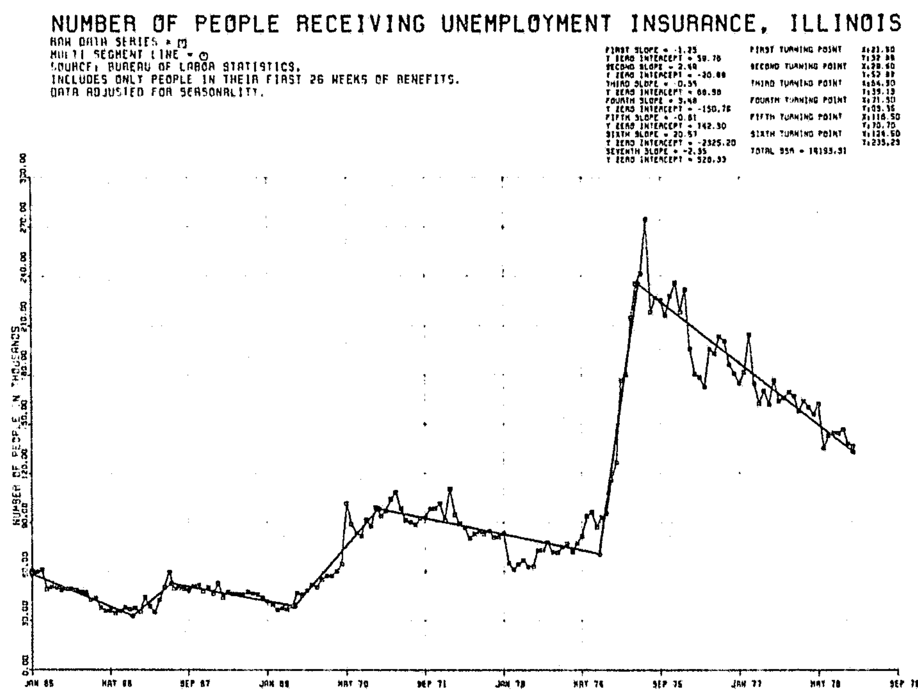
[35] A piecewise regression line is an appropriate pattern description for a discontinuous series. See note 3, page 6 above.

will test the hypothesis that some intervention, such as a legal change, caused a change in the pattern of the series.[36] However, at the initial, descriptive stage of analysis, intervention time series analysis methods are not appropriate. It is not an interrupted time series experiment, in which you have hypothesized that an intervention occurred at a certain time. Instead, it is an empirical description of an apparent discontinuity, which may indicate a change in definition. The proper course of action at this stage is to investigate the source of the data in order to determine whether or not there was a change in definition.

For example, we initially thought that the unemployment insurance benefit series (figure 21) was discontinuous. The number of people receiving unemployment insurance benefits in Illinois tripled over a period of nine months, from about 90,000 in October 1974, to more than 270,000 in June 1975. To determine whether or not the jump in the series was due to a change in the

Figure 21

An Apparent Discontinuity



---

[36]For an overview, see Glass, Willson and Gottman (1975).

way in which unemployment was defined or measured, we compared a number of related data sets, read data manuals, and talked to national and local data collectors. After a lengthy investigation, we concluded that the data were really continuous.

If a series appears to have an abrupt increase or decrease, first determine whether this apparent discontinuity indicates a change in definition. The only way to do this is to investigate the source of the data, which includes talking to the people who collect and maintain the data. If you decide that the data are defined the same way throughout, a segmented line (linear spline) is the appropriate pattern description. Remember, however, that because of the sharp increase or decrease, at least one of the line segments will be very short.

BIBLIOGRAPHY

BIBLIOGRAPHY

Alt, Frank B., Stuart Jay 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. Statistical Inference in Dynamic Economic Models. Tjalling C. Koopmans (ed.) New York: John Wiley & Sons.

Ansley, Craig F. and Wecker, William E.
1981 Extentions and examples of the signal extraction approach to regression. Mimeo, Graduate School of Business, University of Chicago.

