

1-1-1984

Forecasting interest rates using pattern recognition techniques

John S. Pearson
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Let us know how access to this document benefits you.

Recommended Citation

Pearson, John S., "Forecasting interest rates using pattern recognition techniques" (1984). *Dissertations and Theses*. Paper 149.

<https://doi.org/10.15760/etd.149>

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

FORECASTING INTEREST RATES
USING PATTERN RECOGNITION TECHNIQUES

by
JOHN S. PEARSON, JR.

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY
in
SYSTEMS SCIENCE

Portland State University

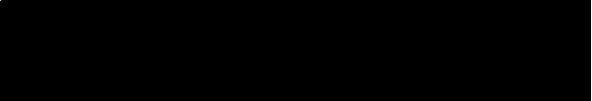
© 1984 JOHN S. PEARSON, JR.

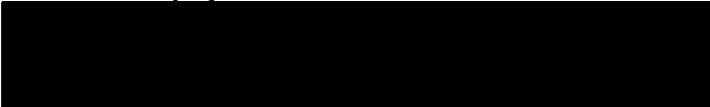
TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the dissertation
of John S. Pearson, Jr. presented July 2, 1984.

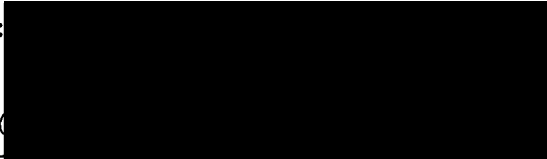

Harold A. Linstone, Chairman


Giles H. Burgess


Christopher James


Thomas H. Tuchscherer

APPROVED:


Jim F. Heath, Associate Vice President for Academic Affairs

AN ABSTRACT OF THE DISSERTATION OF John S. Pearson, Jr. for
the Doctor of Philosophy in Systems Science presented
July 2, 1984.

Title: Forecasting Interest Rates Using Pattern
Recognition Techniques

APPROVED BY THE MEMBERS OF THE DISSERTATION COMMITTEE:

[REDACTED]

Harold A. Linstone, Chairman

[REDACTED]

Giles H. Burgess

[REDACTED]

Christopher James

[REDACTED]

Thomas H. Tuchscherer

Much depends on the future course of interest rates.
The decisions of families to make major purchases, the
willingness of businesses to expand and invest, the rise
and fall of the economy and stock market, the ability of
lesser developed countries to repay their debts, the tenure

of presidents and prime ministers--all of these may turn on whether interest rates increase or decrease in the months ahead.

Several forecasting models developed in the dissertation permit the direction of change of interest rates on long-term U.S. government bonds to be forecast correctly about 60% of the time. When the different models are combined, effectiveness is increased; and when the forecasts are dollar-weighted, performance rates in excess of 70% are possible.

The dissertation reports on one of the first applications of powerful techniques recently developed in cybernetics, engineering, and artificial intelligence to forecasting the direction of change in interest rates. Two forecasting algorithms, called linear decision functions or linear classifiers, are derived using the principles of pattern recognition. Because they are recursively updated, both algorithms operate dynamically and adapt their performance to changes in the economic environment. One classifier, a modification of the widely used least-mean-squared-error algorithm, permits monthly revision of the forecasting model and allows the larger historical movements in interest rates to have greater weight in future decisions. The second algorithm permits refinement of the parameter estimates generated by the first.

These formal, mathematical classifiers use financial

variables suggested by economic theory--leading indicators of inflation and investment activity. The raw economic time series are computer-enhanced with a detrending algorithm to reduce "noise" and extract more information for input into the forecasting models. The direction of change of interest rates on twenty-year, constant-maturity U.S. government bonds is forecast over a one-month time horizon throughout the period 1969-82. All forecasts are unconditional and ex-ante, i.e., all forecasts refer to periods beyond the time used in estimating model parameters and assume no knowledge of future economic events. The results are evaluated in comparison with a Bayesian forecasting model and a 10,000-event Monte Carlo simulation of a random decision rule. The forecasting ability of the algorithms is statistically significant at the 99% level of confidence.

These results have both theoretical and practical importance. They are of practical interest to investment managers who would improve the total return on assets under their control, to speculators who would increase profits and reduce risk, to treasurers of large corporations who would lower the cost of borrowed funds, and to policymakers who would reduce uncertainty. The results are of methodological interest to forecasters who would be satisfied to predict the direction of change in a variable if they were not yet able to make reliable point estimates. Finally, there is material of interest to those who are uneasy about

the economic edifice being constructed on the theory of rational expectations. In particular, the so-called efficient market theory is tested and found wanting.

ACKNOWLEDGMENTS

I appreciate the guidance and support of my dissertation committee: Harold Linstone (chairman), Giles Burgess, and Thomas Tuchscherer, all from Portland State University, and Christopher James of the University of Oregon.

Two others, not members of the committee, George Lendaris (Portland State University) and James Comas (University of Southern California) read an early draft of the work and offered many valuable suggestions.

Several people assisted me in obtaining data and other information: Charlotte Boschan (Economics Department, Citibank), Russell Fogler (University of Florida), P. J. Harrison (University of Warwick), Spyros Makridakis (INSEAD), Audrey Powers (Research Library, Federal Reserve Bank of San Francisco), James Scherer (U.S. Treasury Department), Dorothy Swink (Research and Statistics Division of the Federal Reserve System) and Herbert Taylor (Federal Reserve Bank of Philadelphia).

The Commodity Research Bureau, Inc. and the F.W. Dodge Division, McGraw-Hill Information Systems Company granted permission to use their copyrighted data series.

The staffs of the Portland State University library and the computer center were always helpful.

Printing and plotting equipment for the physical production of the work were provided by Peggy Pearson at the Portland office of Milliman and Robertson, Inc., Consulting Actuaries. Robert Pryor of that office helped resolve equipment interfacing difficulties.

Administrative burdens were made much lighter because Sally Lopez (Systems Science Program, Portland State University) carried most of the load.

Thank you all.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGMENTS	iii
LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER	
I INTRODUCTION	1
II PERSPECTIVES ON INTEREST RATE FORECASTING	8
Multiple Perspectives	8
The Efficient Market Theory	13
What Is An Efficient Market	
The Importance of the Efficient	
Market Theory	
Weaknesses of the Efficient Market	
Theory	
Empirical Tests of The Efficient	
Market Theory	
The Difficulty of Forecasting	
Self-Altering Forecasts	
Survival of the Fittest Forecasters	
Conclusion	
Other Market Perspectives	35
The Fundamentalists	
The Chartists and Cyclists	
The Insiders	
The Traders	
Forecasting Methods	38
Judgmental Methods	
Econometric Models	
Economic Indicator Analysis	
Time Series Analysis	
Market Analysis Methods	

Pattern Recognition

III	PRINCIPLES OF PATTERN RECOGNITION	56
	Pattern and Features	56
	Classifiers	59
	The Perceptron	67
	The Minimum Squared Error Classifier . . .	75
	Weighted Linear Classifiers	80
	Self-Adaptation of the Classifiers . . .	86
	Summary	94
IV	INTEREST RATE FORECASTING	95
	Preliminary Matters	95
	What Is a Forecast	
	What Rate to Forecast	
	The Objective Function	104
	Evaluation Standards	106
	A Random Rule	
	A Bayesian Rule	
	Data Sources	122
	Dynamic Detrending	126
	Economic Theory	133
	Forecasting Models	136
	The Investment Model	
	The Price-Investment Model	
	The Labor Model	
	An Atheoretical Model	
	Evaluating the Forecasting Models . . .	159
	The L-Method Test	
	Analysis and Discussion	170
	Combining Forecasts	175

V	SUMMARY, CONCLUSIONS, IMPLICATIONS AND DIRECTIONS FOR SUBSEQUENT RESEARCH	196
	Summary	196
	Conclusions	198
	Implications	200
	Directions for Subsequent Research . .	203
	REFERENCES	206
	APPENDIX	
A	SUMMARY OF NOTATION	223
B	TECHNICAL NOTES	226
C	USE OF THE HYPERBOLIC TANGENT ALGORITHM .	241
D	USING THE MODELS TO HEDGE A FIXED-INCOME PORTFOLIO	243
E	DATA	248

LIST OF TABLES

TABLE		PAGE
I	Frequency Data from a 10,000-Trial Monte Carlo Simulation: Net Profit	118
II	Frequency Data from a 10,000-Trial Monte Carlo Simulation: Profit/Loss Ratio	118
III	Frequency Data from a 10,000-Trial Monte Carlo Simulation: Number of Hits	119
IV	Frequency Data from a 10,000-Trial Monte Carlo Simulation: Rate of Return	119
V	Frequency Data from a 10,000-Trial Monte Carlo Simulation: Worst Computed Loss	120
VI	Rank Data from a 10,000-Trial Monte Carlo Simulation	120
VII	Month by Month Profit or Loss: Bayesian Rule	125
VIII	Month by Month Profit or Loss: Investment Model	141
IX	Month by Month Profit or Loss: Price-Investment Model	147
X	Month by Month Profit or Loss: Labor Model	153

TABLE		PAGE
XI	Month by Month Profit or Loss: Atheoretical Model	158
XII	Hypothesis Test Statistics	166
XIII	Comparison of Model Performance with the L-Method Test	169
XIV	Summary Statistics for Moving-Sum Models .	178
XV	Summary Statistics for Majority-Rule Models	179
XVI	Summary Statistics for Hedge-Rule Models .	179
XVII	Summary Statistics for Veto-Rule Models .	181
XVIII	Series as Reported by Source	249
XIX	Series as Detrended	255
XX	Coefficients for the Forecasting Models .	261

LIST OF FIGURES

FIGURE	PAGE
1. Linearly Separable Features in a Two-Dimensional Feature Space.	61
2. Linearly Unseparable Features in a Two-Dimensional Feature Space.	64
3. The Pattern Classification Process	66
4. Evolution of the Perceptron Partition Boundary in a Two-Dimensional Feature Space	70
5. Partition Boundary for the Minimum Squared Error Algorithm in a Two-Dimensional Feature Space.	79
6. Hyperbolic Tangent Approximation to the Sign Function	83
7. Partition Boundary for the Minimum Squared Error Algorithm -- First Three Points .	90
8. Partition Boundary for the Minimum Squared Error Algorithm -- First Four Points . .	91
9. Partition Boundary for the Minimum Squared Error Algorithm -- All Five Points . . .	92
10. Interest Rate on 20-Year Constant-Maturity U.S. Government Bonds	99

FIGURE	PAGE
11. The Yield Curve	101
12. Relative Frequency and Cumulative Relative Frequency for the Net Profit in the Monte Carlo Simulation	113
13. Relative Frequency and Cumulative Relative Frequency for the Profit/Loss Ratio in the Monte Carlo Simulation	114
14. Relative Frequency and Cumulative Relative Frequency for the Number of Hits in the Monte Carlo Simulation	115
15. Relative Frequency and Cumulative Relative Frequency for the Rate of Return in the Monte Carlo Simulation	116
16. Relative Frequency and Cumulative Relative Frequency for the Worst Computed Loss in the Monte Carlo Simulation	117
17. Cumulative Net Profit--Bayesian Model . . .	123
18. Profit/Loss Ratio--Bayesian Model	124
19. Ordinary Least Squares Regression	128
20. Axis Translation for Regression in Deviations-from-Means Form	129
21. Relative Weights Assigned to New Points in Weighted and Unweighted Regressions . .	132

FIGURE

PAGE

22.	Construction Contracts for Commercial and Industrial Bulidings--Reported and Detrended Series	138
23.	Cumulative Net Profit--Investment Model . .	139
24.	Profit/Loss Ratio--Investment Model	140
25.	Contract and Orders for Plant and Equip- ment--Reported and Detrended Series . .	143
26.	Commodity Research Bureau Index of Spot Prices on 22 Commodities--Reported and Detrended Series	144
27.	Cumulative Net Profit--Price-Investment Model	145
28.	Profit/Loss Ratio--Price-Investment Model .	146
29.	Civilian Employment Ratio--Reported and Detrended Series	150
30.	Cumulative Net Profit--Labor Model	151
31.	Profit/Loss Ratio--Labor Model	152
32.	Ratio of Coincident to Lagging Indicators-- Reported and Detrended Series	155
33.	Cumulative Net Profit--Atheoretical Model .	156
34.	Profit/Loss Ratio--Atheoretical Model . . .	157

FIGURE

PAGE

35.	Relative Frequency and Cumulative Relative Frequency for the Net Profit in the Monte Carlo Simulation, Comparing the Models	160
36.	Relative Frequency and Cumulative Relative Frequency for the Profit/Loss Ratio in the Monte Carlo Simulation, Comparing the Models	161
37.	Relative Frequency and Cumulative Relative Frequency for the Number of Hits in the Monte Carlo Simulation, Comparing the Models	162
38.	Relative Frequency and Cumulative Relative Frequency for the Rate of Return in the Monte Carlo Simulation, Comparing the Models	163
39.	Relative Frequency and Cumulative Relative Frequency for the Worst Computed Loss in the Monte Carlo Simulation, Comparing the Models	164
40.	Cumulative Net Profit--Investment Model, Labor Model, and Atheoretical Model Combined According to the Majority Rule	182

FIGURE

PAGE

41.	Profit/Loss Ratio--Investment Model, Labor Model and Atheoretical Model Combined According to the Majority Rule	183
42.	Cumulative Net Profit--Investment Model and Labor Model Combined According to the Hedge Rule	184
43.	Profit/Loss Ratio--Investment Model and Labor Model Combined According to the Hedge Rule	185
44.	Cumulative Net Profit--Price-Investment Model and Labor Model Combined According to the Hedge Rule	186
45.	Profit/Loss Ratio--Price-Investment Model and Labor Model Combined According to the Hedge Rule	187
46.	Cumulative Net Profit--Labor Model and Atheoretical Model Combined According to the Hedge Rule	188
47.	Profit/Loss Ratio--Labor Model and Atheoretical Model Combined According to the Hedge Rule	189
48.	Cumulative Net Profit--Investment Model, Price-Investment Model and Labor Model Combined According to the Veto Rule . .	190

FIGURE

PAGE

49.	Profit/Loss Ratio--Investment Model, Price- Investment Model and Labor Model Combined According to the Veto Rule . . .	191
50.	Cumulative Net Profit--Investment Model, Price-Investment Model and Atheoretical Model Combined According to the Veto Rule	192
51.	Profit/Loss Ratio--Investment Model, Price- Investment Model and Atheoretical Model Combined According to the Veto Rule . . .	193
52.	Cumulative Net Profit--Price-Investment Model, Labor Model and Atheoretical Model Combined According to the Veto Rule	194
53.	Profit/Loss Ratio--Price-Investment Model, Labor Model and Atheoretical Combined According to the Veto Rule	195
54.	An Ordinary Least Squares Regression Program Written in the Notation of APL	234
55.	Cumulative Pay-out Under Three Fixed-Income Portfolio Investment Strategies	246

CHAPTER I

INTRODUCTION

It is the most crucial question facing the world economy. The answer will determine whether millions of Americans will be able to afford a new home or a new car, whether U.S. companies will undertake new investment to propel the recovery on to a sustainable growth path, whether the stock market will continue to rise, and whether massive defaults by heavily indebted developing countries will cause the international financial system to collapse. Yet after the wild gyrations of the past year, businessmen, investors, and economists appear more confused than ever over this question: Where will interest rates go?

With that paragraph Business Week (1983) began a cover-featured story on interest rates. The objective of the research reported here is to find a method for answering the "most crucial question"--to find a way to forecast the direction of change of interest rates.

The results should interest several groups: financial decision makers in business and industry who borrow or lend large sums of money in the nation's capital markets; speculators who seek to profit by correctly predicting swings in interest rates; government officials who set economic policy; economic theorists who are concerned about questions of market efficiency; financial forecasters who are not fully satisfied with existing forecasting techniques; and those in the systems field who believe that a systems

perspective can help solve real-world problems.

First consider the financial decision makers. On September 22, 1982, The Wall Street Journal ran a front page article, "Baffled Borrowers: Corporate Treasurers Find the Rate Outlook Unusually Confusing."

A prudent finance executive whose company is borrowing \$100 million at short term can save about \$10,000 a day just by doing nothing.... Over a 10-year period, however, inaction could cost the company \$20 million or more if it misses a window and later has to sell \$100 million of bonds at a higher interest cost.

Forecasting interest rates has always been tricky because nobody knows what will happen next week, much less six months or a year hence.... Moreover, the penalty for being wrong is much higher than it used to be, finance men lament, because interest rates have become so volatile. Twenty years ago, bond rates moved only fractions of a percentage point for months at a time.... But in just 15 months, from mid-1980 to last fall, bond interest rates soared almost six percentage points. A company that missed the 1980 window and borrowed \$100 million last September could pay an extra \$60 million of interest over 10 years.

This Journal article adopts the viewpoint of the borrower, the corporation selling bonds to raise needed funds. The other side of the transaction involves the lender of funds, the bond purchaser, which may be an insurance company, pension fund, or bank trust. The investment managers of these institutions face a similar dilemma: to buy bonds when returns appear low or to hold out for possibly higher returns later. But while higher yields would increase the future income received by lenders making new investments, the higher rates have a seriously negative

impact on the value of debt securities already held in portfolios. Consider the life insurance business as an example. The bond portfolios of life insurance companies were reported to exceed \$200 billion at the end of 1982 (American Council of Life Insurance, 1983). Since bond prices and interest rates are inversely related, a 1% rise in interest rates would reduce the value of these assets by perhaps \$20 billion. Clearly, changes in interest rates have serious financial consequences for borrowers and lenders alike.

It is possible, however, for those who hold fixed-income securities to use the financial futures markets to hedge their risk of capital loss. A futures contract is an agreement to buy or sell a specified quantity of a commodity at a specified price at some future date (Horn and Farah, 1979). The commodity involved in financial futures contracts is a security. For example, futures contracts for U.S. Government bonds are traded in \$100,000 units on the Chicago Board of Trade. Futures contracts allow the risk of price fluctuation to be transferred from the holder of the commodity to someone else willing to bear that risk. Those who hold bond portfolios can sell futures contracts; if interest rates were to rise, the loss on its bond portfolio would be offset by the gain on its futures contracts.

The party to whom the risk is transferred is frequently a speculator hoping to profit by correctly

anticipating the direction that prices will move. The volume of risks so transferred by hedging and speculation is enormous: the value of financial futures traded on the Chicago futures exchanges exceeds by far the total value of stocks bought and sold on New York and American Stock Exchanges combined. Given this volume of trading, it is easy to see why participant speculators as well as commercial hedgers are so seriously concerned about the future course of interest rates.

It is not only borrowers, lenders, and speculators who follow closely the outlook for interest rates. Those who guide and influence national economic policy consider interest rates an important macroeconomic variable. Total debt outstanding in the U.S. economy is measured in trillions of dollars. The interest component of national income has risen from \$4 billion in 1952 to more than \$250 billion in 1982; in percentage terms, from less than 1.5% of national income to almost 11% (U.S. Department of Commerce, 1983).

This trend concerns thoughtful observers. Felix Rohatyn (1984), credited with engineering the fiscal rescue of New York, worries particularly about the public sector:

I think that this country by 1985 is going to look the way New York City looked in 1975, with a runaway budget and a debt service that's going to run up to, I don't know, 20% or 25% of total government spending, which you can't sustain. At some point there are going to be market problems.

The financial effects of interest rate movements are

so widespread that financial regulators, central bankers, and policymakers in general would like to be able to anticipate these changes.

Interest rate forecasting has bearing on another, more theoretical, aspect of public policy. It is a tenet of most free-market advocates that efficient markets contribute to the general welfare of society. Because the question of market efficiency is such an important one in economic theory, it has been subject to extensive empirical investigation. This study provides additional evidence in the continuing debate.

Forecasters are always looking for new and better forecasting methods. The techniques used here, while broadly applied in engineering, have not received much attention in economics. When the primary forecasting task is to gauge the direction of change of an economic variable (as is typical, for example, in the so-called turning point problem), these approaches appear to have substantial merit.

This research may be of interest to some in the systems field. As president-elect of the Society for General Systems Research, John Warfield (1981) issued a statement concerning the criteria for selecting research topics.

Most systems researchers are not studying the hard problems that are constantly in the fore in the media. [We] leave inflation to the economists, leave world stability to the political scientist, leave education to the educators, leave productivity to the business

entrepreneurs, leave organizational design to the personnel people.

... in many of these hard-problem areas, much relevant work has been done in one or more disciplines. If this work were properly integrated and augmented from a system perspective, highly valuable results might be produced.

The research undertaken here demonstrates how techniques developed in engineering can be applied to a hard problem constantly in the fore in the media. Concepts from physics, ecology, physiology, psychology, sociology, and information theory help, from time to time, to illuminate the problem.

In this chapter the author has identified the audiences to whom the work is addressed and has discussed the reasons some of these groups are motivated to forecast interest rates.

Chapter II looks at the interest rate forecasting problem from several perspectives. The major theoretical impediment to forecasting--the efficient market theory--is discussed at length. The nature of the economic environment in which forecasting takes place is considered, and different methods of forecasting prices are reviewed. The second chapter concludes with a general, nontechnical overview of how forecasts can be made using pattern recognition techniques.

Chapter III starts from basic principles and develops the subject of pattern recognition just far enough to

derive the algorithms needed to forecast the direction of change in interest rates.

The substantive results are in Chapter IV. First, it provides unambiguous notions of what a forecast is and which interest rate variable is to be forecast. It then proposes a reasonable forecasting objective and suggests standards by which forecasting success can be judged. A brief discussion of the economic theory underlying the choice of variables to be used for forecasting is followed by the development of a method of detrending those variables so as to better capture information useful in forecasting interest rates. Four specific models for making forecasts are then presented and their performance evaluated in comparison with a random walk simulation and a Bayesian alternative. The four forecasts are then combined in several different ways to yield composite forecasts that are better than their components.

The final chapter summarizes the work. It gathers in one place the major conclusions and develops their implications for the audiences identified here. Since the sections of Chapter V which deal with conclusions and implications are brief and phrased nontechnically, some readers may want to read those sections first before starting Chapter II.

CHAPTER II

PERSPECTIVES ON INTEREST RATE FORECASTING

"1984 interest rates? Ask Ouija board."

This Oregonian headline, which ran in late 1983, exhibits its writer's skepticism about interest rate forecasting: Can it be done at all? If so, what methods work best? The present chapter considers these questions. The author concludes that interest rates can be forecast--though not perfectly and only with difficulty--and he offers pattern recognition as the preferred method.

MULTIPLE PERSPECTIVES

Many phenomena are usefully considered from different, sometimes conflicting, points of view. This is as true in physical science as in social science and everyday life. In physics, for example, the propagation of light is best explained by the electromagnetic wave theory, while the interaction of light with matter is better understood using the corpuscular theory (Ridley, 1976). Physicists no longer ask which theory is true, they only ask which is more useful in a given context.

James Keys (pseudonym of G. Spencer-Brown, 1972)

makes the point this way.

You may look at the world any way you please, through any window you choose.... How the world appears, what you see and what you miss ... depends on which window you are using.

Windows or viewpoints can go by different names. One set of epistemological models has been catalogued under the term inquiring systems. C. West Churchman (1971) labeled his models with the names of the philosophers who espoused them. (Of course, the descriptions of these models, and others that follow, are caricatures. No man's ideas can be summarized in two sentences. But it is the caricatures that makes them memorable and useful.)

Locke, the empiricist, finds truth in data, observation and experience. It is from "hard facts" and consensual agreement about them that reality is known.

Leibniz, the rationalist, finds truth in the analytic, deductive and theoretical. What is needed to understand reality is a grand world model from which all truth can be derived.

Kant finds truth in the synthesis of data and theory. Data and model are complementary and inseparable; the theoretical and empirical must be meshed. Objectives and alternatives are important to Kant.

Hegel, the dialectician, finds truth in the conflict of structured debate. Conflicting views

are synthesized into a broader, more encompassing world view.

Singer, the pragmatist, finds truth in a holistic view which utilizes all of the above approaches. He also asks that both subjective and objective factors be considered and that important values give rise to injunctions and commands.

Another set of windows has been developed by Linstone (1981, 1984). Called multiple perspectives, these have proven especially valuable in the study of socio-technological systems, particularly technology assessment and policy decision making.

The Technical Perspective is closely associated with and encompasses most of the paradigms of science and technology. It emphasizes classical rationality, seeks to optimize product or performance, and looks for cause-effect relationships.

The Organizational Perspective is the viewpoint of a group within the social infrastructure. Challenge-response may replace cause-effect; problem-avoidance may replace problem-solution; "satisficing" takes over from optimizing; and loyalty may be more important than rationality.

The Personal Perspective is the viewpoint of the individual. It is the outlook of the person with his particular knowledge and beliefs, values and biases, fears and ambitions.

A third set of windows has been provided by Martino (1972) in the context of technology forecasting, but its usefulness is certainly not limited to that domain. His sectors or dimensions include the following:

- Technological
- Economic
- Managerial
- Political
- Social
- Cultural
- Intellectual
- Religious/Ethical
- Ecological

Like Linstone's perspectives, Martino's dimensions can be considered as ways of viewing a problem as well as indicating where to look for aspects of it, both things you look through and things you look at.

Other broadly useful sets of perspectives have been given by Allison (1971), Steinbruner (1974), Lilly (1975) and Andersen (1977). A set of perspectives particularly suited to dealing with the topic of interest rate forecasting and making sense of securities markets is that developed by Smith (1981). Smith's perspectives are examined later in this chapter.

When one begins to use multiple perspectives,

several difficulties become apparent. First, as Keys (1972) notes, "The number of different windows is endless." This problem is overcome by using only a few perspectives to spotlight a given problem and resolutely excluding all others. Failure to do so may postpone action through endless analysis. When choosing the perspectives, one should be careful to pick those which fit oneself as well as those which fit the problem.

Not everyone is psychologically able to handle the cognitive dissonance that multiple, conflicting perspectives can create. In extreme situations, some people go crazy (Bateson, 1972). Lilly (1967) has proposed a multi-valued logic that may be of some help here. Rather than the usual dichotomy of truth and falsehood, Lilly proposes a logic with four values -- true, false, as-if-true, and as-if-false. This is really not so radical as it seems: children and mathematicians, for example, use as-if-true frequently.

Another "problem" is that once one admits of multiple perspectives, the question of values cannot be ignored. Perspectives are not scientifically verifiable statements; instead, science (considered as a method of knowing rather than as a body of knowledge) is a perspective. Perspectives are justified by their utility, by how well they work, and by the value system of the user. The organizational and personal perspectives, for example,

require that some organization's or some person's preferences be considered. Moreover, those with opposing views must be given broader range to criticize. Since the views from different windows are not generally isomorphic, those using different perspectives can become involved in endless, sometimes fruitless, debate.

At the highest level, this study has been guided by a pragmatic, Singerian inquiring system. In execution (Chapters III and IV) it falls well within the range of main-stream science, employing primarily the technical perspective of Linstone and looking mainly to the economic sector of Martino. From time to time, it has been enlightened ("cross-cued") by observations taken from other perspectives.

The next section examines the most widely held academic theory on the question of interest rate forecasting. In the view of many economists, the efficient market theory is true (or as-if-true). If one wishes to forecast interest rates, it is far more useful to consider the theory as false (or as-if-false).

THE EFFICIENT MARKET THEORY

The efficient market theory (EMT) is extremely important to would-be interest rate forecasters since, if true, it effectively puts their goal beyond reach. The theory implies that meaningful trading profits cannot be

made by forecasting interest rates using publicly available information.

What Is An Efficient Market?

In academic circles an efficient market is defined as one in which there are a large number of well-informed participants actively competing to maximize profits. In this kind of market all existing information and all anticipated information is "discounted" in the prevailing market price. In an efficient market only new information or new anticipations can affect price. Since new information arrives randomly, is disseminated rapidly, and is fully and correctly reflected in the market price almost instantaneously, no one can consistently outperform the market in the sense of earning excess profits as a result of price prediction based on publicly available information. This is the efficient market theory. It does not deny the existence of causes and explanations for prices changes; it only says that these causes, and therefore the resulting price movements, cannot be known in advance on the basis of public information.

The efficient market theory comes in several varieties. In its "pure" form it is a statement about a mental construct that does not exist in the real world. Just as physicists postulate imaginary substances (such as an ideal gas to which Boyle's law applies by definition), economists postulate situations where their theories would

hold of necessity. To ensure an efficient market requires a large number of market participants with rational expectations. No single participant is permitted to buy or sell in quantity sufficient to significantly affect the price. It is also required that information be costless, and that transaction fees, the costs of buying or selling in the market, be nil (Carlozzi, 1983). The term rational expectations has a technical meaning.

The concept of rational expectations is based on the belief that economic agents are utility maximizers. Thus, market participants form expectations that fully reflect all available information. More formally, rational expectations imply that individuals' subjective probability distribution of possible outcomes is identical to the objective probability distributions that actually occur (Hafer, 1983).

Market efficiency, however, is more than a theoretical concept. Whether markets are efficient is not only an issue to be debated abstractly, but an empirical question to be tested with evidence from the real world. In order to transform the "pure" form into a testable hypothesis it is necessary to take into account the unreality of the initial postulates. Consequently, markets are not deemed inefficient if the "inefficiency" can be attributed to the cost of acquiring relevant information or to the cost of executing a transaction. Agreement on the meaning of the terms excess profit and consistently is also needed.

The difficulty of testing how closely a theory

approximates reality exists with respect to many questions, but it does not prevent reasonable men and women from agreeing if the evidence is sufficiently strong. A statement of the theory that permitted a test which could elicit such agreement might be termed the "operational form" of the EMT. It is this type of statement of the EMT that is considered relevant here.

There are "weak" and "strong" versions of the EMT. The weak version is sometimes called the random walk hypothesis; it asserts that market price changes cannot be distinguished from a statistical random walk in which changes are serially uncorrelated (after adjustment for any bias in the mean). The random walk hypothesis is "weak" in the sense that it allows only past price history to be used to obtain superior profits; in other words, it asserts that price history alone will not help one predict the course of future prices. The strong version of the efficient market theory goes further to assert that there is no public information helpful in predicting price changes.

These different statements of the EMT stress a common theme: markets are to be judged as efficient or not on the basis of some agreed-upon set of information. Since market efficiency must be empirically tested relative to a specified information set, it follows that a given market could be judged efficient with respect to one information

set but inefficient with respect to another.

The Importance of the Efficient Market Theory.

Because the question of market efficiency is a pivotal point in the economic theory of free-market capitalism, it is easy to understand why the theory has been the subject of so much debate and empirical testing. The argument, using Carlozzi's (1983) words, goes something like this:

In an efficient market the price "accurately reflects all relevant information." In other words, all the factors that matter to buyers and sellers in the market, including their expectations of future events, are built into the market price. In a sense, the price is always right in an efficient market....

The price changes generated by new information are viewed favorably by economists. They signal to everyone who looks at the price that something has occurred that calls for people to rethink their decisions on how to allocate their resources. When all markets are efficient the reactions of individuals to price changes will produce the best economy-wide allocation of resources--no one could be made better off without injuring someone else.

The social welfare aspect makes it clear that the efficient market theory is not just a technical question of interest only to specialists but one whose resolution carries ideological ramifications important in contexts other than price forecasting.

The focus of this research is not on the laissez-faire/central-planning controversy, it is on interest rate forecasting. But the the implications of the EMT for

interest rate forecasters are equally crucial: if the efficient market theory is correct in its "operational form," then it is not possible to forecast interest rates well enough to make meaningful profits.

Weaknesses of the Efficient Market Theory.

Although the EMT is an empirical hypothesis to be tested with evidence from the real world (some of which is discussed later in this chapter and in Chapter IV), the efficiency question is of sufficient importance that other arguments, not all of them economic, ought to be considered.

It is not necessary to argue against the EMT when it is put forth in either of two forms. In the "pure" form the theory is a statement about "an imaginary substance," and there is no need to make an issue.

At the other extreme, if the theory is interpreted merely to say that price forecasting is only difficult (what might be labeled the "asymptotic" form of the EMT), then there is no controversy. The author agrees that price forecasting is difficult and almost no one is so naive as to believe that markets routinely shower fortunes helter-skelter on casual investors.

Information. The EMT is a statement about the relationship of prices and information, "prices accurately reflect all relevant and available information."

But what if the data are wrong? Churchman (1977)

warns that bad data make for bad forecasts, and it is well known (Webb, 1983; Morgenstern, 1963) that most economic data are attributed a spurious accuracy when, in fact, they are plagued with errors of observation and conceptual ambiguities.

How much relevant information is available? "Masses of information, data, news, analyses, research reports, computer-assisted paraphernalia are available on virtually anything by way of investment products" (Gerbino, 1982; emphasis in the original). That is the problem. There is so much relevant data the its volume exceeds the ability of any individual or group of individuals to process it effectively.

Miller (1978) has done extensive research, from the level of a cell to the level of a society, on how systems respond when gluts of information produce what he calls information input overload. His general conclusion is that as more and more information is input to a system it finally becomes unable to process all of the data correctly. When this happens, stress occurs and the system compensates by adjusting its behavior.

The following appear to be the chief adjustment processes.

Omission--failing to transmit certain randomly distributed signals in a message.

Error--incorrectly transmitting certain signals in a message.

Queuing--delaying transmissions of

certain signals in a message, the sequence being temporarily stored until transmission.

Filtering--giving priority in processing to certain classes of messages.

Abstracting--processing a message with less than full detail.

Multiple channels--simultaneously transmitting messages over two or more parallel channels.

Escape--acting to cut off information input.

Chunking -- transmitting meaningful information in organized "chunks" of symbols rather than symbol by symbol (Miller, 1978; emphasis in the original).

Psychology. Information overload is also likely to produce confusion.

Confusion triggers off an immediate search for meaning or order to reduce anxiety inherent in any uncertain situation. The result is ... a readiness to assume causal connections even where such connections may appear to be quite nonsensical.... Anybody who is confused is likely to jump to conclusions by holding to the first apparently reliable piece of evidence that he detects through the fog of his confusion (Watzlawick, 1976).

Adam Smith (1967), despite the pseudonym, is a real market participant who complains that "most of what has been written about the market tells you the way it ought to be, and the successful investors I know do not hold to the way it ought to be, they simply go with what is." What determines market value, Gerald Loeb (1965) says, is not balance sheets and income statements but the "hopes and fears of humanity ... greed, ambition, ... stress and

strain, fashion ..."

Evidence on this score is to be found in Charles MacKay's (1932) classic, Extraordinary Popular Delusions and the Madness of Crowds, a fascinating account of the Holland tulip mania, the South Sea bubble, and other market anomalies. Other case histories are recounted by Charles Kindleberger (1978) in Manias, Panics, and Crashes. But one need only recall the frenzied precious metal speculation of 1980, when silver skyrocketed to \$50 per ounce only later to plunge below \$5, to realize that markets are sometimes more emotional than rational.

Regardless of one's views of his economic theories, John Maynard Keynes must be acknowledged as a very bright fellow and a perceptive observer of human behavior. He made fortunes for himself and his school, Kings College, through astute investment and speculation. Keynes (1983) was under no illusion about efficient markets.

Very few American investors buy any stock for the sake of something which is going to happen more than six months hence, even though its probability is exceedingly high; and it is out of taking advantage of the psychological peculiarity of theirs that most money is to be made.

The winners, says Keynes (1967), are those who "guess better than the crowd how the crowd will behave."

Make the (heroic) assumption that the data considered by the market participant are accurate and not

so voluminous as to overwhelm the system. Assume also that fear, greed, and the herd instinct are not operative so that the investor is coldly logical in his or her decision. Can it then be expected that the "individual's subjective probability distribution is identical to the objective probability distribution that actually occurs?" Not quite.

Tversky and Kahneman (1974) and Kahneman and Tversky (1979) demonstrate that there are extreme biases when people make judgments under conditions of uncertainty. They find an insensitivity to prior probabilities, misconception of chance, insensitivity to sample size, illusions of validity, misconceptions of regression, and insufficient adjustment to new information. Any of these errors will usually lead to subjective probability distributions that are extremely different from the objective distributions. The investigators are careful to point out that these biases and misconceptions are not attributable to motivational effects (such as wishful thinking or skewed payoffs) and that some errors are found as much among experienced researchers with extensive training in statistics as among statistically naive laymen.

Tversky and Kahneman's experiments dealt with situations where the objective probability distributions were relatively simple. In more complex circumstances

another difficulty compounds the problem of arriving at a correct view of the world--the counterintuitive behavior of large social-economic-financial systems. Forrester (1976) argues, as a basic theme, that the human mind is not adapted to interpreting how these systems behave. Since they contain multiloop nonlinear feedback channels, the cause-effect relationships are very hard to disentangle. He notes particularly that the short-term and long-term consequences of an innovation are frequently quite different. If Forrester is correct, it is not easy to see how markets can "accurately reflect" the consequences of a new and existing information.

None of these research results--physiological, psychological, or sociological--nor the testimony of the market participants themselves support the efficient market theory. Homo economicus is as "extinct" as the Piltdown Man.

Statistics. Four somewhat technical statistical notions bear on the efficiency question. The first goes to the very heart of the matter, to the concept randomness itself. Randomness is a curious, fundamentally undecidable proposition. It is easy to construct deterministic series which are statistically indistinguishable from random sequences. On the other hand any series, no matter how regular in appearance, could be produced by a random process (Gardner, 1968). G. Spencer-Brown (1957) considers

the profound difficulties and paradoxes that the issue of randomness raises.

Second, the idea that the "real world" may always be accurately represented by objective probability distributions is faulty. The real world is both risky and uncertain.

What's in a name? That which we call Uncertainty by any other name can be handled more easily. So Uncertainty were it but called Risk, would obtain that dear perfection without which scientific quantification and probability analysis is inapplicable (Davidson, 1978).

Davidson (but not Tversky and Kahneman) uses the word uncertainty in its technical sense to convey an understanding of a state such that it is impossible to assign probabilities to outcomes on the basis of either historical frequencies or subjective estimates. Risk, on the other hand, means a situation where, although the outcome is unknown, the probability distribution can be estimated. In the former case,

our decision to do something ... can only be taken as a result of animal spirits--of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities (Keynes, 1936).

The third notion is an important statistical law, the arc sine law, which seems to have received insufficient attention from the EMT's proponents and detractors alike. Feller (1957) examines the law in detail and demonstrates its counterintuitive nature. Consider

the cumulative result of a sequence of two-person zero-sum games in which an unbiased coin is tossed. The first player always bets heads, the second tails and the loser of each round pays the winner \$1. It might be expected that each player would, on a cumulative basis, be in the lead about one-half the time and that the lead would change hands fairly frequently. This is incorrect. Feller shows, for example, that in a game sequence of twenty rounds the probability that one player takes the lead and never relinquishes it exceeds .35.

There are warnings here for both camps. The efficient market advocate must see that it is not only possible for an investor to outperform the market on a long-term, cumulative basis, it is almost guaranteed that some will do so. The lesson for those who disdain the theory because of their personally superior market performance must be more willing to ascribe their good fortune to good luck.

The last statistical question is considered in connection with the empirical evidence presented below.

The EMT Paradox. The efficient market theory has a paradoxical nature and potential for self-falsification. If it were universally believed that nothing could be gained by security analysis and other attempts at price prediction, such efforts would cease; at that point, of course, these activities would again yield worthwhile

information. What kind of proposition is this, one which is claimed to be true only if it is not believed?

Right now of course, there is no danger of the efficient market theory being universally accepted. Large fees are charged by (and paid to) advisory services, brokers, investment fund managers and the like for forecasting future market prices. Clearly, many market participants with money to spend don't believe the academicians. The truth of the EMT will not be determined by a plurality of votes of market participants, but since economic reality is largely socially created, it can be perilous to ignore contrary views.

There is more paradox. The EMT says market participants have rational expectations. One such expectation (according to the theory itself) is that attempts to forecast security prices will not yield superior returns. Therefore, economic agents with rational expectations will not expend effort in attempting to forecast these prices. Even the most casual observation shows that many people do try to forecast prices. Consequently (again according to the theory itself) these people lack rational expectations. Thus the theory leads to its own denial; the snake eats its tail.

Empirical Tests of the Efficient Market Theory.

There is a vast literature on the efficient market theory. That there are so many studies and that they

continue to be made in increasing numbers suggests that the results are inconclusive. After all, if economists had surveyed 847 markets and concluded that they were all efficient (or, for that matter, all inefficient) who would be eager to run Test Number 848?

An important general review of the evidence is given by Fama (1970). More recent work, related particularly to interest rates, is discussed by Mullineaux (1981) and additional material (with many references) can be found in Gay and Kolb' (1982).

Some generalizations may be possible. First, with respect to the weak version of the EMT (the random walk hypothesis), the evidence indicates that most attempts to forecast price changes on the basis of their own history have failed, but some critics ("chartists" in the jargon of Wall Street) argue that the models employed have not been sufficiently sophisticated. Theory and evidence are given by Cootner (1964). The author's own attempts to forecast future interest rates on the basis of only their past behavior offer no convincing evidence that the random walk hypothesis should be rejected.

When the information set is broadened and the range of statistical techniques is increased, the picture becomes less clear. While many of these tests still fail to reject the EMT, a growing number suggest that there is an inefficient use of pertinent information in economic

forecasting.

One example, particularly relevant here, is from Hafer (1983).

The impact of unanticipated changes in the weekly money supply on short-term interest rates has been investigated extensively.... The evidence clearly indicates that unanticipated changes in the money stock have an important effect on interest rates....

Market participants still wager large sums and reputations on correctly anticipating the elusive weekly money figure.... Forecasters could have improved upon their ability to predict changes in the money stock by incorporating information contained in the series on loans, demand deposits and the adjusted base. Thus, over the recent period, the forecasts do not meet the broader efficiency criterion tested here.

Other recent evidence contradicting the EMT may be found in Brown and Maital (1981), Puglisi (1978), Lang and Rasche (1978), Branch (1978), Chow and Brophy (1978), Shiller (1979), Vignola and Dale (1979) and Gay and Kolb (1982).

In the author's opinion some proponents of the efficient market theory are too cavalier in their treatment of the empirical evidence and the conclusion it supposedly supports. Poole (1976), for example, says:

Numerous investigators have analyzed an enormous amount of data using many different statistical techniques, and no serious departures from the predictions of the (efficient market) hypothesis have been found. Thus, there is very strong evidence in favor of the hypothesis (emphasis added).

This is extreme indeed! First, there is contrary

evidence. Some very serious departures from the predictions of the theory are reported by Seligman (1983), a long-time partisan of the theory who confesses his own loss of faith.

Question: how close to reality is [the EMT]? Having now resurveyed the basic case made for it in the business schools, and also looked at some recent findings that seem inconsistent with it, I find myself still answering that [the EMT] is extremely useful for understanding the stock market--but doubting that it's as close to reality as I had previously assumed. It seems fairly clear that some superior investors are out there beating the market systematically.

Second, Poole seems to take liberties with the whole thrust of hypothesis testing as presented in any introductory statistical text. The hypothesis is set up in the form:

H_0 : Markets are efficient.

Then evidence from a sample is gathered and, if it is inconsistent with the hypothesis, leads to rejection of H_0 ; otherwise the hypothesis is "accepted." But,

it is important to understand that the rejection of a hypothesis is to conclude it is false, while the acceptance of a hypothesis merely implies that we have no evidence to believe otherwise (Walpole and Myers, 1978; emphasis added).

There is quite a difference between "strong evidence in favor" and "no reason to believe otherwise."

Moreover, efficient market tests are always made with respect to a given, usually quite limited, information set. To find that one or two variables are not useful in

price forecasting fails to justify the conclusion that other variables would fair no better.

The Difficulty of Forecasting.