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

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. Statistical Analysis Center, Illinois Criminal Justice Information Authority.
1983 How to Handle Seasonality, Introduction to the Detection and Analysis of Seasonal Fluctuation in Criminal Justice Time Series. Statistical Analysis Center, Illinois Criminal Justice Information Authority.

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 Criminal Justice Information Authority.

Block, Carolyn Rebecca, Louise S. Miller, Richard L. Block, and Douglas Hudson
1981 Explaining Patterns of Change Over Time in Chicago Homicides with a Gun. Statistical Analysis Center, Illinois Criminal Justice Information Authority.

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

Brown, R.P., J. Durbin and J.M. Evans
1975 Techniques for testing the constancy of regression relationships over time. Journal of the Royal Statistical Society 37:149-192, including discussion.

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.

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.

Cattell, R.B.
1966 The scree test for the number of factors. Multivariate Behavioral Research 1:245-276.

Chernoff, H.
1971 The use of faces to represent points in n-dimensional space graphically. Technical Report No. 71, Stanford University Department of Statistics.
1973 The use of faces to represent points in k-dimensional space graphically. Journal of the American Statistical Association 68:361-368.

Chow, Gregory C.
1960 Tests of equality between sets of coefficients in two linear regressions. Econometrica 28(July):591-605.

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 (July.)

Cowden, Dudley J.
1963 The perils of polynomials. Management Science 9(4,July): 542-550.

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

Daniel, Cuthbert and Fred S. Wood
1980 Fitting Equations to Data: Computer Analysis of Multifactor Data. Second edition. New York: John Wiley and Sons. Contains a number of examples of the use of Mallows's C_p plot to choose among alternative models.

Davis, James A.
1971 Elementary Survey Analysis. Englewood Cliffs, N.J.:Prentice Hall, Inc. See pp. 99-103.

Deutsch, Stuart Jay
1978 Stochastic models of crime rates. International Journal of Comparative and Applied Criminal Justice 2(2):127-151.
Builds stochastic models for seven Index crimes in each of ten U.S. cities.
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): 543-568.
See Hay and McCleary (1979) for a criticism and Deutsch (1979) for the rejoinder.

Deutsch, Stuart Jay and Lu Ann Sims
1978 An identification algorithm for dynamic intervention modeling with application to gun control. Unpublished report, Series No. J-78-29, Georgia Institute of Technology, Atlanta 30332.

Ehrenberg, A.S.C.
1978 Data Reduction: Analysing and Interpreting Statistical Data. Revised reprint. New York:John Wiley and Sons, Ltd.
1981 What we can and can't get from graphs, and why. Paper presented at the American Statistical Association meetings, Detroit.
1982 How good is best? The Journal of the Royal Statistical Society Series A (General) 145 (3):364-366.

Emerson, John D. and David C. Hoaglin
1982 Resistant lines for y versus x. Pp. 129-165 in Hoaglin, Mosteller and Tukey (eds.) Understanding Robust and Exploratory Data Analysis. New York: John Wiley & Sons.

Ertel, James E. and Edward B. Fowlkes
1975 Some Algorithms for Linear Spline and Piecewise Linear Regression. Bell Laboratory Memorandum.
Contains more detail than the article below.
1976 Some algorithms for linear spline and piecewise multiple linear regression. Journal of the American Statistical Association 71(355):640-648.

Feder, Paul I.
1975a On asymptotic distribution theory in segmented regression problems--identified case. The Annals of Statistics 3(1): 49-83.
1975b The log likelihood ratio in segmented regression. The Annals of Statistics 3(1):84-97.

Felson, Marcus
1981 Review of "Forecasting Crime Data: An Econometric Analysis" by James Alan Fox. Contemporary Sociology 10(2):230-231.

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.)

Fienberg, Stephen E.
1979 Graphical methods in statistics. The American Statistician 33(4,November):165-178.