It is one thing to argue that excess profits are possible, quite another to say they are easy to obtain. Listen to the legendary Bernard Baruch (1980), speculator extraordinaire:

If you are willing to give up everything else, and will study the market and every stock listed there as carefully as a student studies anatomy, and will glue your nose to the ticker tape at the opening of every day of the year and never take it off till night; if you can do all that, and in addition have the cool nerves of a gambler, the sixth sense of a clairvoyant and the courage of a lion--you have a Chinaman's chance.

To succeed on Wall Street, claims Gerbino (1982), requires grueling research, market savvy, and economic analysis.... The pro will push himself beyond the limits to be successful in any field. And that's only one of the steps to success on Wall Street. One must also search and find a technology and methodology of investing that works and then ... learn it and make it work.

Many of the arguments advanced above in opposition to the efficient market theory serve just as well to support the contention that price forecasting is likely to be very difficult. It has been argued here (contrary to the EMT) that market participants do not in the aggregate correctly perceive what prices should be. The logic works just as well at the level of the individual: if it is hard for the market as a whole to determine the correct price, it is not going to be easy for the participants as

individuals to do so. To summarize those arguments:

There is incorrect information.

There is too much information to process it all correctly.

Hopes and fears, greed and ambition determine market prices as much as economic fundamentals.

People are subject to extreme biases when making judgments under uncertainty; their estimates of the probability distributions are faulty.

The distinction between uncertainty and risk, the terms used technically, may invalidate much of the use of the concept of probability distributions in connection with the EMT.

The human mind is not adapted to interpreting correctly how some types of complex systems operate.

There is more. Not only is the economy extraordinarily complex, involving a large number of variables interacting with nonlinear feedbacks, it sometimes exhibits instability and discontinuity, with unpredictable results. The unstable nature of securities markets has been discussed by Zeeman (1974), who used catastrophe theory, a branch of differential topology, to deal with the discontinuities. Moreover, the economy is an evolving system, adapting and changing over time (Nelson, 1982). Georgescu-Roegen (1975) argues this point forcefully and proposes a biological rather than a mechanistic model of the economic

system. Evolution may proceed smoothly through natural selection or dramatically through mutation. The growth of the underground economy and the 1973 oil price increase, respectively, seem to fit these descriptions. As systems grow more complex, qualitatively new properties (emergents), unpredictable beforehand, come to the fore. The economy in the United States is a living, evolving system, the parts of which do not fully understand the whole.

Self-Altering Forecasts.

Were the interest rate forecaster to overcome all of these obstacles and succeed in predicting interest rates accurately, a Catch 22 still looms--the problem of the self-altering forecast. In one version, the self-altering phenomenon would assure the forecast; in another version, the forecast would be self-defeating. For either phenomenon to occur, the forecast must be public and must be believed.

In the first instance, market participants, believing the forecast and seeking financial gain, enter the market and through their buying or selling of financial instruments bring about the predicted change. The forecast becomes self-fulfilling. In the second case, one can imagine economic policymakers, say the Federal Reserve Board, confronting a forecast which they believe but do not like. In response they manipulate monetary policy to

prevent the forecast's realization. It becomes self-defeating. Henshel and Kennedy (1973) argue that "the purest example of self-altering predictions in economics is the operation of the stock market, with its self-generated markets reflecting investor confidence as well as objective conditions."

Not long ago the bond market provided a striking example of the self-fulfilling prediction phenomenon. Henry Kaufman is partner and chief economist of the brokerage firm, Salomon Brothers. Mr. Kaufman's views are respected in money markets and his interest rate forecasts are widely reported in the financial press. A Kaufman speech can move bond prices by several percentage points. The Wall Street Journal (1980) reported a self-fulfilling prediction by Mr. Kaufman on the front page:

Bond prices, which move inversely to interest rates, surged in response to two developments: the prime-rate cut by Chase and a few smaller banks, and a memorandum to portfolio managers from Henry Kaufman in which the oracular economist at Salomon Brothers withdrew his forecast for even-higher interest rates and conceded that lower rates might lie ahead. Some long-term Treasury bonds soared by as much as 5%, the biggest rise for any day on record.

Survival of the Fittest Forecasters.

When individuals possessed of differential ability engage in competition, the more able tend to win and the less able tend to lose. This is so evident from experience and common sense that Darwinian arguments are unnecessary.

In virtually every field of human activity those who possess superior information and process it more effectively outperform those who do not. Price forecasting is not an exception.

The implication is that some will consistently earn excess profits because they will be competing against others less able to acquire and interpret price-relevant information. But since they will also be competing against those of equal or greater ability, the excess profits will not be achieved easily.

Conclusion.

For all these reasons, the author views the efficient market theory as incorrect in its operational form. This is no minor quibble over the precise wording of the theory. The contention to be advanced and defended in this study is this: it is possible, albeit difficult, for some people to consistently outperform the market and earn financially significant excess profits.

Having emphasized this conclusion, the author does not necessarily suggest that economists abandon the hypothesis. Like the ideal gas or the corpuscular theory of light, the EMT is, for many purposes, a highly useful fiction and it will frequently be advantageous for economists to consider it as-if-true.

But given the evidence presented here, the interest rate forecaster has good reason to think of the EMT

as-if-false.

OTHER MARKET PERSPECTIVES

The efficient market theory is only one view of the securities markets, the one held by many economists but not generally shared by market participants themselves. The market is a confusing place. It is time to see how some of those involved attempt to make sense of the confusion.

The material of this section is based on the book by sociologist Charles A. Smith, The Mind of the Market (1981). The present author has added some additional distinctions, buried others, and otherwise reinterpreted some of the material, but the debt to Smith is clear.

Smith presents his different perspectives in the form of characters, archetypes, who are represented as holding the respective views. Most real people reflect more than one of the perspectives, or different ones at different times, or perhaps mutations of the pure strain. But some people, it is claimed, are fairly represented by the archetypes--these real people are the true believers.

The Fundamentalists.

If any perspective can be said to have a plurality of adherents among market participants, it is probably the fundamentalist view. The fundamentalist perspective stresses economic conditions as the determinant of market prices.

Economic conditions encompass the usual national and international macroeconomic variables as well as the balance sheets and earnings statements of specific companies. The fundamentalist expects to profit by buying undervalued securities and holding those securities until their true value is more generally recognized. Fundamentalists read and react to economic and political news. They follow the teachings of Graham and Dodd (1962) in Security Analysis, which emphasizes profits, book value, market share and other traditional financial ratios. They agree with the conservative market adage, "investigate, then invest."

The Chartists and Cyclists.

Chartists might be called anti-random-walkers. The chartist believes that the future of market prices can be divined from their past history, a position antithetical to the random walk hypothesis. Chartists keep charts, sometimes hundreds of charts on many variables--but always related to market statistics, not primary economic data. Chartists believe that their charts reveal patterns that portend future price movements.

In many ways cyclists are closely akin to the chartists (in fact, Smith lumps the two together as a single species). Cyclists believe that the market is ruled by a transcendental order and that prices rise or fall as if with the tides. Profits are made by understanding these

rhythms and acting in accordance with them.

While the fundamentalist see economic events and supply and demand forces influencing market prices, the chartists and cyclists are likely as not to reverse the order of causality. They give great weight to the market's power to determine both future economic conditions and the forces of supply and demand within the market itself.

The Insiders.

The insider generally agrees (with the efficient market theory) that prices cannot be forecast on the basis of publicly available information. Consequently, he works to develop personal contacts who can supply him with inside information. Profits are to be made by buying securities before their prices have reacted to the yet-to-be-disseminated information. The information itself is not so important as how it will be perceived and acted upon. As a result, the insider emphasizes the importance of "sponsorship" or lack thereof. What counts, he believes, are the buying and selling of big-money coalitions and large institutional investors. Winning is a result of being in the know.

The insider lives on "tips" and frequently skirts securities laws which prohibit the use of inside information to make profits.

The Traders.

The trader emphasizes his intuition and personal feelings about the market. He sees the market as much emotionally as cognitively. True, the market reacts to economic information, to the buying and selling pressures of large institutions, and to its own internal rhythms and patterns; but it is the correct interpretation of these factors that leads to profits. By being in tune with the psychological mood of the market, the trader is better able to act correctly on the news and information he has received.

The views held by Smith's market participants conform closely to the taxonomic scheme of Linstone--the fundamentalists, the chartists, and the cyclists hold Technical Perspectives; the insiders hold an Organizational Perspective; and traders maintain a Personal Perspective.

FORECASTING METHODS

Many methods can be used to try to forecast interest rates. One classification of these methods, intended to be broadly representative but certainly not exhaustive, is outlined below:

A. Judgmental Methods

B. Econometric Models

1. Single-Equation Models
2. Multiple-Equation Models

C. Economic Indicator Analysis

D. Time Series Analysis

1. Smoothing Methods
2. Box-Jenkins ARIMA Models
3. Stepwise Autoregression
4. Harrison-Stevens Bayesian Approach
5. Spectral Analysis

E. Market Analysis ("Technical") Methods

F. Pattern Recognition Approaches

Judgmental Methods.

All forecasting involves an ingredient of judgment and intuition. The choice of the forecasting method, the variables included in the model, and the functional form to represent the relationships are all matters of judgment. But here the term judgmental methods is used to encompass all approaches that do not involve explicit techniques; sometimes the approach is called "expert forecasting" or "intuitive forecasting."

The forecast may take the form of a simple statement, but more often it is incorporated within a scenario. The following are typical:

The release of the minutes of the October FOMC [the Federal Open Market Committee, the policy-making arm of the Federal Reserve Board] meeting and the press release accompanying last week's cut in the discount rate suggest that concern about the economy and international and domestic financial tensions have become increasingly important to FOMC members. Indeed, monetary policy over the next few months should be conducted with a bias toward accommodation. The path will prove somewhat erratic, but interest rates should continue downward well into 1983 (Mastrapasqua, 1982).

Behind its recovery-is-just-around-the-corner facade, the Administration is now joining the Fed in panicking. Treasury Secretary Regan wants to push the July 1st income tax cut ahead by six months. Such a speedup may or may not be more effective than the first two installments. But one thing is sure: It will further widen the Treasury's already monstrous fiscal gap and will thereby worsen the strain on the money market. In fact, interest rates have already started climbing anew. Earlier this week, three-month Treasury bills reached the highest level in three months and bond prices softened across the board. Note that this latest rise in rates is taking place in spite of the liberal Fed and the economy relapse. Implication: The demand for credit generated by deficit-ridden governments at all levels and by illiquid private borrowers is becoming too strong for the Fed to offset (Holt, 1982).

The lack of reproducibility, evidenced above, is one reason the approach is frequently criticized. The fallibility of human judgment (Tversky and Kahneman, 1974) and the inability to obtain quantitative confidence intervals or probability estimates are also used as arguments against the method. Nevertheless, sound judgment will always be called for and some retrospective studies of economic forecasting accuracy have given a slight edge to the judgmental approach over competing large-scale econometric models (McNees, 1973; Haitovsky, Treys, and Su, 1974).

Econometric Models.

Econometric models, as the term is used here, include only single- and multi-equation statistical regression models.

A multi-equation model is intended to represent the structure of the economy in greater or lesser detail. The so-called St. Louis model is relatively simple, involving perhaps a half-dozen variables and equations. This is dwarfed by many of the large-scale models developed by consulting firms like DRI and Chase Econometrics and educational institutions like MIT and Michigan, all of whose models run to hundreds of equations.

A single-equation regression model relates a dependent variable to a set of explanatory variables chosen on the basis of economic theory or hunch. The function is usually a linear combination of the explanatory variables. If the statistical assumptions are satisfied, then the ordinary least square regression technique yields parameter estimates that are "best," linear, and unbiased. Standard measures for developing confidence intervals and testing hypotheses are easily obtained. Unfortunately, the mathematical assumptions underlying the statistical basis of regression models are almost never satisfied. This point cannot be overemphasized just because the power, ease of application, and sophistication of the technique make it so seductive in comparison with cruder approaches.

Econometric regression methods are usually classified as causal; they purportedly represent cause-effect relationships in a quantitatively

sophisticated manner. This argument is too easily overemphasized. The method merely computes correlations. If the economy is a jungle of mutually causal feedbacks, a simple single-equation econometric model, even if it includes the right variables, will not be able to capture them.

Examples of attempts to forecast interest rates with single-equation econometric models may be found in Hunt (1973), Schott (1973, 1977) and Horan (1978).

Economic Indicator Analysis.

A general approach to forecasting is to relate the current situation to an analogous historical precedent. In technological forecasting the method is associated with the terms "quantified analogy" and "precursor events." In economics the method is widely used and goes under the label, "indicator approach."

In its simplest form it makes perfectly good sense. In many businesses, for example, orders precede production, which in turn precedes actual income from sales. It is not unreasonable, then, to forecast short-term gross income on the basis of recent orders.

In the late 1930's Wesley Mitchell and Arthur Burns at the National Bureau of Economic Research undertook a monumental empirical analysis of some 500 economic series, searching for timing regularities among them. Their original work, which focused on the business cycle has

been revised and expanded several times in the intervening years, but the conceptual basis remains unchanged. The idea is to classify economic series into those which consistently lead the economy in expansions and contractions (called, reasonably enough, leading indicators), those which more or less parallel the economy as a whole (the coincident indicators), and those which follow the rest of the economy in its ups and downs (the lagging indicators). The emphasis, it should be added, is on identifying turning points of the economy--peaks and troughs, as they are called--beginnings and endings of periods of recession and expansion.

Strongly empirical, the method virtually guarantees that the forecaster who employs it will have an up-to-date feel for what is actually going on in the economy. As might be expected from an approach so blatantly data-oriented, it has been a source of controversy almost from the beginning. In a classical confrontation of empiricists and rationalists, T.C. Koopmans' "Measurement Without Theory" (1947) attacked the work for giving inadequate attention to economic theory.

Interest rates have been identified as one of the lagging indicators in the economy, so the method has potential value. Cagan (1972) provides the historical record through 1971.

Time Series Analysis.

In each of the methods considered above, at least a trace of economics could be found. Most time series models forego explanation entirely. Prediction of a variable with time series models is usually based solely on the past behavior of that variable. Most time series methods are univariate, although multivariate time series models are becoming more common and may include causal explanatory variables. (These causal models usually are called transfer functions and intervention models.)

There are obvious problems with noncausal time series methods. No matter how sophisticated the mathematics, the essence of the method is extrapolation of the past behavior of the series. There is a total lack of underlying theory.

Burdened with such disadvantages, the method frequently performs surprisingly well. When time series forecasts are compared with those of the large structural models, the former often prove more accurate. Given their lower cost, smaller data requirements, relative ease of computation, and equal or superior accuracy, time series models deserve all the attention they have received.

There are many approaches to time series modeling.

Smoothing methods. Smoothing is generally accomplished by means of filters, which are linear transformations or weighted averages of historical data.

Smoothing techniques usually don't forecast very

well, but they are accurate enough when one must cheaply forecast 10,000 different items in inventory (Makridakis and Wheelwright, 1978).

Box-Jenkins autoregressive, integrated, moving average (ARIMA) models. The first step in developing a Box-Jenkins model is to transform the original series so that it becomes stationary. Stationarity requires that the series have a constant mean, constant variance and constant covariances over time. Stationarity can frequently be achieved by differencing the original series one or more times.

This transformed series is then represented as a linear combination of its own past values and past error terms. The use of past series values makes the method auto-regressive, accounting for the "AR" in the acronym. The use of a moving average of the past errors accounts for the "MA" part. If the original series was differenced, then its representation must be restored by summing or "integration." This gives the "I" in ARIMA. The objective of ARIMA models is to decompose a series into two parts, the first dependent on its own past history, the second part purely random and unpredictable. The method proceeds in stages. After stationarity has been achieved one uses correlations and partial correlations between lagged values of the series to "identify" the number of autoregressive and moving average terms. The next stage is to estimate the

model's parameters. The preliminary model thus obtained is then checked for adequacy. Recycling through the steps may be necessary.

The advantage of ARIMA models is that they have frequently yielded good short-term forecasts, in general better than those of other univariate techniques. Its major disadvantage, aside from those shared by all time series methods, is that it is relatively difficult to apply. Substantial judgment is required and the computations are not trivial. Both disadvantages have been reduced somewhat by the recent availability of package programs that lead the modeler through the process.

The standard reference is the Box-Jenkins (1970) original, Time Series Analysis: Forecasting and Control. It is easier to understand Nelson's (1973) presentation in Applied Time Series Analysis for Managerial Forecasting. Miller and Hickman (1973) report an unsuccessful attempt to forecast the interest rate on Treasury bills using ARIMA methods.

Stepwise autoregression. Stepwise autoregression proceeds as follows. At the first step the lagged value of the series which contributes the most toward explaining the variance of the series is introduced in the regression equation. At the second step the lagged variable that best improves the fit is added. The process is continued until further lagged variables fail to produce significant

improvement in fit. It frequently happens, however, that the addition of new lagged variables reduces below significance other lagged variables previously introduced. This difficulty is eliminated by specifying the order of variable introduction on subsequent iterations.

Stepwise autoregression is the "poor man's" approach to ARIMA forecasting. The moving average part is ignored; the parameter estimation procedure is simpler, using ordinary least squares regression rather than the Marquardt nonlinear optimization algorithm; and the procedure is computer-automated, requiring no analytical intervention. On occasion stepwise autoregression produces good forecasts.

Harrison-Stevens' Bayesian Approach. During the 1970's Harrison and Stevens (1971, 1976) developed a new approach to short-term time series forecasting based on Bayesian principles. Generally ignored on this continent until recently, it received critical acclaim in Europe.

The method assumes that the series can be represented by a "variable" straight line contaminated by noise. The line is variable in that it is subject to occasional changes in intercept and slope. In addition to normal observational noise, the series is sometimes subject to major transient shocks which perturb it momentarily. These events--intercept change, slope

change, transient shock--are modeled explicitly and successive data points are used to estimate the posterior probabilities at each point in time.

Its developers claim the following advantages:

The system can recognize, and respond appropriately to, transient errors and sudden changes in level and slope.

The system is truly adaptive in both level and slope, i.e. its sensitivity increases when changes occur (i.e. when uncertainty increases) and has a good response to transients.

The system produces not merely a single-figure forecast but a joint parameter distribution, thus expressing the inherent uncertainty of the estimates of level and slope (Harrison and Stevens, 1971).

Unfortunately, the algorithms are not simple, probability distributions must be estimated, and computer programs are not readily available.

Spectral analysis. The literature of spectral analysis, which considers a series represented in the frequency domain rather than the time domain, does not encourage its use as a forecasting tool. Kendall (1979) expresses the opinion that anything which can be achieved with spectral analysis is more easily and efficiently done by ARIMA techniques.

Market Analysis Methods.

What is called market analysis is known in the jargon of Wall Street as "technical analysis." Much of market analysis differs from the methods discussed above in that

it has failed to achieve academic respectability.

There are several schools of technicians. Perhaps the best known and most followed are the chartists who look for patterns in the price history of the commodities or securities they follow. In their charts are found "head-and-shoulder-formations," "flags," "pennants," "island reversals" and a host of other forms. Of these patterns most fundamentalists can make "neither head nor tail."

Some technical tools are more statistical in nature. Moving averages are used to identify trends and changes in trends, and forecasts made accordingly. For example, the forecast may be for higher prices if the 10-day moving price average exceeds the 40-day average.

Indexes are computed which measure rates of price change. They are called momentum indicators, oscillators, or relative strength indexes. For example, a six-month rate of change index might be computed as the current price less the price six months ago, or alternatively, as the current price divided by the price six months ago.

Most "technicians" would agree that forecasting on the basis of charts requires a fair measure of intuition and judgment.

Another school of technicians are the cyclists. Some of the seemingly sound cyclical studies have found statistically significant cycles in interest rates (Shirk,

various dates). The problem for forecasters is that the percentage of change in interest rates accounted for by the cycles are generally too small to allow meaningful profits.

Other cyclists go beyond empiricism to mysticism. Some believe that the market is ruled by a transcendental order and that prices rise or fall as if with the tides or cycles of the moon. Profits are to be made by understanding these rhythms and acting in accordance with them.

Are these "technical" tools useful or is the whole scheme mumbo-jumbo? One thing is certain. The forecaster who uses charts containing price history, momentum indicators, and moving averages has before him a record of that which he is trying to forecast. That cannot be harmful. Whether these measures foretell continues to be a matter of controversy (Levy, 1971).

It is not difficult to agree with Horn and Farah (1979):

Regardless of the opinion held as to the value of charting, many traders use this technique, and their market influence must not be ignored.... When tens of thousands of these individuals act in unison when certain price levels are penetrated, their action can have a very potent influence upon the market.

Representative literature includes Pring (1980), Kaufman (1980), Bernstein (1982), and Barnes (1979).

Pattern Recognition.

Pattern recognition (PR) is a relatively new and potentially powerful method for selecting and processing information.

For several reasons PR has not received wide attention or application in economics. Knowledge disseminates rapidly within a given discipline because people active in the field share a common vocabulary and read the same journals. More time is required to move knowledge across fields, however, because intellectual contacts are fewer and the process is inhibited by unfamiliar notation and specialized terminology. Pattern recognition originated and had its early development in engineering, where it continues to be used extensively. Kalman filtering, spectral analysis, and ARIMA modeling are other techniques that demonstrate the lag between the development of a method in engineering and its application in economics. Moreover, since economists deal primarily with continuous variables, they have less motivation to learn a new method which, like PR, is designed only to handle dependent variables that are categorical in nature.

Pattern recognition deals with classes or sets.

Recall the definition of set:

A set is a collection of definite distinct objects of our perception or of our thought, which are called elements of the set. (Breuer, 1958)

Pattern recognition is concerned with the broad

injunction: Given the description of an object, recognize it as coming from or belonging to a set.

This general problem comes in different forms. One subdivision of pattern recognition techniques is known as clustering. Clustering requires that a given set of heterogeneous objects be split into subsets that are in some sense more homogeneous. The goal is to discover these "more homogeneous" groups and develop a scheme for distinguishing them.

A second subdivision, called scene analysis, deals with highly complex problems like recognizing objects in a photograph. Because of the complexity involved, the picture is broken down into subobjects and sub-subobjects in a hierarchical or tree-like structure. Methods of this sort are referred to as syntactic or structural approaches.

Another subdivision within pattern recognition, the one of interest here, is pattern classification. Pattern classification takes a given description of an object and attempts to assign the object to its proper class. Some approaches to classification make assumptions about the probability distributions of different classes and attempt to estimate the parameters of these distributions. These approaches are termed parametric, statistical or Bayesian. Methods not requiring the estimation of the parameters of probability distributions are called

nonparametric or deterministic. Only nonparametric methods of pattern classification are used here.

PR and its extensions have been applied in many areas of human endeavor. One of the earliest tasks to be tackled was reading alpha-numeric characters, both typed and handwritten. An important advance in PR was made in the late 1950's when Rosenblatt (1957) devised a linear decision function modeled on a simplified understanding of the neuron. He called his classification machine a perceptron. The perceptron has the ability to learn from its classification errors and modify its structure so that future errors of the same type become less likely. Perceptrons and more complex classifiers have been used to interpret electrocardiograms, to analyze and classify chromosomes, to identify fingerprints, and to help monitor the safety of nuclear reactors (Tou and Gonzalez, 1974).

There have been only a few uses of pattern recognition in finance and economics. Blin (1973) used PR to examine questions of consumer preference and welfare economics. Fogler (1974) found PR helpful in forecasting industrial production. Felsen (1975a, 1975b) used a perceptron to forecast stock market averages and select stocks with above average potential for appreciation. In addition, there are reports, without documented evidence, that PR has been used to forecast commodity prices (Aronson, 1980; Liversidge, 1983).

The goal of pattern recognition, then, is to take a series or group of object descriptions and make "sense" of them. Sense means different things in different contexts. An optical character reader makes sense of a mark on a paper by recognizing it as the numeral three. A diagnostic machine makes sense of the P-QRS-T complex of an electrocardiogram by distinguishing those evidencing structural damage from normal patterns. A chromosome analyzer may make sense by building an evolutionary tree based on similarities and differences in the chromosome structure of animals. The current task is to make enough sense of economic data to be able to forecast the direction of change of interest rates on long-term bonds.

PR is the product of many unions. Psychology, linguistics, cybernetics, information theory, engineering, statistics and data processing have all played important roles in its conception and development. An eclectically conceived offspring, pattern recognition has given rise to both conflicting custodial claims and disavowals of parentage.

Practitioners of the subject tend to be a pragmatic lot. Heuristics and ad hoc procedures are welcomed when they work, not criticized for lack of rigor. Mathematical statisticians sometimes shudder at such approaches, but the problems encountered cannot always await elegant mathematical proofs. Each problem seems to have its own

peculiarities and is best approached with a flexible attitude. Nonetheless, many people in the PR field recognize the discipline's deficiencies and seek to remedy them, but they refuse, in Milton Friedman's phrase, to let the best become the enemy of the good.

CHAPTER III

PRINCIPLES OF PATTERN RECOGNITION

This chapter starts from basic principles and develops the subject of pattern recognition just far enough to derive the classifiers needed to forecast the direction of change in interest rates. The important results are the weighted least squares classifier (and the recursive method of updating it to take into account changes in the economic environment) and the hyperbolic tangent algorithm which can be used to refine the parameter estimates generated by the weighted least squares classifier.

PATTERNS AND FEATURES

A pattern may be defined as a description of an object, process, or event. A pattern represents and contains information about the object. It is the input data actually or potentially available concerning the object to be classified.

Consider an epidemiological screening program where the objective is to identify those persons who have a heightened susceptibility to some particular disease. A

participant's "work-up sheet" would provide data on the individual's sex (female), age (32), weight (120), and height (5' 2"). It would include information about her medical and family history, blood pressure, pulse rate and so on. Extensive testing could provide hundreds of measurements on her body chemistry. All of this data together constitutes a pattern and is potentially useful in distinguishing those who are susceptible to the disease from those who are not.

It is convenient to represent patterns as vectors of real numbers called pattern vectors. The participant's pattern vector, based on the data provided above, could be coded as (1, 32, 120, 5.17) where the numbers represent, respectively, her sex, age, weight and height. In general, the numerical elements of a pattern vector may be from ratio, interval, or ordinal scales or may be nominal in nature, representing qualitative aspects of the object (Stevens, 1968).

The first major problem in designing a pattern classifier is to determine which elements of the pattern are to be used in the classification. The elements actually chosen are called features and the task of choosing them is called feature selection. One aspect of feature selection is economic and practical: some desirable features may be too difficult or too expensive to obtain. Complex, invasive clinical tests, for example,

would not be appropriate if an disease being screened for were not life-threatening.

Other aspects of feature selection have to do with technical considerations. At first it would seem that all pattern elements possessing any independent information useful in classifying the pattern should be utilized by the classifier. This turns out to be erroneous unless the number of samples is infinite (Duda and Hart, 1973) and in fact the optimal number of features is remarkably small. The goal is to obtain discriminating information while reducing redundancy and noise. The only guaranteed way to find the best subset from a set of properties is to try all possible combinations. This is impractical for even a modest number of properties and computationally impossible for larger sets. Mucciardi and Gose (1971) have studied this problem and tested seven techniques of pattern selection. Two desirable attributes of discriminating features are that they be individually capable of classifying with low error probabilities and that they be relatively uncorrelated. Mucciardi and Gose propose an index composed of a weighted sum of these two attributes.

Features need not be limited to simple unaltered choices from the elements of the pattern. Mathematical transformations and functional combinations of pattern elements frequently provide features with greater discriminating power.

Features, like patterns, are conveniently represented as vectors of real numbers. These are called feature vectors. In this paper lower case bold letters will always represent vectors; thus \mathbf{w} , \mathbf{x} , and \mathbf{y} are vectors. Feature vectors will always be assumed to be $(n \times 1)$ column vectors and will be denoted by \mathbf{x}_0 , where

$$(1) \quad \mathbf{x}_0 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

and x_i represents the i th feature. The notation

$$(2) \quad \mathbf{x}_0 = (x_1, x_2, \dots, x_n)'$$

where the prime (') indicates transposition, is equivalent.

Similarly,

$$(3) \quad \mathbf{x}'_0 = (x_1, x_2, \dots, x_n)$$

For reference, the principal symbols used in this work are listed in Appendix A, Summary of Notation.

CLASSIFIERS

The second major problem in pattern classification is to determine an algorithm, called a classifier, which assigns a pattern to a set or class based on the information contained in the feature vector. The classes are assumed to be mutually exclusive and will be denoted by C_1, C_2, \dots, C_M , where M is the number of classes. Since the direction of

change of interest rates can only be up or down (ignoring the case of no change), two classes will suffice here; but most of what follows can be generalized to problems with more than two classes.

If the feature vectors are considered to be points in an n -dimensional Euclidean space, then (in the two-class case) the classification problem can be seen as one of partitioning that space into regions such that all points in one class lie on one side of the partition boundary, while the rest of the points lie on the other side. A simple case is illustrated in Figure 1 where the feature space is a plane, the feature vectors from classes C_1 and C_2 are distinctly marked, and the partition boundary is the straight line.

Let

$$(4) \quad d(x) = d(x_1, x_2, \dots, x_n) = 0$$

be the equation of any partition boundary that correctly partitions the space so that the points in the two different classes fall on different sides of the boundary. Then it turns out that the function

$$(5) \quad d(x) = d(x_1, x_2, \dots, x_n)$$

evaluated at the points in feature space will be greater than zero for all points in one of the classes and less than zero for all points in the other. Consequently, $d(x)$ is useful as a classifier and is called a decision function or a discriminant function. Any continuous

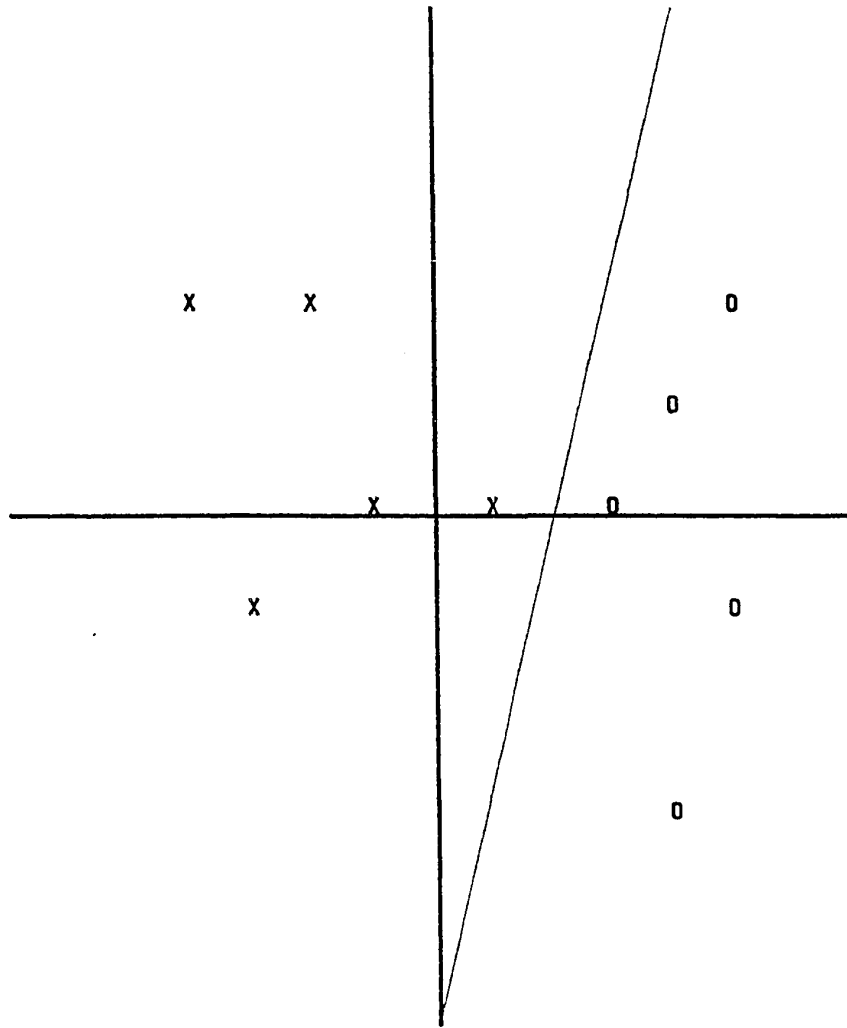


Figure 1. Linearly separable features in a two-dimensional feature space.

function can be used as a decision function, but it is frequently convenient to restrict attention to functions that are linear in x_i . Since the individual features, x_i , may be combinations or transformations of the original pattern elements, the linear limitation is much less restrictive than it first appears. In the two-dimensional illustration of Figure 1 the linear decision boundary is a straight line; in higher dimensions, linear decision boundaries are planes or hyperplanes. Nonlinear decision functions are not discussed here.

The general form of a linear discriminant function in n dimensions is

$$\begin{aligned} d(\mathbf{x}) &= w_1x_1 + w_2x_2 + \dots + w_nx_n + w_{n+1} \\ (6) \quad &= \mathbf{w}_0' \mathbf{x}_0 + w_{n+1} \end{aligned}$$

where the vector $\mathbf{w}_0 = (w_1, w_2, \dots, w_n)'$ is called the parameter vector or weight vector. In order to employ vector notation it is conventional in pattern recognition to augment the the vectors by appending a 1 to the feature vector and w_{n+1} to the weight vector. In what follows the terms feature vector and weight vector refer to their augmented versions, and the subscript $(_0)$ is dropped, i.e.,

$$(7) \quad \mathbf{x} = (x_1, x_2, \dots, x_n, 1)'$$

and

$$(8) \quad \mathbf{w} = (w_1, w_2, \dots, w_n, w_{n+1})'$$

With this notation the decision or discriminant function

is

$$(9) \quad d(x) = w'x$$

Having chosen the form of the discriminant function, one tries to determine values for the weight vector so that the decision function partitions the feature space correctly. A set of feature vectors (representing elements of the two different classes) which can be correctly partitioned by a linear decision function is called a linearly separable set. In practice this partitioning objective is frequently unattainable. Figure 2 illustrates a situation where the elements of the two classes cannot be separated by a linear decision function in the given feature space.

There are many methods of computing a weight vector. An important group, which take into account statistical knowledge about the relative frequency of the features within the classes, are the parametric methods. As standard references (Tou and Gonzalez, 1974; Young and Calvert, 1974; Duda and Hart, 1973) demonstrate, these approaches rest on a sound basis of statistical decision theory and can be shown to possess desirable optimality properties, when the underlying assumptions are valid. One of the important assumptions required is that the statistical distributions remain stable.

The present author holds the opinion that the U.S. economy is not stable in the usual statistical sense; it is evolutionary. What is needed here, then, is a method

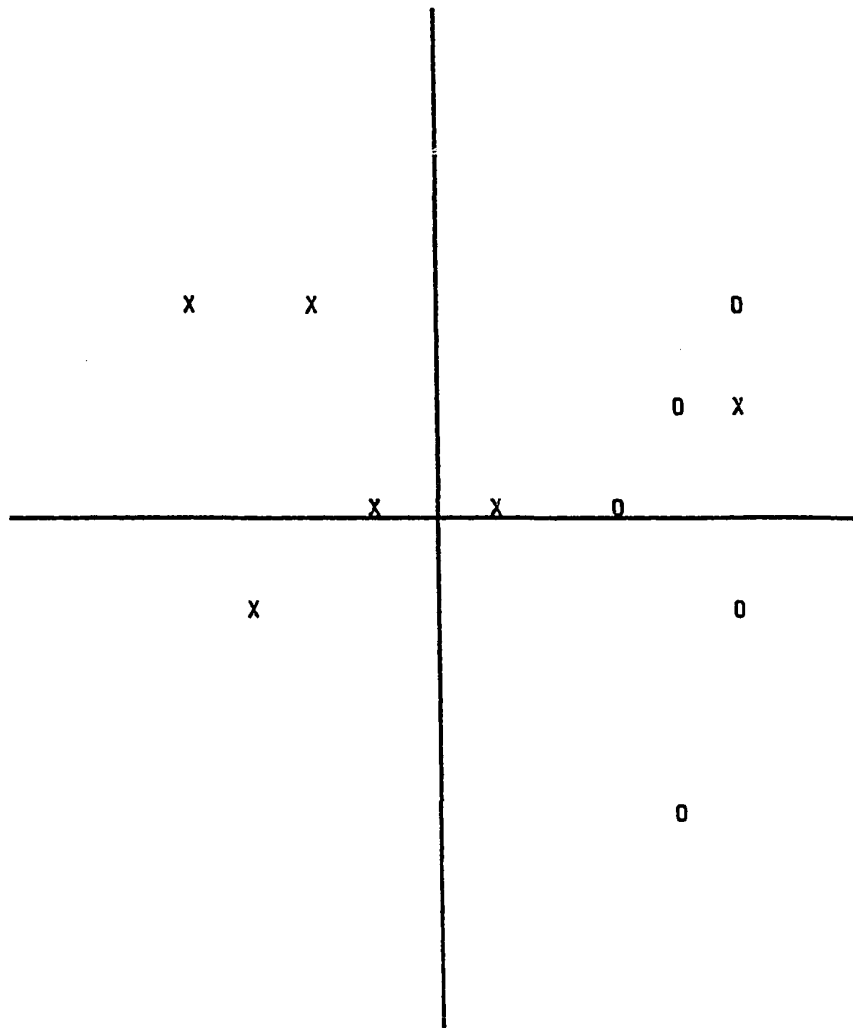


Figure 2. Linearly inseparable features in a two-dimensional feature space.

of determining the weight vector which can learn from its errors and change with the evolving economy. In a more general context, Jantsch (1976) discusses what is required if a system is to learn.

The prerequisite for learning is a certain plasticity of the system which allows it to have a history. In evolving systems ... each state depends on the past There exists therefore a "system memory"

One group of approaches for determining weight vectors which satisfy this prerequisite is the class of nonparametric methods. With these methods, each new feature vector is processed using the then-current weight vector. The computed class is compared with the actual class and if an error is made, the classifier adjusts its weight vector so that similar errors become less likely. Each state of the classifier depends on the past and the system memory is retained in the values of the weight vector. Advanced versions of this type of classifier can also extract information from correctly classified feature vectors, that is, the performance of these classifiers may be expected to improve even without their having made a classification error.

The pattern classification process is represented schematically in Figure 3.



Figure 3. The pattern classification process.

THE PERCEPTRON

In this subsection the perceptron algorithm is derived and discussed not only because it is historically important but also because it clearly illustrates important principles.

Let $(x(1), x(2), \dots, x(n))$ be a set of augmented feature vectors for which the correct classification between two sets is known. Assume the set is linearly separable, i.e., that there exists at least one weight vector w_* , called a solution vector, such that $w_*'x(i) > 0$ for all $x(i)$ in class C_1 and $w_*'x(i) < 0$ for all $x(i)$ in class C_2 . Let $w(1)$ be an arbitrarily chosen initial estimate of w_* . Successive estimates are denoted $w(2), w(3), \dots$. Take the first feature vector, $x(1)$, compute $w'(1)x(1)$ and proceed according to the following rule:

If the result is positive and $x(1)$ is a member of class C_1 , then the pattern is correctly classified and no change is made to w , so $w(2) = w(1)$.

If the result is negative and $x(1)$ is a member of C_2 , then the pattern is correctly classified so no change is required and $w(2) = w(1)$.

If the result is positive or zero and $x(1)$ is a member of C_2 , then an error has occurred.

Set $w(2) = w(1) - ax(1)$, where a is an arbitrary positive constant.

If the result is negative or zero and $x(1)$ is a member of C_1 , there is an error. Set $w(2) = w(1) + ax(1)$.

This process is continued with $x(2)$, $x(3)$, ..., $x(n)$, recycling through the whole set as many times as is necessary to correctly classify all patterns. An example of the process is given below. It can be shown (Tou and Gonzalez, 1974) that if the set is linearly separable, the process will correctly assign all feature vectors after a finite number of steps. If the set is not linearly separable, the algorithm will oscillate forever.

The algorithm can be expressed more concisely. Define a new set of vectors ($y(1)$, $y(2)$, ..., $y(n)$) equivalent to the old group of feature vectors except that those which are members of C_2 are multiplied by -1 . In other words, $y(i) = x(i)$ if $x(i)$ is a member of class C_1 , and $y(i) = -x(i)$ if $x(i)$ is in C_2 .

The condition for correct classification, $w'y(i) > 0$ for all $y(i)$, is obviously equivalent to $w'x(i) > 0$ for $x(i)$ in class C_1 and $w'x(i) < 0$ for $x(i)$ in C_2 .

Then the perceptron algorithm is

$$\begin{aligned}
 (10) \quad w(k+1) &= w(k) && \text{if } w'(k)y(k) > 0 \\
 &= w(k) + ay(k) && \text{if } w'(k)y(k) \leq 0
 \end{aligned}$$

where the vectors $x(k)$ in class C_2 have been multiplied by -1 to obtain $y(k)$, a is an arbitrary positive constant and $w(1)$ is chosen arbitrarily.

The perceptron can be viewed as a reward-punishment scheme where the algorithm is "punished" by having its weight vector changed every time it makes a mistake and "rewarded" by being left alone if it classifies correctly.