Flury, Bernhard, and Hans Riedwyl
1981 Graphical representation of multivariate data by means of asymmetrical faces. Journal of the American Statistical Association 76(376):757-765.

Fox, James Alan
1978 Forecasting Crime Data. Lexington, Mass.:Lexington Books. See Felson (1981).

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.
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,) these are the classic time series intervention literature. Also see Tryon (1982) for a nonparametric method.

Gorman, J.W. and R.J. Toman
1966 Selection of variables for fitting equations to data. Technometrics 8(1,February):27-52.

Granger, C.W.J. and Paul Newbold
1977 Forecasting Economic Time Series. New York: Academic Press.

Harman, Harry H.
1976 Modern Factor Analysis. Third edition. Chicago: The University of Chicago Press.

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.

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):277-314.
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.

Hinkley, D. V.
1969 Inference about the intersection in two-phase regression. *Biometrika* 56(3):495-504.
Suggests a slight improvement in Hudson's (1966) algorithm, using the difference between the SSR's for a straight line and a two segment line, instead of the two segment SSR, in the iterative search for the maximum likelihood join point.

Hudson, Derek J.
1966 Fitting segmented curves whose join points have to be estimated. *American Statistical Association Journal* 61:1097-1129.
See Hinkley (1969).

Hyman, Herbert H.
1955 *Survey Design and Analysis*. New York: The Free Press.
See Chapters vi-vii.

Kendall, Sir Maurice
1976 *Time-Series*. Second edition. New York: Hafner Press.

Kendall, M.G. and A. Stuart
1966 *The Advanced Theory of Statistics*. Vol. 3. New York: Hafner Publishing Co., Inc.
Chapter 46 outlines trend and seasonality.

Kendall, Patricia L. and Paul F. Lazarsfeld
1950 Problems of survey analysis. Pp. 133-196 in Robert K. Merton and Paul F. Lazarsfeld (eds.) *Continuities in Social Research*. Glencoe, Illinois: The Free Press.
1955 Interpretation of statistical relations as a research operation. Pp. 115-125 in Paul F. Lazarsfeld and Morris Rosenberg (eds.), *The Language of Social Research*. New York: The Free Press.

Kennard, Robert W.
1971 A note on the C_p statistic. *Technometrics* 13(4,November): 899-900.

Kleiner, B. and J.A. Hartigan
1981 Representing points in many dimensions by trees and castles *Journal of the American Statistical Association* 76:260-276.

Land, Kenneth C.
1980 Modeling macro social change. Ch.8, pp. 219-278 in Sociological Methodology, 1980.

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

Leinhardt, Samuel and Stanley S. Wasserman
1978 Exploratory Data Analysis: An Introduction to Selected Methods. Sociological Methodology.

Macaulay, Frederick R.
1931 The Smoothing of Time Series. New York: National Bureau of Economic Research.

McCain, Leslie J. and Richard McCleary
1979 The statistical analysis of the simple interrupted time-series quasi-experiment. Pp. 233-293 in Quasi-experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally.

McCleary, Richard and Richard A. Hay, Jr.
1980 Applied Time Series Analysis for the Social Sciences. Beverly Hills:Sage Publications.

McGee, Victor E. and Willard T. Carleton
1970 Piecewise regression. Journal of the American Statistical Association 65(Sept.):1109-1124.

Mallows, Colin L.
1964 Choosing variables in a linear regression: A graphical aid. Presented at the Central Regional Meeting of the Institute of Mathematical Statistics, Manhattan, Kansas.
1973 Some comments on C_p. Technometrics 15(4,November):661-675. Also see Gorman and Toman (1966); Kennard (1971); Daniel and Wood (1980).
1980 Some theory of nonlinear smoothers. Annals of Statistics 8(4):695-715.

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

Marshall, Clifford W.
1977a Application of Time Series Methodology to Crime Analysis. The Polytechnic Institute.
1977b The State Space Forecasting Technique Applied to Reported Crime Data. Supplement to 1977a, above.