The following simple example, illustrated in Figure 4, shows the operation of the perceptron algorithm. Let the feature vectors be $(-1,-2)$, $(0,1)$, $(-1,-3)$, $(0,0)$, and $(2,2)$. Then the augmented feature vectors are

$$x(1) = (-1,-2,1)$$

$$x(2) = (0,1,1)$$

$$x(3) = (-1,-3,1)$$

$$x(4) = (0,0,1)$$

$$x(5) = (2,2,1)$$

Let $x(1)$, $x(3)$ and $x(4)$ belong to class C_1 and let $x(2)$ and $x(5)$ belong to C_2 . Multiplying the vectors in C_2 by -1 , the full set of y vectors is

$$y(1) = (-1,-2,1)$$

$$y(2) = (0,-1,-1)$$

$$y(3) = (-1,-3,1)$$

$$y(4) = (0,0,1)$$

$$y(5) = (-2,-2,-1)$$

Arbitrarily let $w(1) = (1,1,1)'$ and $a = 1$. Then, applying the algorithm,

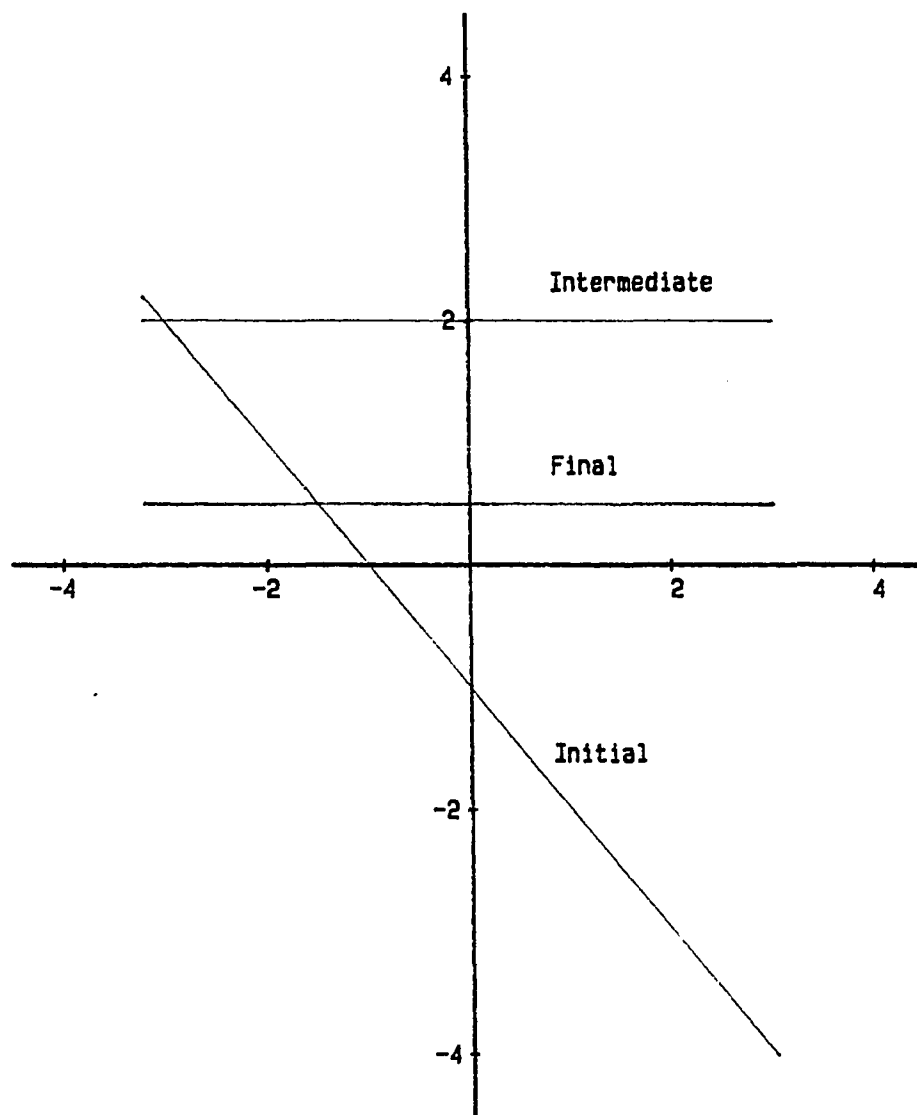


Figure 4. Evolution of the perceptron partition boundary in a two-dimensional feature space.

$$w'(1)y(1) = (1,1,1)(-1,-2,1)' = -1-2+1 = -2$$

$$w(2) = w(1) + ay(1)$$

$$= (1,1,1)' + (1)(-1,-2,1)' = (0,-1,2)'$$

$$w'(2)y(2) = (0,-1,2)(0,-1,-1)' = 0+1-2 = -1$$

$$w(3) = w(2) + ay(2)$$

$$= (0,-1,2) + (1)(0,-1,-1)' = (0,-2,1)'$$

$$w'(3)y(3) = (0,-2,1)(-1,-3,1)' = 0+6+1 = 7$$

$$w(4) = w(3) = (0,-2,1)'$$

$$w'(4)y(4) = (0,-2,1)(0,0,1)' = 0+0+1 = 1$$

$$w(5) = w(4) = (0,-2,1)'$$

$$w'(5)y(5) = (0,-2,1)(-2,-2,-1)' = 0+4-1 = 3$$

$$w(6) = w(5) = (0,-2,1)'$$

$$w'(6)y(1) = (0,-2,1)(-1,-2,1)' = 0+4+1 = 5$$

$$w(7) = w(6) = (0,-2,1)'$$

$$w'(7)y(2) = (0,-2,1)(0,-1,-1)' = 0+2-1 = 1$$

Corrections were made after the first and second steps because the vectors $y(1)$ and $y(2)$ were misclassified. The final weight vector $w = w(7) = (0,-2,1)$ correctly classifies all feature vectors. The corresponding decision function is $d(x)$,

$$(11) \quad d(x) = w'x = -2x_2 + 1$$

and the decision boundary is the equation

$$(12) \quad d(x) = -2x_2 + 1 = 0$$

Confusion over notation is avoided by remembering that $x(i)$ and $y(i)$ are vectors, but x_i and y_i are components of vectors.

The perceptron is just one of a family of iterative algorithms that can be derived analytically from a properly defined criterion function, $J(w)$, which is a mathematical specification of the objective to be achieved. In other fields the criterion function is sometimes called a cost function, objective function, or a loss function.

Consider a particular criterion function J_p defined by

$$(13) \quad J_p(w) = .5 (|w'y| - w'y)$$

where $|w'y|$ is the absolute value of $w'y$ and it is assumed that the feature vectors are linearly separable and that the augmented feature vectors belonging to class C_2 have been multiplied by -1 to obtain y . From the previous discussion it is known that if $w'y > 0$, then the pattern is correctly classified. Note that the criterion function, J_p , is minimized if $w'y > 0$. Ignoring the trivial case where $w = 0$, the zero vector, it is clear that the function will be minimized when a pattern is correctly assigned. The converse is also true. Consequently, minimization of J_p provides a weight vector that correctly classifies y .

Of the many minimization techniques that have been developed (Beveridge and Schechter, 1970; Cooper and Steinberg, 1970), a simple approach employing gradients is used here to develop a recursive formula for finding the

minimum of J_p . Recall from the calculus that the gradient vector is defined as the partial derivatives of a function with respect to its arguments.

$$\begin{aligned}
 \text{grad } f(\mathbf{z}) &= \frac{df(\mathbf{z})}{d\mathbf{z}} \\
 (14) \qquad &= \left(\frac{df}{dz_1}, \frac{df}{dz_2}, \dots, \frac{df}{dz_n} \right)'
 \end{aligned}$$

The gradient of a scalar function with a vector argument is a vector. The components of the vector give the rate of change of the function in the direction of the corresponding function argument. The gradient vector has the property of pointing in the direction of greatest increase of the function when the arguments increase; conversely, the negative of the gradient vector points in the direction of steepest descent.

Starting with an arbitrary weight vector, $\mathbf{w}(1)$, the minimization scheme increments \mathbf{w} in the direction of steepest descent, i.e., in the direction of the negative of the gradient vector. Mathematically, this process can be represented, at any step, by

$$(15) \qquad \mathbf{w}(k+1) = \mathbf{w}(k) - a \text{ grad } J(\mathbf{w})$$

where the gradient is evaluated at $\mathbf{w}(k)$ and a is an arbitrary positive constant. Call (15) the general gradient algorithm.

Now apply the general gradient algorithm to the

criterion function given in (13). The gradient is

$$(16) \quad \text{grad } J_p(w) = .5 (y \text{sgn}(w'y) - y)$$

where the sign function, $\text{sgn}(w'y)$, is defined as

$$(17) \quad \begin{aligned} \text{sgn}(w'y) &= 1 && \text{if } w'y > 0 \\ &= -1 && \text{if } w'y \leq 0 \end{aligned}$$

Substitution of (16) in (15) yields

$$(18) \quad \begin{aligned} w(k+1) &= w(k) - .5a (y \text{sgn}(w'y) - y) \\ &= w(k) && \text{if } w'y > 0 \\ &= w(k) + ay && \text{if } w'y \leq 0 \end{aligned}$$

Obviously this is the perceptron algorithm.

As a second example of the application of the general gradient algorithm, consider a criterion function which is one-half the square of the one used above, i.e., let

$$(19) \quad J_R(w) = .125 (|w'y| - w'y)^2$$

The gradient is

$$(20) \quad \text{grad } J_R(w) = .25 (|w'y| - w'y)(y \text{sgn}(w'y) - y)$$

Substitution in the general gradient algorithm (15) gives

$$(21) \quad \begin{aligned} w(k+1) &= w(k) - .25a (|w'y| - w'y)(y \text{sgn}(w'y) - y) \\ &= w(k) && \text{if } w'y > 0 \\ &= w(k) + ay |w'y| && \text{if } w'y \leq 0 \end{aligned}$$

This is called the relaxation algorithm.

Many algorithms of this type have been developed and are limited only by the ability to provide meaningful criterion functions.

The classifiers considered above are recursive

schemes which are designed by processing one feature vector at a time and making adjustments to the weights only if the pattern is misclassified. In nonseparable situations, the process does not converge. Even in linearly separable cases, the number of steps required to reach a solution may be very large. In practice one is seldom able to know in advance whether the set is linearly separable. This gives rise to the obvious problem of deciding when to stop the operation of the algorithm. In the separable case one may stop too soon; in the nonseparable case one may stop at a point where the weight vector is far from its optimal value. Moreover, "good" choices for the initial weight vector, $w(1)$, and the correction constant, a , as well as proper scaling of the components of the feature vectors, significantly affect the ability of these classifiers to move rapidly toward a solution. Finally, the solution vector depends on the order in which the feature vectors are presented to the algorithm.

The minimum squared error classifier developed below overcomes most of the perceptron's disadvantages by analytically determining a unique weight vector using all feature vectors simultaneously.

THE MINIMUM SQUARED ERROR CLASSIFIER

The presentation and programming of the minimum squared error classifier is simplified by matrix notation

(Anscombe, 1981; Helzer, 1983). Uppercase bold letters at the end of the alphabet represent matrices; thus X and Y are matrices. The transpose of a matrix X is denoted X' and its inverse, if defined, is X^{-1} .

With this notation let Y represent a matrix the rows of which are the vectors y' , previously defined. The classification problem requires that a solution vector w be found such that

$$(22) \quad Yw > 0$$

where 0 is the zero vector. Now consider the more stringent, but mathematically and computationally more tractable task of finding a vector w that satisfies the set of simultaneous linear equations

$$(23) \quad Yw = a$$

where $a_i > 0$ for all i . Clearly a solution to (23) is a solution to (22).

Let the matrix Y be of size $(t \times n+1)$, that is, let there be t equations in $n+1$ unknowns (w_1, w_2, \dots, w_{n+1}) where, in general, t is significantly greater than $n+1$. In other words, there are t patterns of known classification, each represented by a vector of n features and the augmentation constant 1 . The system of equations (23) will generally be overdetermined and no exact solution will exist. Consequently, an approximate solution that minimizes some function of the errors, defined as

$$(24) \quad e_i = a_i - w'y(i)$$

is required. An obvious goal would be to minimize the sum of the absolute values of the errors,

$$(25) \quad \sum_i |e_i| = \sum_i |(a_i - w'y(i))|$$

but this is computationally difficult. A somewhat less satisfactory alternative is to minimize the sum of the squared errors, represented by the criterion function,

$$(26) \quad J_S(w) = \sum_i (a_i - w'y(i))^2 = (a - Yw)'(a - Yw)$$

The gradient of J_S is

$$(27) \quad \text{grad } J_S = -2(Y'a - Y'Yw)$$

Since there are no constraints, the gradient can be set equal to zero and the equations solved for the vector w which minimizes the criterion function. Hence, equating the gradient to 0 and manipulating,

$$(28) \quad w = (Y'Y)^{-1}Y'a$$

The matrix $(Y'Y)^{-1}Y'$ is sometimes called the pseudoinverse or generalized inverse of Y and will be denoted by $Y^\#$. The vector w is the solution to the so-called normal equations of ordinary least squares regression. If a is set equal to 1, the unit vector, Duda and Hart (1973) show that (28) yields an asymptotic approximation to an optimal classifier.

An equivalent formulation, using the original feature vectors x , is

$$(29) \quad w = (X'X)^{-1}X'b = X^\#b$$

where the rows of X are the vectors x' and b is a vector such that $b_i = 1$ if $x(i)$ is a member of class

C_1 and $b_i = -1$ if $x(i)$ is in class C_2 .

The weight vector determined by this classifier is not guaranteed to correctly classify the patterns even when they are linearly separable. The algorithm provides a unique solution, however, and represents a reasonable compromise for separable and nonseparable pattern classification problems. The minimum squared error algorithm has received wide attention in the literature. With certain modifications, it will be applied here to the interest rate forecasting problem.

To illustrate the operation of the minimum squared error classifier, reconsider the example presented above. The matrix X and its generalized inverse $X^\#$ are, respectively,

$$X = \begin{vmatrix} -1 & -2 & 1 \\ 0 & 1 & 1 \\ -1 & -3 & 1 \\ 0 & 0 & 1 \\ 2 & 2 & 1 \end{vmatrix}$$

$$X^\# = (.009) \begin{vmatrix} -14 & -63 & 31 & -18 & 64 \\ -3 & 42 & -33 & 12 & -18 \\ 21 & 39 & 9 & 27 & 15 \end{vmatrix}$$

and the vector b is $(1, -1, 1, 1, -1)'$. The solution vector is

$$w = (-.018, -.432, .027)'$$

The example is presented graphically in Figure 5. Different solutions can be provided by the perceptron and the minimum squared error classifier, particularly when only a few points are to be separated.

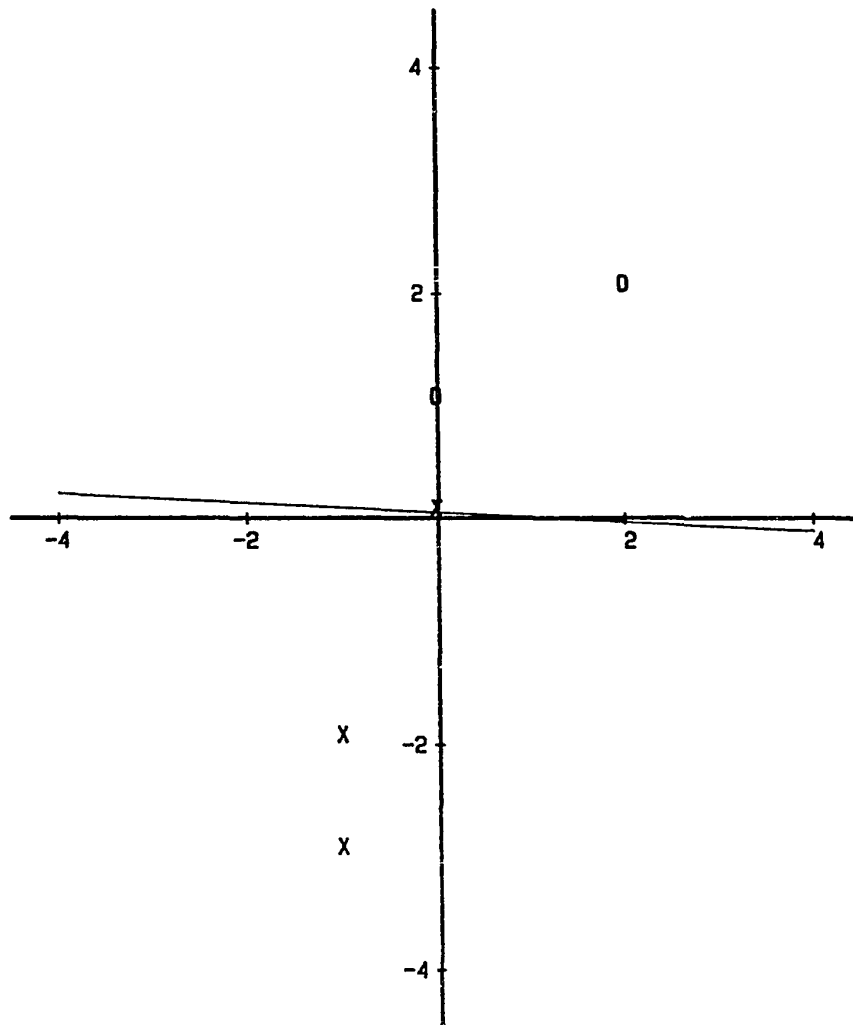


Figure 5. Partition boundary for the minimum squared error algorithm in a two-dimensional feature space.

WEIGHTED LINEAR CLASSIFIERS

It was suggested in Chapter II that actual pattern recognition problems seldom come so nicely packaged that standard approaches can be routinely applied. Rather, each problem has its unique aspects that may demand the creative use of ad hoc procedures and specially designed algorithms. Now it becomes necessary to leave behind the textbook techniques that have characterized the PR presentation so far and develop methods that help with the actual interest rate forecasting problem.

What should be the objective of an interest rate forecaster? Ideally one would prefer an accurate point forecast of the interest rate at some future date. An alternative, which suffices equally well in many practical decision situations, is to forecast instead the direction of change in interest rates. Given this more modest objective, how might performance be judged? One frequently used measure is the probability of error, but this is too limited since all errors are not equally costly. In other words, it is possible to conceive of situations where the direction of change is forecast correctly when the amount of change is small, but incorrectly when the amount of change is large. Clearly some weighting scheme is needed.

An initially reasonable demand on an interest rate forecaster is that he be required always to have a forecast about the direction of future interest rate

movement, i.e., "no opinion" is not allowed. An initially reasonable utility function is linear, so that the marginal utility (disutility) of a gain (loss) is constant. Under such conditions the forecaster will try to forecast so as to maximize the sum of the gains less the sum of the losses incurred by following the forecasts. An initially reasonable assumption is that gains and losses are proportional to the amount of change in the interest rate.

A criterion function, J_N , incorporating these conditions and assumptions is

$$(30) \quad J_N(\mathbf{w}) = \sum_i |v_i| - \sum_i |v_i| |\operatorname{sgn}(v_i) - \operatorname{sgn}(\mathbf{w}'\mathbf{x})|$$

where $|v_i|$ is the magnitude of the change in the interest rate and the rest of the symbols are as previously defined.

If $\operatorname{sgn}(v_i) = \operatorname{sgn}(\mathbf{w}'\mathbf{x})$, then the forecast of the direction of change is correct, the second term in the expression above is zero, and the contribution to the criterion function by the first term is $|v_i|$, the gain achieved by the forecast. If the signs are not equal, the forecast is in error and the net contribution to J_N is $-|v_i|$, the amount lost by following the forecast. Thus maximizing J_N maximizes the net gain.

It is equivalent to minimize the sum of the losses. The appropriate criterion function to be minimized is

$$(31) \quad J_L(\mathbf{w}) = .5 \sum_i |v_i| |\operatorname{sgn}(v_i) - \operatorname{sgn}(\mathbf{w}'\mathbf{x}(i))|$$

Unfortunately, the presence of the term, $\operatorname{sgn}(\mathbf{w}'\mathbf{x}(i))$,

makes either of these two criterion functions, (30) or (31), discontinuous. Consequently, the gradient is not everywhere defined and the optimization methods previously used cannot be directly applied. One way of overcoming the difficulty is to substitute a continuous approximation for $\text{sgn}(\mathbf{w}'\mathbf{x}(i))$. Many functions might serve satisfactorily. A particularly convenient choice uses the hyperbolic tangent

$$(32) \quad \tanh z = \frac{\exp(z) - \exp(-z)}{\exp(z) + \exp(-z)}$$

Let s be an arbitrary positive constant and let $\tanh(\mathbf{sw}'\mathbf{x}(i))$ replace $\text{sgn}(\mathbf{w}'\mathbf{x}(i))$ in (31) above.

Define J_T as

$$(33) \quad J_T(\mathbf{w}) = .5 \sum_i |v_i| |\text{sgn}(v_i) - \tanh(\mathbf{sw}'\mathbf{x}(i))|$$

Call this the hyperbolic tangent algorithm. It is easy to show that J_T approaches J_L in the limit as s approaches infinity. Figure 6 illustrates the approximation of $\text{sgn}(z)$ by $\tanh(sz)$.

The gradient of J_T is

$$(34) \quad \text{grad } J_T = -.5s \sum_i |v_i| \mathbf{x}(i) \text{sech}^2(\mathbf{sw}'\mathbf{x}(i))$$

Conceptually it is easy to set the gradient equal to zero and solve the simultaneous equations for a minimum. In practice, this is difficult. Because the criterion function has many local minima, other optimization techniques encounter problems as well. These minimization difficulties, particularly the possibility of achieving a local

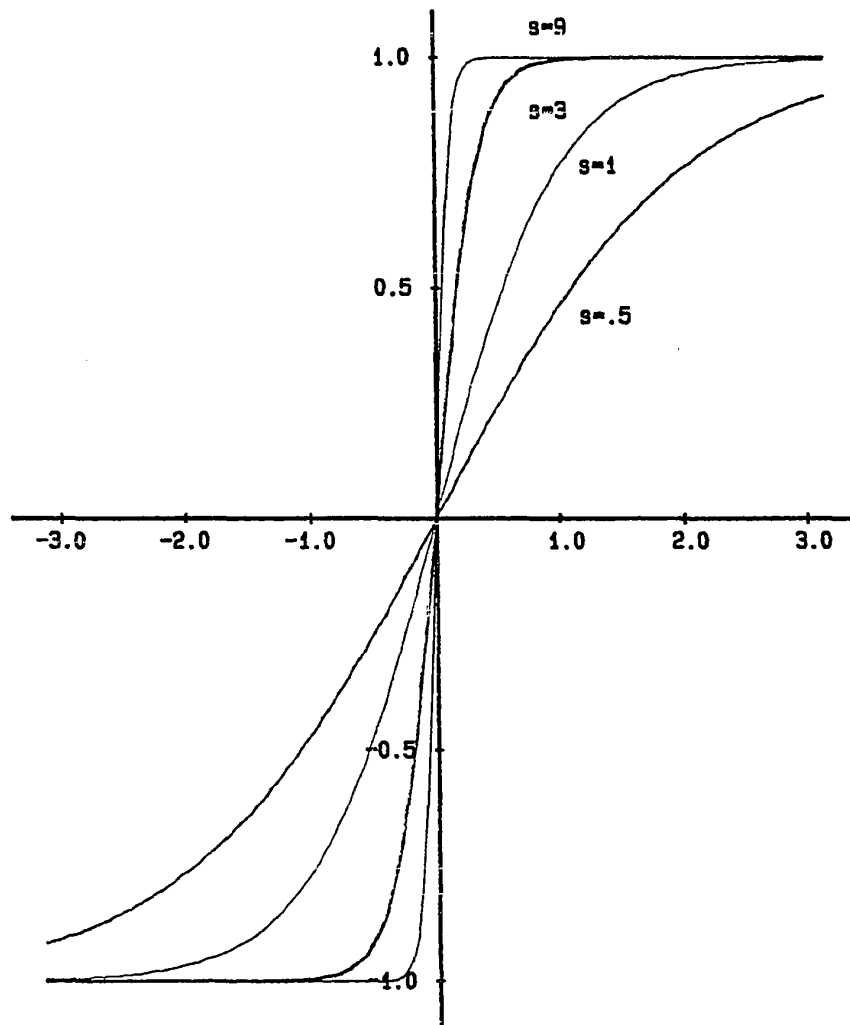


Figure 6. Hyperbolic tangent approximation to the sign function. $\tanh(sx)$ for $s = .5, 1, 3$, and 9 .

rather than a global minimum, motivate the search for an alternate formulation which assures that a global optimum will be found. The guarantee of a global optimum is achieved only at the cost of some misspecification of the objective.

Recall the criterion function for the minimum square error classifier

$$(35) \quad J_S(w) = \sum_i (a_i - w'y(i))^2$$

Following the weighting scheme used with the hyperbolic tangent algorithm, insert into the summation of (35) an additional term representing the absolute value of the change in interest rates, $|v_i|$, and define a new criterion function

$$(36) \quad J_W(w) = \sum_i |v_i| (a_i - w'y(i))^2$$

If u_i is set equal to the positive square root of $|v_i|$, then u_i can be taken inside the squared term and distributed, giving

$$(37) \quad J_W(w) = \sum_i (u_i a_i - u_i w'y(i))^2$$

Define a diagonal matrix, U , with nonzero elements, u_i .

The matrix equivalent of (37) is

$$(38) \quad J_W(w) = (Ua - UYw)'(Ua - UYw)$$

The gradient of J_W is

$$(39) \quad \text{grad } J_W = -2Y'U'Ua + 2Y'U'UYw$$

Setting the gradient equal to 0 and solving for w gives

$$\begin{aligned} w &= (Y'U'UY)^{-1}Y'U'Ua \\ (40) \quad &= (YU)^{\#}Ua \end{aligned}$$

This is the weighted least squares algorithm, and is sometimes denoted WLS.

If class C_1 is defined by $v_i > 0$ (i.e., interest rates increase) and class C_2 by $v_i < 0$, then $\text{sgn}(v_i)$ specifies the class of $x(i)$. Let $b = \text{sgn}(v_i)$. With these definitions, an equivalent formulation in terms of the original feature vectors is

$$(41) \quad w = (XU)^{\#}Ub$$

A pattern is assigned to C_1 if $w'x > 0$ or to C_2 if $w'x < 0$.

The device of weighting the individual terms by the amount of the change in interest rates was motivated by the desire to achieve the defined objective. An alternative viewpoint may be taken, namely that those feature vectors which lead to greater changes are somehow more informative (i.e., contain more information, are less subject to error) than those which correspond to smaller interest rate changes.

The solution to (41) is straightforward. All vectors are considered simultaneously and the result is unique since the matrix is almost never singular. But the price in terms of distortion of the objective has been paid.

SELF-ADAPTATION OF THE CLASSIFIERS

Many PR problems concern processes that are inherently stable. This means that if a classifier giving adequate performance can once be developed, it need not later be changed. A reasonable approach under such circumstances is to select a large, representative sample of feature vectors, present them to the classifier, accept the resulting weight vector as definitive, and put the classifier into operation. In PR terminology the first phase of this process (when the weight vector is being determined) is called training--whether it be multi-step, as with the perceptron, or one-step, as with the minimum squared error classifier. The separation of the training phase from the operating phase is a matter of convenience and practicality. There is no conceptual reason that the classifier could not continue to learn and modify its weight vector on the basis of feedback about its operating performance.

Indeed, the author has emphasized the necessity of a classifier being able to use such feedback to continually modify itself when the system or environment is undergoing change. However, to reduce mathematical and conceptual complexity while the algorithms were being derived above, that ability was not stressed. For the same reason the crucial distinction between retrospective and prospective performance was allowed to remain blurred. It is time to

correct the first of these lapses and examine how the classifiers are expected to change themselves as new information becomes available during their operation.

The easiest way to view the adaptation process is to assume that, as each new pattern of known classification becomes available, the classification problem is created de novo and that a new solution must be found without reference to prior results. The sequence of solutions (weight vectors) so determined represents the adaptation of the classifier over time. Depending on the algorithm, it may turn out that that this formulation of the adaptation process is highly inefficient computationally, but for now it keeps the concepts straight.

Consider a sequence of augmented feature vectors, $x(i)$, $i = 1, 2, \dots, m$. Consider also a sequence of problems, P_j , $j = 1, 2, \dots, m$. The problems are related to the vectors as follows: problem P_j consists of finding a weight vector, $w_{<j>}$, which classifies the vectors $x(k)$, for k less than or equal to j . (Conceptually and notationally, $w_{<j>}$ is distinct from $w(k)$ used in conjunction with the perceptron and relaxation algorithms; $w_{<j>}$ is a final solution to problem P_j , $w(k)$ is an interim result occurring only if certain kinds of algorithms are used; $w_{<j>}$ is the result of a major loop or iteration, $w(k)$ is the result of a minor loop or iteration.) Let the

matrix, W , be constructed such that its rows are the weight vectors, $w^{<j>}$, sequentially obtained by solving the problems, P_j . The matrix provides a history of the evolution of the classifier.

Refer back to the example illustrating the perceptron. The matrix W for that case is

$$W = \begin{bmatrix} -1 & -3 & 3 \\ 0 & -2 & 1 \\ 0 & -2 & 1 \\ 0 & -2 & 1 \\ 0 & -2 & 1 \end{bmatrix}$$

because $w^{<1>} = (-1, -3, 3)'$, (not previously computed) correctly classifies $x(1)$ if presented by itself; $w^{<2>} = (0, -2, 1)'$ correctly classifies $x(1)$ and $x(2)$; $w^{<3>} = (0, -2, 1)'$ correctly classifies $x(1)$, $x(2)$ and $x(3)$; $w^{<4>} = (0, -2, 1)'$ correctly classifies $x(1)$, $x(2)$, $x(3)$ and $x(4)$; finally $w^{<5>} = (0, -2, 1)'$ correctly classifies all five vectors.

A special convention is required for the minimum squared error and weighted least squares algorithms since a solution vector is not defined by the algorithm when the number of vectors is less than the number of elements in the augmented feature vector (when, in other words, the number of equations in (23) is less than the number of unknowns to be determined). Arbitrarily let the undetermined rows of W be set equal to $0'$. For the minimum squared error example above, let the first two

rows be 0'. Additional computation shows that the third and fourth rows are, respectively, $(-2, 0, -1)$ and $(2, -1, .5)$. The last row is the solution vector of the problem as initially given, $(-.018, -.432, .027)$. Hence the matrix is

$$W = \begin{bmatrix} .000 & .000 & .000 \\ .000 & .000 & .000 \\ -2.000 & .000 & -1.000 \\ 2.000 & -1.000 & .500 \\ -.018 & -.432 & .027 \end{bmatrix}$$

This adaptation is illustrated in Figures 7, 8 and 9.

In problems with a large number of feature vectors, examination of the matrix W provides worthwhile insight into the evolution of the coefficients of the classifier. It is also sometimes desirable to standardize the rows of W so that the length of all the vectors is 1.

The computational burdens imposed by starting over from the beginning each time a new observation becomes available can be severe. Naturally, the efficient approach is to use the final solution vector of the previous major iteration as the starting point for the new problem. This is a trivial matter for algorithms like the perceptron where a recursive formulation is already being used. It is not immediately obvious how to proceed with either minimum squared error or weighted least squares algorithms. Fortunately, recursive schemes have been devised so that the repeated matrix inversions seemingly called for by (29) and (41) are unnecessary. Adapting Harvey (1981) to

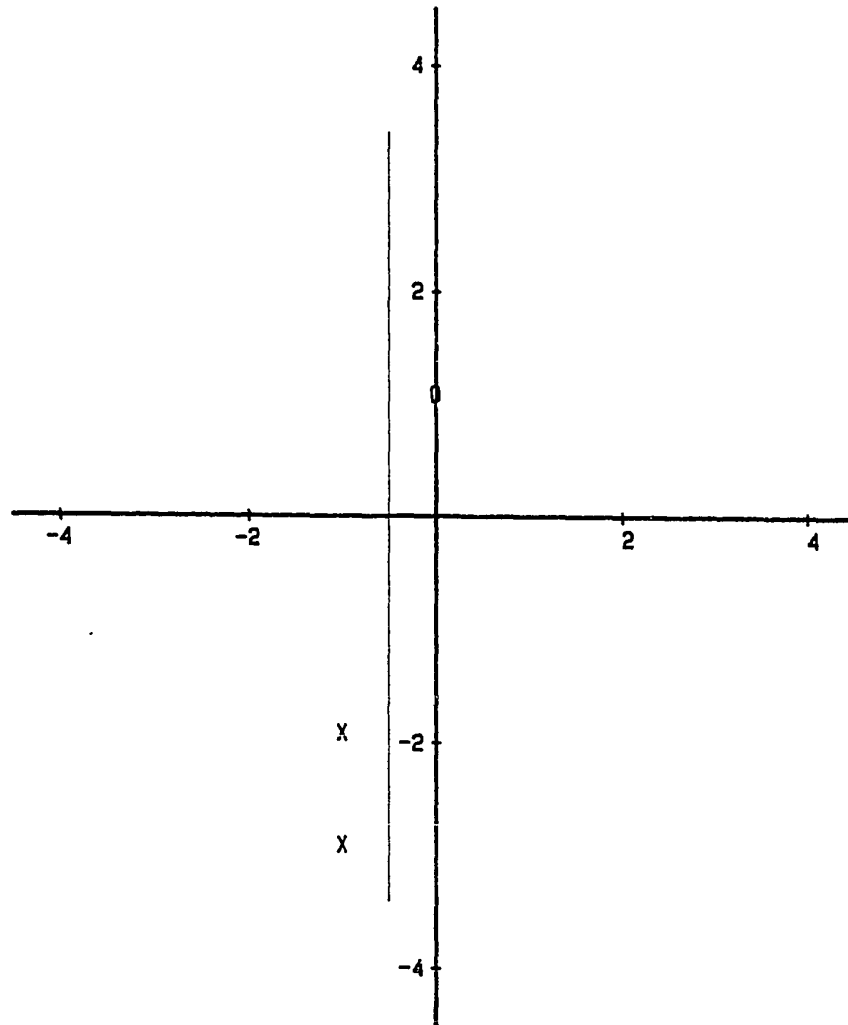


Figure 7. Partition boundary for minimum squared error algorithm
-- first three points.

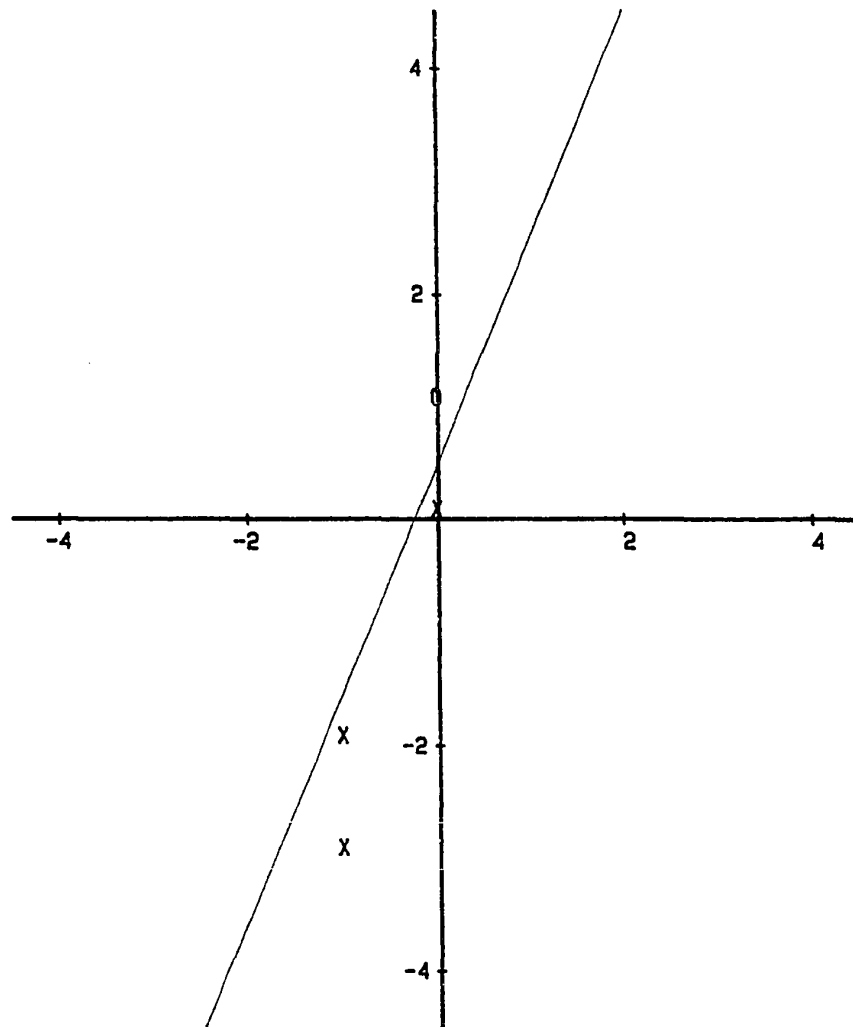


Figure 8. Partition boundary for minimum squared error algorithm
-- first four points.

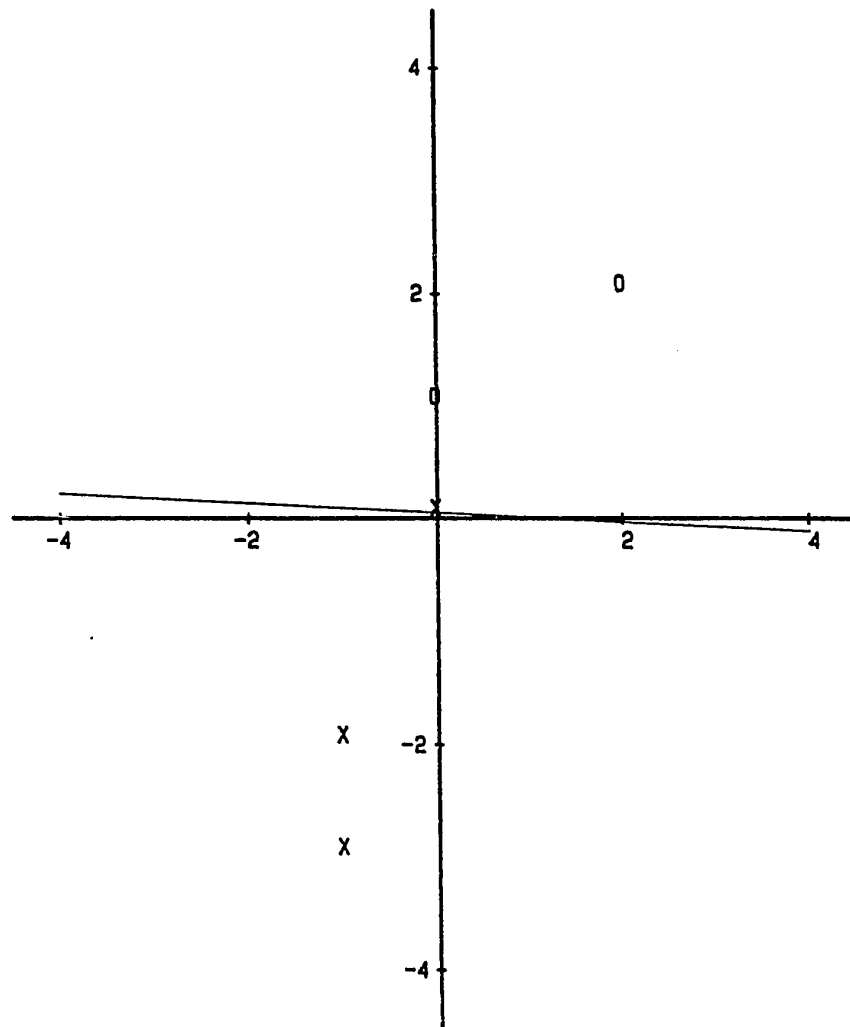


Figure 9. Partition boundary for minimum squared error algorithm
— all five points.

the notation used here, let

$$(42) \quad w^{<j-1>} = (X'^{<j-1>} X^{<j-1>})^{-1} X'^{<j-1>} b^{<j-1>} \\ = X^{\#<j-1>} b^{<j-1>}$$

represent the solution vector obtained from the minimum squared error algorithm using only the first $j-1$ feature vectors. Now assume the j th feature vector and its classification become available. Then the updated solution vector $w^{<j>}$ is

$$(43) \quad w^{<j>} = w^{<j-1>} + \\ (X'^{<j-1>} X^{<j-1>})^{-1} x(j) (b_j - w'^{<j-1>} x(j)) / f(j)$$

where

$$(44) \quad f(j) = 1 + x'(j) (X'^{<j-1>} X^{<j-1>})^{-1} x(j)$$

The inverted cross-product matrix needed for the next cycle is

$$(45) \quad (X'^{<j>} X^{<j>})^{-1} = (X'^{<j-1>} X^{<j-1>})^{-1} \\ - (X'^{<j-1>} X^{<j-1>})^{-1} x(j) x'(j) (X'^{<j-1>} X^{<j-1>})^{-1} / f(j)$$

While notationally formidable, this method is actually not difficult to implement computationally and it saves a large amount of computer time. Harvey also shows how to compute w without any direct matrix inversion.

Extension of these procedures to the weighted least squares algorithm is straightforward.

In operation, the recursively updated weighted least squares algorithm is used to obtain reasonable estimates for the weight vector. These estimates can then be refined using the hyperbolic tangent algorithm. The former has an

analytic solution and is computationally efficient; the latter requires the use of numeric search techniques that are slow and may not converge. For these reasons the recursively updated weighted least squares algorithm is the one used here. An example of the use of the hyperbolic tangent algorithm to improve the WLS classifier is given in Appendix C.

SUMMARY

This chapter shows how, starting from basic principles, it is possible to derive mathematical algorithms that will act as pattern classifiers. Three traditional classifiers were derived: the perceptron, the relaxation algorithm, and the least mean squared error classifier. Then two additional algorithms were developed to deal with the specific problem at hand. The recursively updated weighted-least-squared-error algorithm is a combination of known mathematical techniques here applied to the classification problem. The hyperbolic tangent algorithm was developed by the author following techniques suggested by Highleyman (1962).

CHAPTER IV

INTEREST RATE FORECASTING

This chapter reports the substance of the research.

PRELIMINARY MATTERS

What Is a Forecast?

There is a proverb, variously attributed, as follows:

To prophesy is difficult -- especially about the future.

It may be inscrutable Chinese wisdom; more likely it is wry commentary on what passes, these days, as economic forecasting; certainly it is more tactful than the alternative:

To prophesy is easy -- especially about the past. Much of the literature on interest rate forecasting concerns prophecy about the past.

Any article on interest rates appearing in an economic journal is likely to exhibit equations of the form

$$(46) \quad R_t = a_0 + a_1X_t + a_2Y_t + a_3Z_t$$

where R_t is the interest rate at time t and X_t , Y_t and Z_t are the values of explanatory variables contemporaneous

with the interest rate, i.e., at time t . If the goal of research is the understanding, explanation, or verification of the economic factors that determine interest rates, this formulation may be appropriate; see, for example, Hamburger and Silber (1969) and Feldstein and Eckstein (1970).

Rewrite equation (46) with different subscripts

$$(47) \quad R_{t+h} = a_0 + a_1 X_{t+h} + a_2 Y_{t+h} + a_3 Z_{t+h}$$

where R_{t+h} is the forecast (made at time t) of what the interest rate will be h periods ahead (at time $t+h$) and X_{t+h} , Y_{t+h} and Z_{t+h} are estimates (made at time t) of what the explanatory variables will be at time $t+h$. Obviously the forecaster proposing to use equation (47) is required to provide these estimates. It may be that the equation is part of a multi-equation structural model which itself generates the needed estimates (Hunt, 1976); if so, the formulation is not to be faulted. Of course, the quality of the forecast then depends as much on the quality of the estimates as on the validity of the equation.

As a single equation for forecasting, however, (47) simply will not do. Unfortunately many articles on interest rate forecasting use just this form with no hint whatsoever as to how to obtain estimates of the explanatory variables used in the equation. When the actual values of X , Y and Z become available, they are duly entered into the equation and, with straight face, the forecaster pronounces his "forecast" a success.

A possibly proper form for a single forecasting equation is

$$(48) \quad R_{t+h} = a_0 + a_1 X_t + a_2 Y_t + a_3 Z_t$$

where R_{t+h} is the forecast (made at time t) of what the interest rate will be at time $t+h$, and X_t , Y_t and Z_t are the values of relevant variables at time t .

This discussion suggests the concept of an information set, a term used informally earlier. An information set at time t , denoted I_t , consists of all the data, knowledge, theories and assumptions available as of time t (Granger and Newbold, 1977). Economic data from the future cannot really be included in the information set, I_t ; in forecasting fantasyland it often is.

Frequently, economic time series relating to time t may not actually be compiled and published until some time thereafter. In such cases the forecasting formula must incorporate the appropriate lags so that the proper equation becomes

$$(49) \quad R_{t+h} = a_0 + a_1 X_{t-k} + a_2 Y_{t-k} + a_3 Z_{t-k}$$

where the subscript $t-k$ refers to data relating to time $t-k$ which does not become available until time t .

Sometimes future data sneak into the information set through the back door. If, for example, a data series is standardized to zero mean and unit variance, the value at $t+h$ influences the transformed value at time t . Similarly, if a series is detrended, there is a danger that the

future will be permitted to influence the past.

The models developed here produce real forecasts. They are not contingent or conditional in any sense. In particular, they do not require the use of an information set that does not exist when the forecast is being made. There is no retrospective forecasting here, only prospective forecasting.

What Rate to Forecast?

The interest rate variable considered in this study is a yield rate on long-term U.S. government securities, in particular, the interest rate on twenty-year, constant-maturity U.S. government bonds. This series is computed daily by the U.S. Treasury and published by the Federal Reserve Board in statistical releases H.15 and G.13 (Board of Governors of the Federal Reserve System, various dates). The rate is shown, for the years 1967 through 1982, in Figure 10.