Miller, Louise S.
1983 Forty Years of Illinois Prison Population. Statistical Analysis Center, Illinois Criminal Justice Information Authority.

Nelson, Charles R.
1973 Applied Time Series Analysis. San Francisco: Holden-Day.

Nerlove, Marc, David M. Grether and Jose' L. Carvalho
1979 Analysis of Economic Time Series: A Synthesis. New York: Academic Press.

Nie, Norman H.
1975 SPSS: Statistical Package for the Social Sciences. New York: McGraw-Hill Book Co.

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

Pierce, David A.
1980 A survey of recent developments in seasonal adjustment. The American Statistician 34(August,3):125-134.

Pierce, Glenn L. and William J. Bowers
1979 The impact of the Bartley-Fox gun law on crime in Massachusetts. Boston: Center for Applied Social Research, Northeastern University, Boston, Massachusetts 02115.

Poirier, Dale J.
1976 The Econometrics of Structural Change, with Special Emphasis on Spline Functions. Amsterdam:North Holland Publishing Co.
This is an excellent introduction to the subject of spline functions.

Selvin, Hanan C.
1958 Durkheim's Suicide and problems of empirical research. American Journal of Sociology 63:607-613.
Reprinted in R.A. Nisbet (ed.), Emile Durkheim (Englewood Cliffs, N.J.: Prentice-Hall, 1965), and Lazarsfeld, et al. (eds.) Continuities in Social Research. New York: The Free Press, 1972).

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.
1982a 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.

1982b The multivariate case of the Shine-Bower single-subject ANOVA. Educational and Psychological Measurement 42:453-457.

1982c Multi-subject behavior functions logically based on Shine's actualized and pure single-subject behavior functions. Educational and Psychological Measurement 42:783-788.

Shiskin, Julius, Allan H. Young, and John C. Musgrave

1967 The X11 Variant of the Census Method II Seasonal Adjustment Program. U.S. Department of Commerce, Bureau of the Census.

Stone, Charles J.

1977 Consistent nonparametric regression. The Annals of Statistics 5(4):595-645 (with commentary).

Tintner, Gerhard, P. Whittle, Herman Wold and Julius Shiskin

1968 Time series. Pp. 47-88, International Encyclopedia of the Social Sciences. Vol. 16. David L. Sills (ed.)

Tishler, Asher and Israel Zang

1981 A new maximum likelihood algorithm for piecewise regression Journal of the American Statistical Association 76(376, December):980-987.

Tryon, Warren W.

1982 A simplified time-series analysis for evaluating treatment interventions. Journal of Applied Behavior Analysis 15 (3,Fall):423-429.

Tukey, John W.

1962 The future of data analysis. Annals of Mathematical Statistics 33:1-67.

1977 Exploratory Data Analysis. Reading, Mass.: Addison-Wesley Also see Emerson and Hoaglin (1982), Velleman and Hoaglin (1981).

Velleman, Paul F. and David C. Hoaglin

1981 Applications, Basics, and Computing of Exploratory Data Analysis. Boston:Duxbury Press.
A very basic introduction to Exploratory Data Analysis.

Vigderhous, Gideon

1978 Forecasting sociological phenomena: Application of Box-Jenkins methodology to suicide rates. Pp. 20-51 in Sociological Methodology 1978.

Wahba, Grace

1977 Discussion of Stone (1977). The Annals of Statistics 5(4): 637-640.

1978 Improper priors, spline smoothing and the problem of guarding against model errors in regression. Journal of the Royal Statistical Society, Series B 40(3):364-372.

Wecker, William E. and Craig F. Ansley
1983 The signal extraction approach to nonlinear regression and spline smoothing. Journal of the American Statistical Association 78(381,March):81-89.

Wegman, Edward J. and Ian W. Wright
1983 Splines in Statistics. Journal of the American Statistical Association 78(382,June):351-366.

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.