"Information About the Treasury Constant Maturity Yield Series" (U.S. Treasury, no date) explains the process by which the series is computed. First the Treasury obtains quotations on all actively traded government issues as reported to the Federal Reserve Bank of New York by five U.S. government securities dealers. The bid price quotes are converted to yield rates and plotted on a graph. The horizontal axis shows the maturity date of each security and the vertical axis measures the

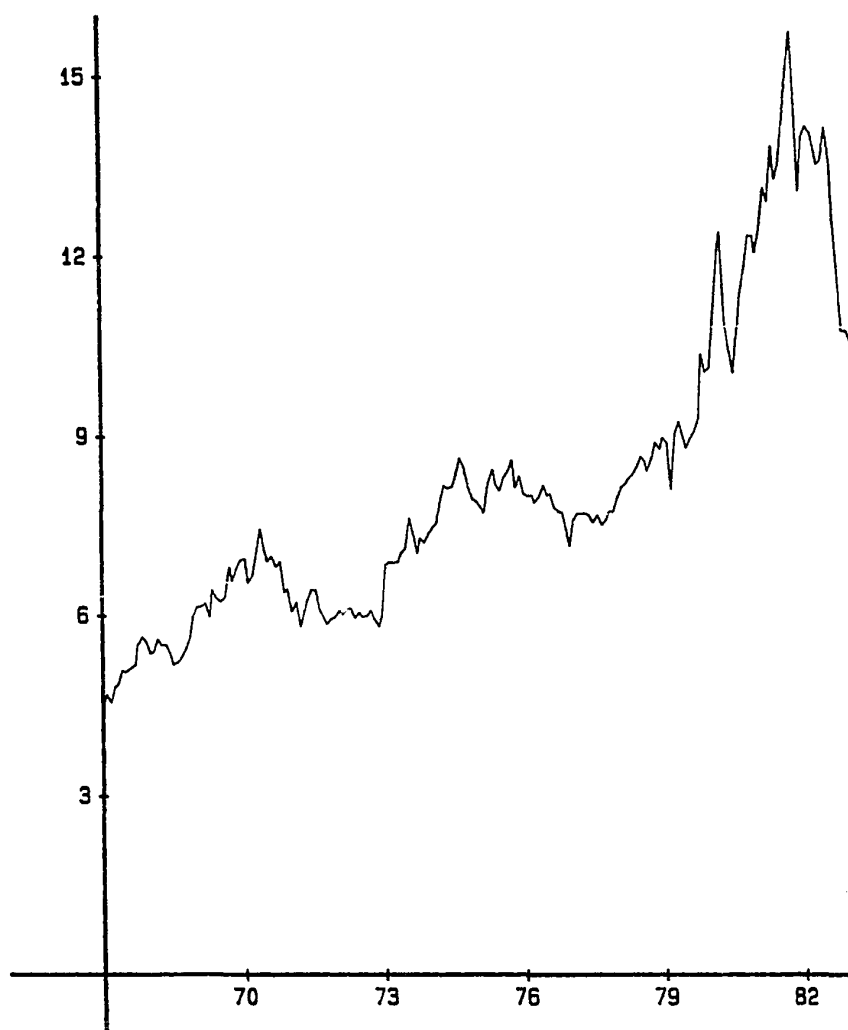


Figure 10. Interest rate on 20-year constant-maturity U.S. government bonds, 1967-82.

yield rate. A yield curve (see Figure 11) is derived with the greatest weight being given to the more recently issued, most actively traded securities. In particular, so-called flower bonds which have low coupons, which sell at deep discounts, and which possess inheritance tax advantages receive very little weight. From the curve it is possible to obtain the yield for a twenty-year bond even if no actual issue with that exact maturity is being traded.

Because of legislatively imposed interest rate ceilings, no Treasury bonds with maturities exceeding seven years were sold between June 1965 and August 1971, and no bonds with maturities exceeding fifteen years were issued between June 1965 and December 1972. As a result of increasing coupon rates and higher yields over that period, the twenty-year series increasingly understated the "true" level of rates. When a new twenty-year bond was again auctioned in January 1973, a discontinuity of about 75 basis points (.75%) appeared and is evident in Figure 10. This aberration has been removed by the simple expedient of increasing all rates prior to January 1973 by 75 basis points. The resulting distortion in the change in rates for any month probably does not exceed one or two basis points and is properly ignored.

The series described above is artificial, in

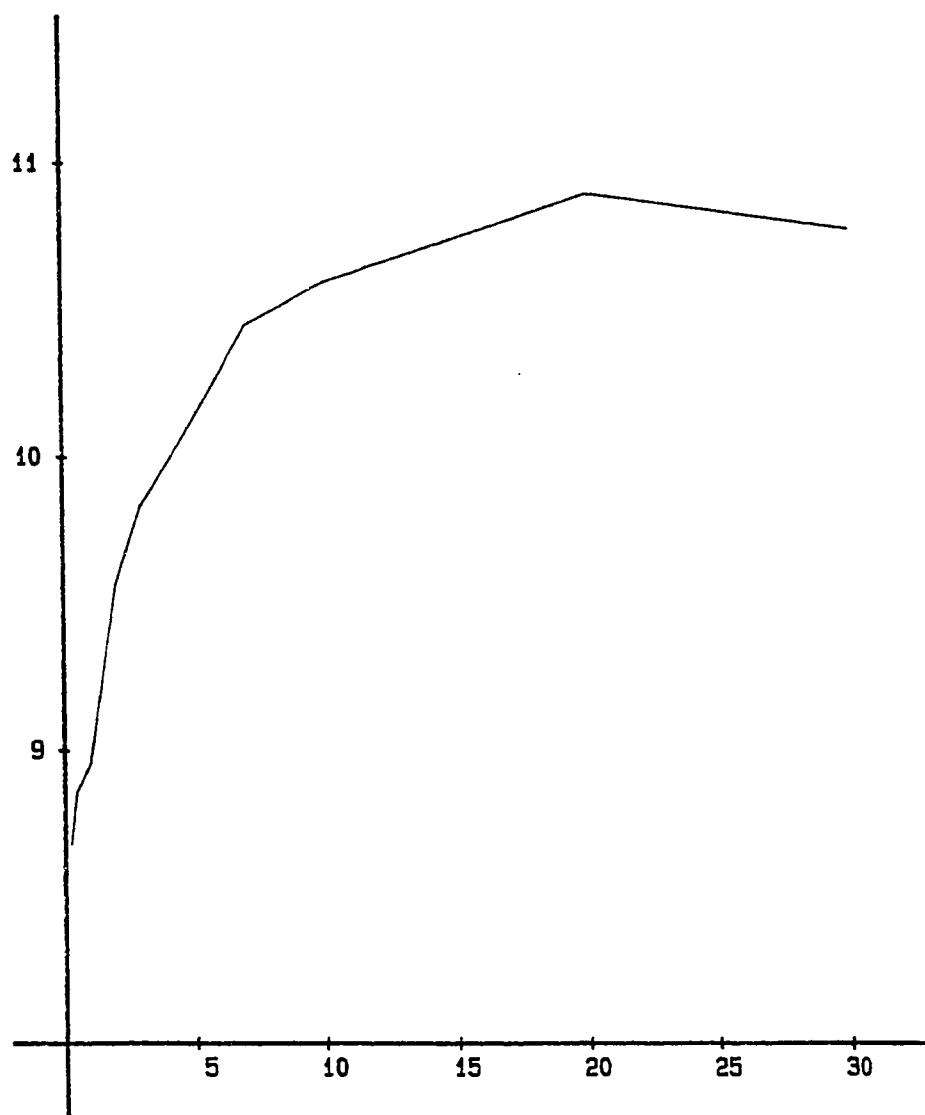


Figure 11. The yield curve shows interest rates as a function of time-to-maturity.

the sense that it is not the rate on any actual bond, and consequently it does not represent an actual security that could be bought or sold. It does, however, have many favorable attributes. It is closely representative of many government bonds at the long end of the maturity spectrum and thus is less subject to random fluctuations influencing any particular bond. By its nature, it eliminates time-to-maturity as a variable affecting rate changes. Moreover, the use of government securities eliminates variation in rates due to changes in the credit quality of the issuer.

Once a series has been chosen, there remain a number of other important practical questions to be answered. Over what time period is the forecast to be made: a day, week, month, quarter or year? A one-month forecast horizon is used here. It coincides with the reporting interval of a very large number of economic time series and allows a reasonable number of data points for statistical analysis.

While many users of an interest rate forecast would prefer an accurate point estimate, the direction of change in interest rates between month-ends is forecast here. Knowledge of the direction of change would serve just as well in many practical situations (Fraser, 1977; McCracken, 1976). For example, the following decisions could be made on the basis of a binary forecast:

Whether or not to hedge a fixed-income portfolio.

Whether to buy long-term bonds or invest in short-term cash-equivalents.

Whether to issue debentures currently or borrow from a bank and await possibly lower rates.

An important practical reason for forecasting the direction of change is that it is easier than forecasting the amount of change. This follows from a formal result in information theory and cybernetics, called the law of requisite variety (Ashby, 1964; Hare, 1967). The law is of wide applicability, but interpreted for the context at hand it says that complexity of the forecasting system can be reduced if the range of outcomes to be forecast is smaller.

The decision to forecast the difference between month-end values (rather than changes in monthly averages) is motivated by statistical considerations. Changes in the monthly average of a series will be correlated even if the individual daily values are uncorrelated (Working, 1960). The existence of correlation in changes of monthly averages can lead the unsuspecting to assert the existence of patterns in data that are really spurious; and even if recognized, it makes evaluation of the forecasting results much more difficult.

To summarize: the objective of this study is to forecast the direction of change between month-end values

of the interest rate on twenty-year, constant-maturity U.S. government bonds.

THE OBJECTIVE FUNCTION

One way to approach the interest rate forecasting problem is to search for a set of variables, x_i , (which have been called features) and a set of weights, w_i , such that, when a new feature vector is presented, the weighted sum

$$(50) \quad d(x) = w_1x_1 + w_2x_2 + \dots + w_nx_n + w_{n+1} = w'x$$

is always greater than zero when interest rates increase and is always less than zero when rates decrease.

Unfortunately such a perfect classifier is not to be found. Given the inevitability of errors, what lesser goal should be sought? It was argued above that under reasonable conditions the objective should be to forecast so as to maximize the sum of the gains less the sum of the losses that would be incurred by following the forecasts or, equivalently, to minimize the sum of the losses. A criterion function that expresses this latter objective is

$$(51) \quad J_L(w) = .5 \sum_i |v_i| |\text{sgn}(v_i) - \text{sgn}(w'x(i))|$$

where v_i is the change in the interest rate, $|v_i|$ is its absolute value and the sign function, $\text{sgn}(z)$, is defined as

$$(52) \quad \begin{aligned} \text{sgn}(z) &= 1 && \text{if } z > 0 \\ &= -1 && \text{if } z \leq 0 \end{aligned}$$

If $\text{sgn}(v_i) = \text{sgn}(w'x(i))$, then the forecast of the direction of change is correct and the contribution to the criterion function is zero. If the signs are not equal, the forecast is in error and the contribution to J_L is $|v_i|$, the amount lost by following the forecast. Thus minimizing J_L minimizes the sum of the losses.

Now a crucial temporal distinction: the forecaster wishes to minimize J_L for future new pattern vectors; his hope is that the set of weights that would have minimized J_L for past patterns will continue to give good performance on new patterns. Thus the criterion function serves two purposes--a statement of the future forecasting objective and a mathematical expression to be minimized in computing the weight vectors from past data.

Finding a set of weights that minimizes J_L is a difficult analytical and computational problem. Consequently J_L is replaced by a mathematically more tractable criterion function. The hope and reasonable expectation is that the solution to the new problem will not differ greatly from the solution of the old (or, more importantly, that the weight vector established by minimizing the new criterion function on past data will perform adequately in minimizing the old criterion function for new patterns). The new criterion function is

$$(53) \quad J_W(w) = \sum_i |v_i| (b_i - w'x(i))^2$$

where $b_i = \text{sgn}(v_i)$. This is a problem of weighted least

squares which has the solution

$$\begin{aligned} \mathbf{w} &= (\mathbf{X}'\mathbf{U}'\mathbf{U}\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}'\mathbf{U}\mathbf{b} \\ (54) \qquad &= (\mathbf{X}\mathbf{U})^\# \mathbf{U}\mathbf{b} \end{aligned}$$

where \mathbf{U} is a diagonal matrix with nonzero elements equal to the positive square root of $|v_i|$.

To make the forecasting system adaptive and take advantage of new information as it becomes available, \mathbf{w} is recomputed when the true outcome for the most recent previous pattern becomes known. In other words, each month the feature vector is assembled and combined with the most recently computed weight vector to make a forecast using (50). When the actual result becomes known at the end of the month, the weight vector is revised using (54) and the process continued. This recursive weighted least squares algorithm will be used as the forecasting method in all of what follows. As a further refinement, the gradient of (34) can be used as the basis of a numerical search for a weight vector that achieves an even lower score on the objective function J_L . An example is given in Appendix C.

EVALUATION STANDARDS

The most obvious and crucial point about evaluating a forecasting model is that it should never be judged on the basis of the data used to develop it. In the pattern recognition field this mistake is often called testing on

the training data. Consequently, the models developed here are to be judged not on the basis of how well they were able to classify the training patterns used to estimate the weight vectors, but only on how well the estimated weight vectors were able to classify patterns subsequently presented to the classifier. The former information is not relevant to the forecasting task.

It is easy to define the perfect forecast; it always equals the actual result. Here that requires that J_L equal zero. In a less than fully predictable world some other measures are needed.

When the half dozen or so classical statistical assumptions underlying ordinary least squares regression are satisfied, there are a number of statistics with well known properties that can guide the forecaster. But these standard measures are not appropriate here because the underlying assumptions are invalid. The dependent variable is dichotomous, not continuous; the objective function is not the usual unweighted quadratic form; and the regression is run recursively. These considerations lead to the development of evaluation measures that appeal to common sense:

1. Net profit. If the direction of change in interest rates is forecast correctly for any month, call the amount of change a profit. Otherwise, call it a loss.

Let the sum of the profits less the sum of the losses over the entire fourteen-year forecasting period (1969-1982) be called a net profit (or a net loss, if negative)

Profits and losses are measured in basis points (one basis point equals .01%).

Profits and losses are expressed in basis points merely as a matter of convenience.

The "10-for-1" rule can be used to convert basis point profits (losses) into dollar profits (losses). This rule says that a one percent increase in the interest rate reduces the present value of a long series of future cash flows by about ten percent. Thus a one basis point change in the yield to maturity changes the price of a \$1000 bond by about \$1. Consequently, on a \$1 million bond portfolio a 500 basis point net profit translates into about \$500,000. The actual dollar value corresponding to a one basis point change depends on the coupon rate and yield to maturity.

2. Profit/loss ratio. Call the ratio of the sum the profits to the sum of the losses the profit/loss ratio.
3. Number of hits. If the direction of change

is forecast correctly call the forecast a hit. Otherwise, it is a miss. Months for which the interest rate did not change (there were three such months during the forecasting period) are arbitrarily counted as misses.

4. Rate of return. Another statistic, especially useful in testing the efficient market theory, is the rate of return earned by assets invested in accordance with a specified strategy. The rate of return for any month depends on how funds are assumed to have been invested during that month. Two alternatives are assumed to be available: to invest in long-term government bonds or to invest in federal funds, the choice determined by whether rates are forecast to increase or decrease. If rates are forecast to increase, funds are assumed to be invested in federal funds and the rate of return is the monthly average rate on federal funds. If rates are forecast to decrease, funds are assumed to be invested in long-term bonds and the return is composed of two pieces: interest earned on the bond during the month; and the capital gain or loss incurred because of changes in the bond's price. All rates of return are

expressed here as monthly effective rates,
i.e., the rates have not been annualized.

5. Worst computed loss. The worst computed loss is the answer to the following question: Suppose the model had been implemented just before and discontinued just after its worst interim performance during the fourteen-year period (judged on a net profit basis and measured only at month-ends), how much would the net loss have been over that period? This figure is important because it measures the risk incurred in following the forecasts of a given model.

All of these are absolute indicators; and others of the same type, such as the probability of error, might be derived. Relative or comparative measures are also desirable. When there are no published alternative forecasts to measure against, how can one tell whether a forecast is "good" or "bad" compared to others? The usual approach is to set up a straw man, a naive forecast which seems obvious, and then try to improve upon it. Two reasonable naive standards are derived below.

A Random Rule.

One way to assess the forecasting power of the models developed here is to compare their performance with what might have been achieved if the random walk hypothesis were

true. One way to implement a random walk is to suppose the following rule: Flip an unbiased coin. If the result is heads interest rates are forecast to increase; if the result is tails interest rates are forecast to decrease. Implementing the rule 168 times (once for each month in the fourteen-year period, 1969-82) would provide one sample of the action of the rule. By repeating the whole process many times, a distribution of results can be obtained. This method of simulated sampling is sometimes called Monte Carlo simulation. Obviously, a good forecasting rule should outperform most of the random walks.

A computerized analogue of the coin-flipping rule was developed to simulate a random walk. The model was run through 10,000 trials. Each trial consists of the following steps:

A random number generator produces a series of 168 digits from the set (+1,-1) corresponding to months in the fourteen-year period.

If the digit is +1, the random decision function forecasts that rates will increase during that month.

If the digit is -1, the opposite forecast is made.

The net profit, the profit/loss ratio, the number of hits, the rate of return, and the worst computed loss are calculated and

retained.

These sets of values form five vectors, each with 10,000 elements. The vectors can be sorted and various statistics computed.

S-shaped cumulative frequency functions for the net profit, the profit/loss ratio, the number of hits, the rate of return, and the worst computed loss as generated by the simulation are shown in Figures 12 through 16. The bell-shaped curves superimposed on these graphs are the corresponding relative frequency functions, which have been rescaled in the vertical dimension to improve visual presentation.

Entries in Tables I through V correspond to the values of the cumulative relative frequency functions of Figures 12 through 16 (or equivalently, to the areas under the relative frequency functions). It is anticipated that the net profit, the number of hits, and the rate of return follow a normal distribution; thus Tables I and III also provide expected frequencies under the normality assumption. Table VI provides selected rank data for each of the five variables.

A Bayesian Rule.

A second standard against which the models might be measured is the result that would have been achieved by

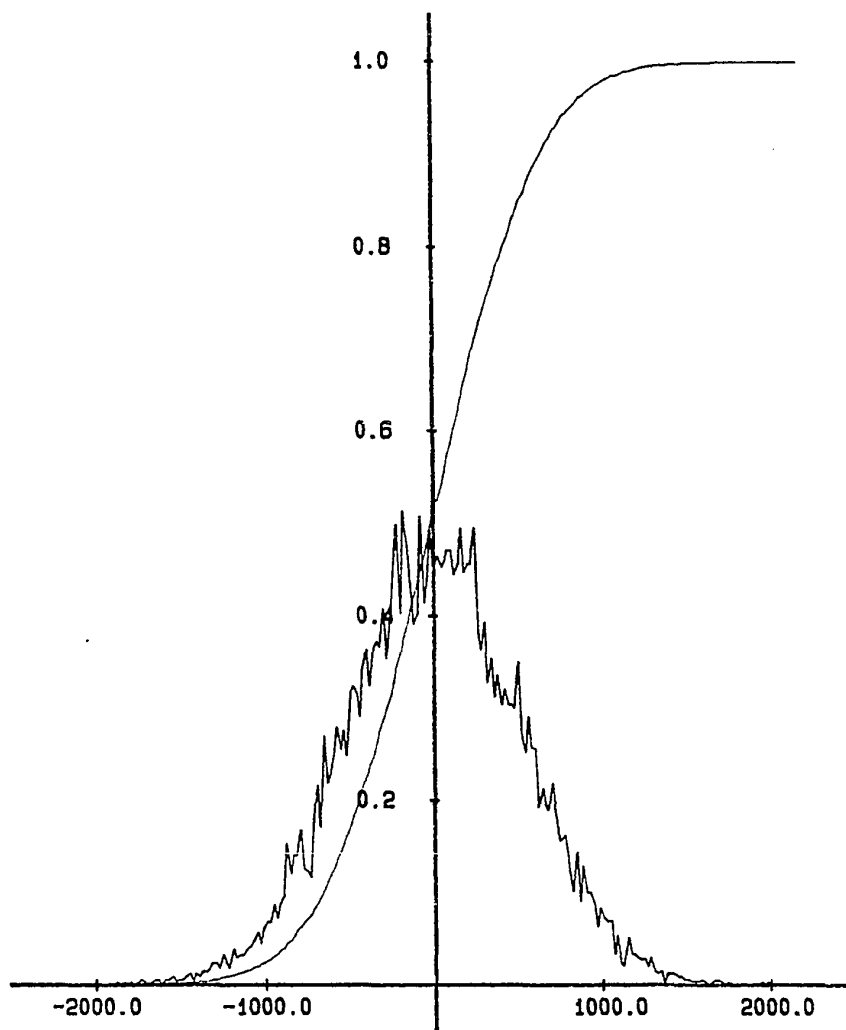


Figure 12. Relative frequency (rescaled) and cumulative relative frequency for the net profit in the Monte Carlo simulation.

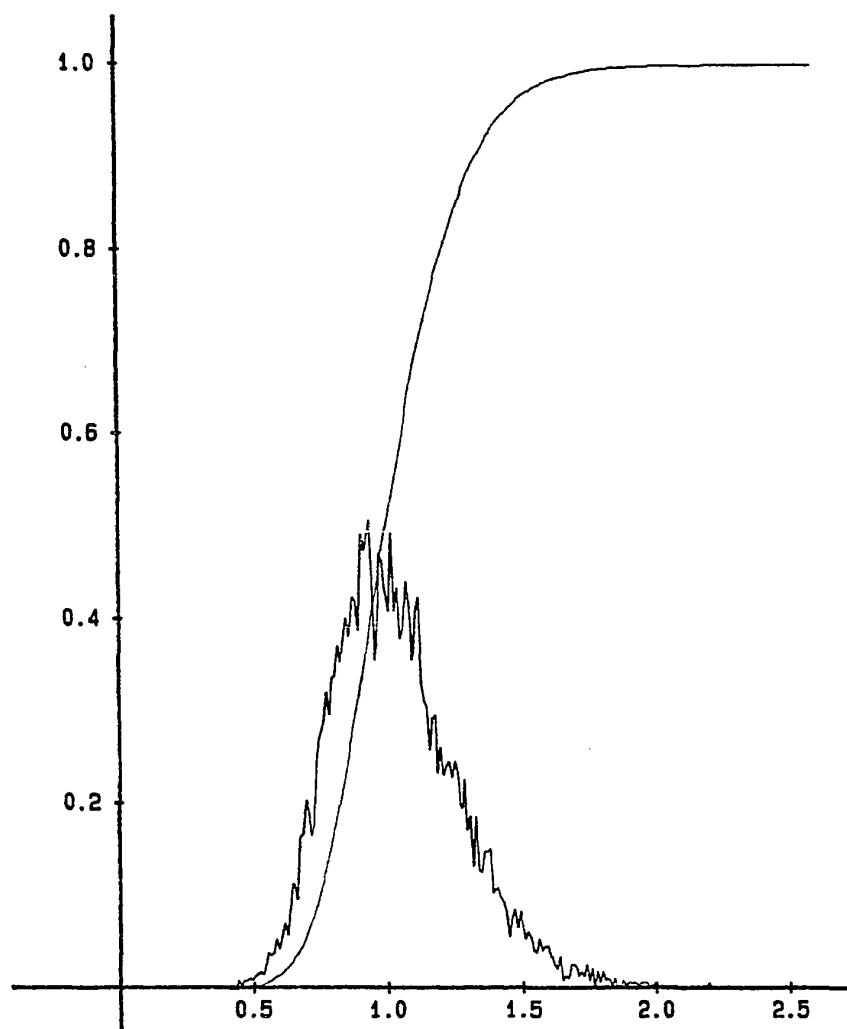


Figure 13. Relative frequency (rescaled) and cumulative relative frequency for the profit/loss ratio in the Monte Carlo simulation.

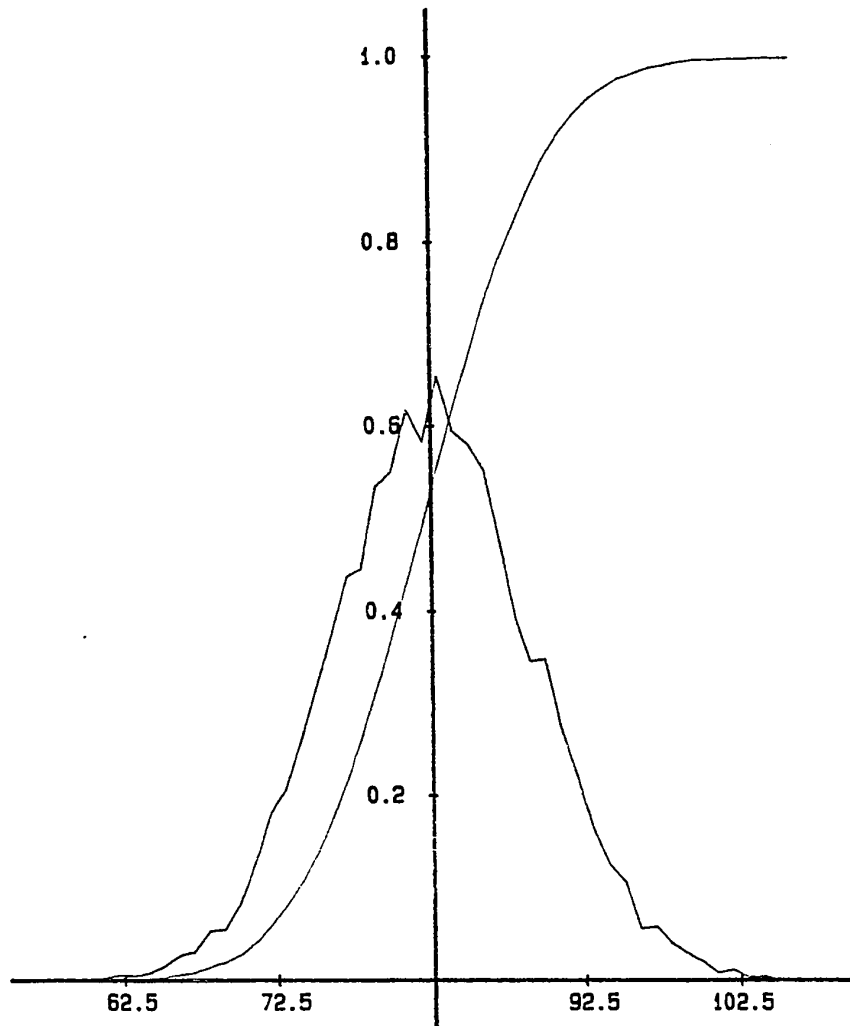


Figure 14. Relative frequency (rescaled) and cumulative relative frequency for the number of hits in the Monte Carlo simulation.

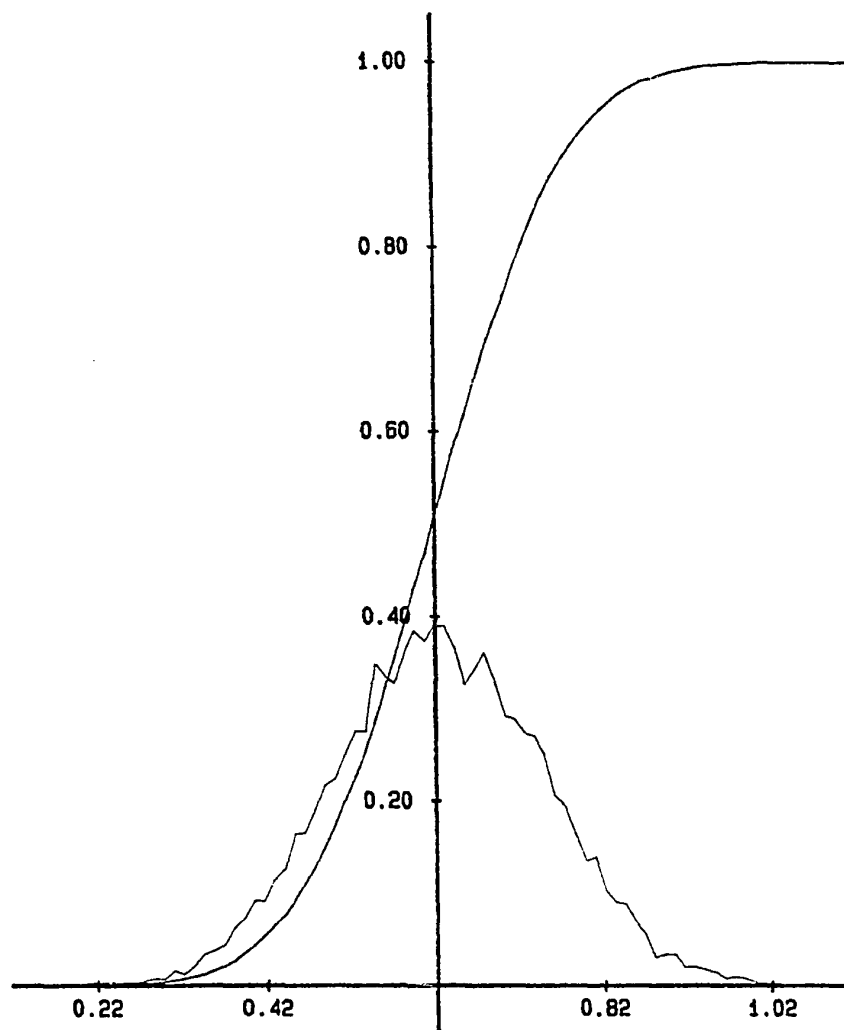


Figure 15. Relative frequency (rescaled) and cumulative relative frequency for the rate of return in the Monte Carlo simulation.

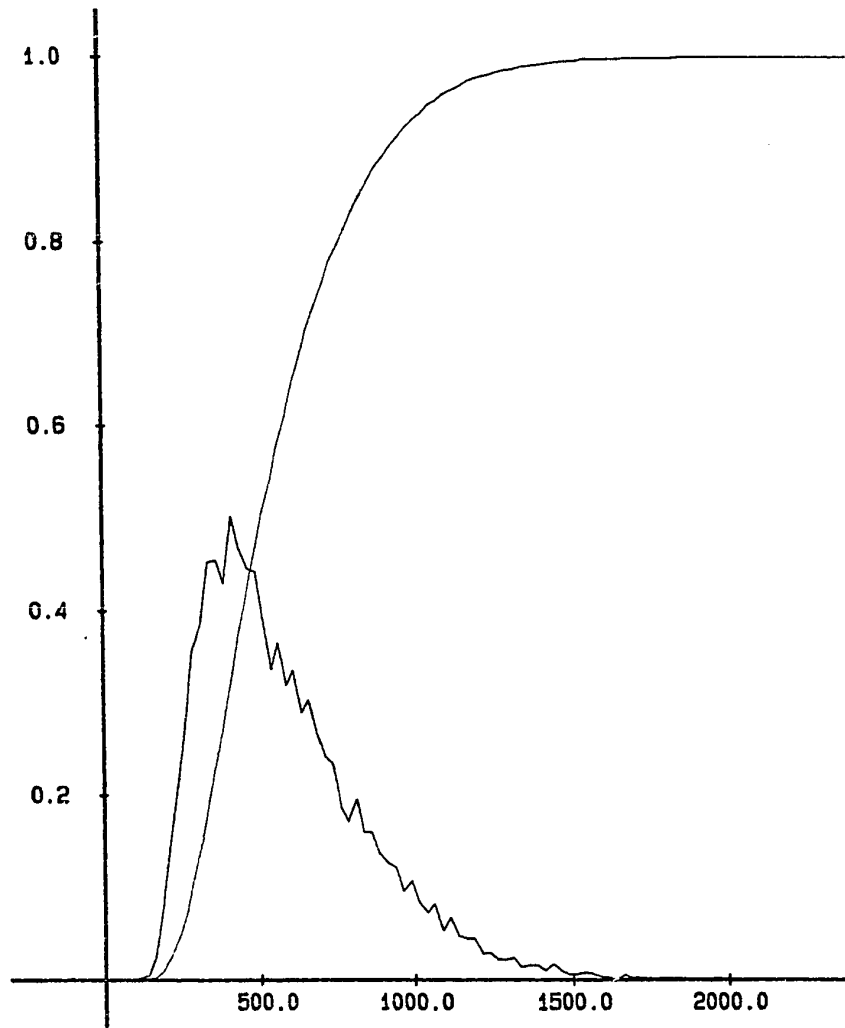


Figure 16. Relative frequency (rescaled) and cumulative relative frequency for the worst computed loss in the Monte Carlo simulation

TABLE I

FREQUENCY DATA FROM A 10,000-TRIAL
MONTE CARLO SIMULATION
NET PROFIT

Net Profit $p \geq$	Observed Frequency	Expected Frequency
-2000	10000	10000
-1600	9988	9992
-1200	9898	9908
-800	9411	9418
-400	7841	7838
0	5044	5000
400	2215	2162
800	588	582
1200	96	93
1600	10	8
2000	0	0
Mean	3.8514	0.0000
Std. Dev.	512.76	509.46
Minimum	-1861.00	
Maximum	1985.00	

TABLE II

FREQUENCY DATA FROM A 10,000-TRIAL
MONTE CARLO SIMULATION
PROFIT/LOSS RATIO

Profit/Loss Ratio $plr \geq$	Observed Frequency
.4	10000
.6	9863
.8	8397
1.0	5044
1.2	2145
1.4	700
1.6	191
1.8	52
2.0	15
2.2	3
2.4	2
2.6	0
Minimum	.419
Maximum	2.551

TABLE III
 FREQUENCY DATA FROM A 10,000-TRIAL
 MONTE CARLO SIMULATION
 NUMBER OF HITS

Number of Hits $h \geq$	Observed Frequency	Expected Frequency
57.5	9999	10000
62.5	9991	9991
67.5	9905	9903
72.5	9399	9403
77.5	7797	7819
82.5	5065	5000
87.5	2209	2182
92.5	626	598
97.5	115	98
102.5	8	10
107.5	0	1
Mean	82.57	82.50
Std. Dev.	6.48	6.42
Minimum	55.00	
Maximum	106.00	

TABLE IV
 FREQUENCY DATA FROM A 10,000-TRIAL
 MONTE CARLO SIMULATION
 RATE OF RETURN

Rate of Return $rr \geq$	Observed Frequency
.2	10000
.3	9948
.4	9629
.5	8320
.6	5712
.7	2644
.8	754
.9	109
1.0	17
1.1	0
Mean	.621
Std. Dev.	.125
Minimum	.217
Maximum	1.074

TABLE V
 FREQUENCY DATA FROM A 10,000-TRIAL
 MONTE CARLO SIMULATION
 WORST COMPUTED LOSS

Worst Computed Loss $wcl \geq$	Observed Frequency
100	10000
300	8913
500	5326
700	2714
900	1215
1100	467
1300	159
1500	47
1700	14
1900	4
2100	1
2300	0

TABLE VI
 RANK DATA FROM A 10,000-TRIAL
 MONTE CARLO SIMULATION

Rank	Net Profit	Profit/Loss Ratio	Number of Hits	Rate of Return	Worst Computed Loss
1	-1861	.419	55	.217	111
10	-1645	.478	63	.255	162
100	-1207	.580	68	.330	198
250	-995	.641	70	.378	224
500	-835	.690	72	.416	253
1000	-651	.749	74	.462	294
2500	-341	.860	78	.535	380
5000	7	1.003	83	.621	521
7500	351	1.167	87	.705	720
9000	663	1.342	91	.782	944
9500	849	1.459	93	.826	1082
9750	997	1.563	95	.863	1212
9900	1193	1.713	98	.905	1386
9990	1585	2.102	102	1.020	1784
10000	1985	2.551	106	1.074	2126

following a simple Bayesian rule. Consider the following problem.

A coin is flipped and one is required to bet \$1 on whether the outcome will be heads or tails.

A correct guess yields a payoff of \$2, an incorrect guess yields nothing. How should one bet?

A Bayesian forecaster is likely to proceed as follows. Initially he will assume equal probabilities for heads and tails either on the basis of his prior experience with coins or on the basis of an assumption of "noninformative priors." Then he will maintain a frequency count of the results and always bet in favor of the outcome having shown the greater frequency.

A similar strategy might be followed in forecasting the direction of change in interest rates: maintain a count of the number of increases and decreases in rates and forecast that they will move in the direction of the more frequent change. For the case at hand, the number of interest rate increases always exceeded the number of decreases (arc sine law?!), so this rule always called for a forecast that interest rates would increase. The overall results of the Bayesian rule, evaluated using the measures previously discussed, are:

Net Profit:	389 basis points
Profit/Loss Ratio:	1.19
Number of Hits:	93

Rate of Return: .690 percent

Worst Computed Loss: 516 basis points

Figures 17 and 18, respectively, show the net profit and profit/loss ratio as they developed over time and Table VII gives the month-by-month profit or loss in basis points.

Although the Bayesian rule is extremely simple, one should not assume that its performance is easily bettered. After all, the effective investment strategy of many financial intermediaries has been to buy bonds and hold them to maturity; this strategy, to a close approximation, is the opposite of what the profitable Bayesian rule would have called for.

DATA SOURCES

Historical data for the interest rate series were obtained directly from the Board of Governors of the Federal Reserve System. Weekly and monthly updates are provided in the Board's statistical releases, H.15 and G.13.

All other series were obtained from Citibase, a machine readable economic database maintained by the Economics Department of Citibank (1983). Citibase contains about 5000 monthly, quarterly, and annual time series, covering the period after 1945. The data are national aggregates and are obtained from over a hundred

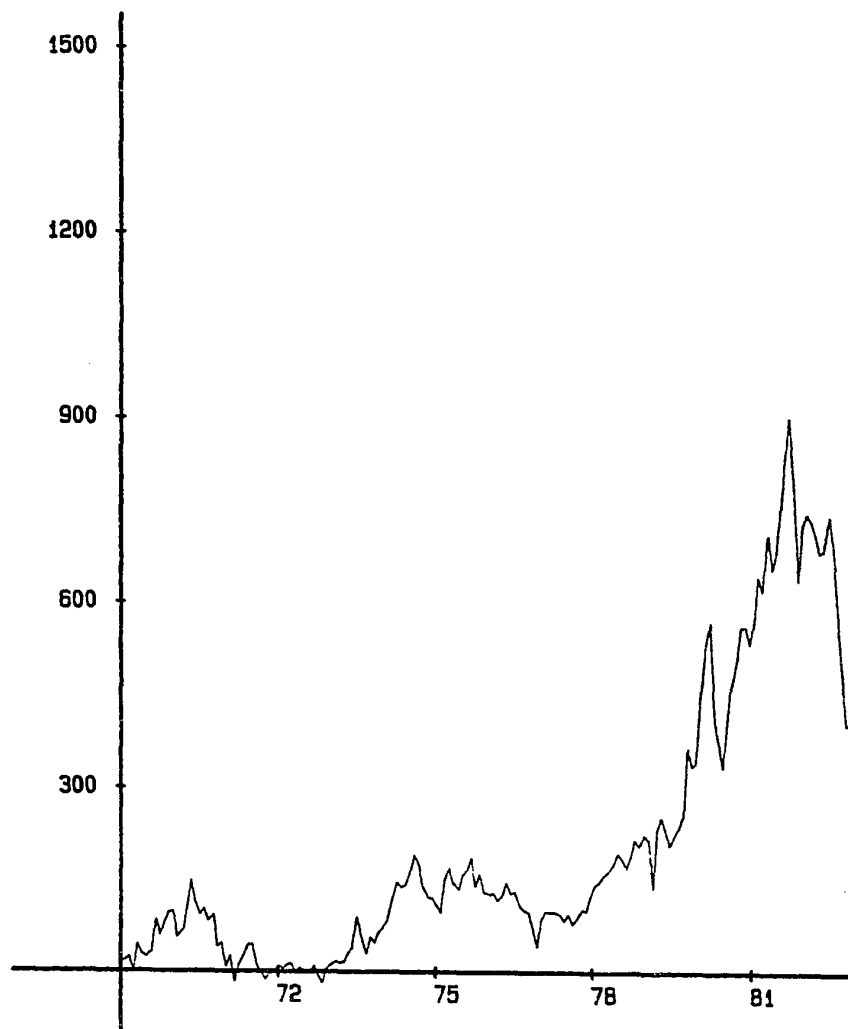


Figure 17. Cumulative net profit -- Bayesian model.

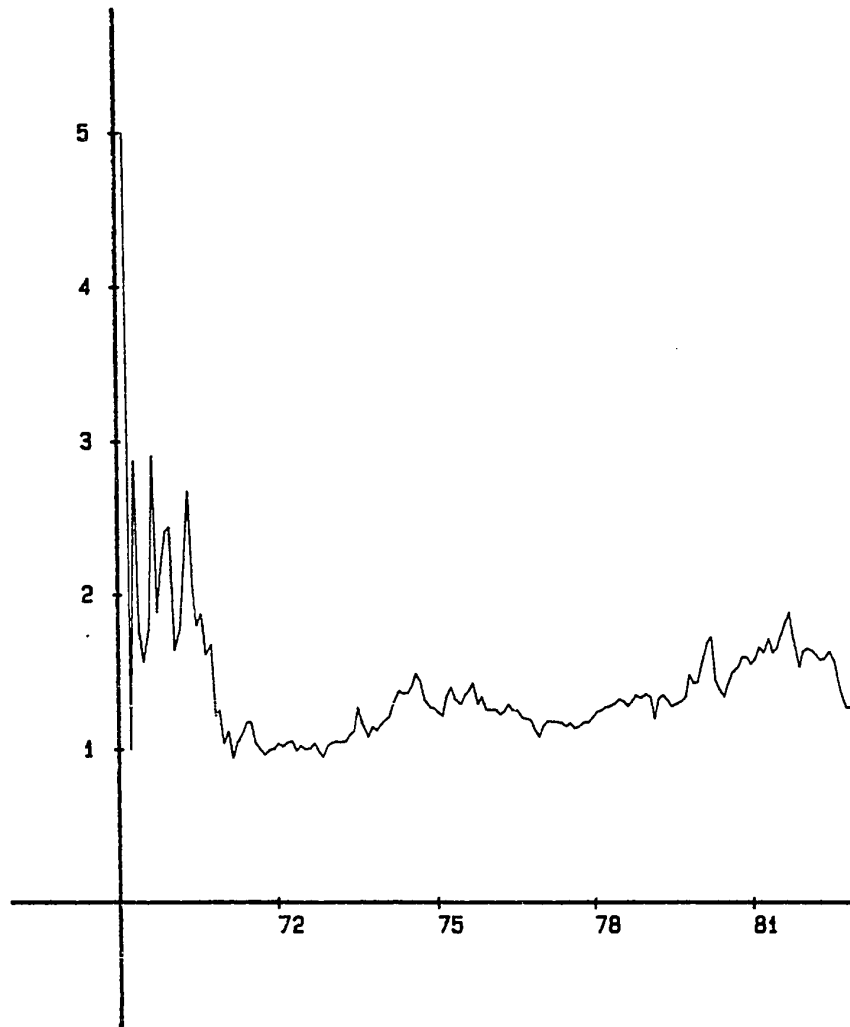


Figure 18. Profit/loss ratio -- Bayesian model.

TABLE VII

MONTH BY MONTH PROFIT OR LOSS -- BAYESIAN RULE

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Tot
1969	16	3	5	-24	45	-15	-5	9	50	-24	22	14	96
1970	2	-41	13	39	39	-35	-20	9	-19	9	-51	6	-49
1971	-38	17	-41	27	17	17	-1	-34	-12	-10	8	4	-46
1972	10	-6	8	1	-16	9	-7	1	10	-16	-11	21	4
1973	8	4	-1	2	14	8	50	-28	-30	25	-8	16	60
1974	10	9	32	30	-6	4	16	31	-15	-37	-16	-3	55
1975	-11	-10	52	20	-24	-11	22	11	19	-46	19	-30	11
1976	-4	2	-12	9	20	-17	2	-23	-6	-3	-28	-27	-87
1977	44	12	-1	0	-4	-11	11	-15	8	15	-1	23	81
1978	18	5	10	6	13	17	-8	-16	20	27	-10	19	101
1979	-9	-78	95	19	-18	-26	16	12	21	109	-30	6	117
1980	100	96	32	-152	-46	-38	72	62	42	54	0	-29	193
1981	37	72	-23	92	-56	28	72	84	63	-103	-162	91	195
1982	17	-12	-20	-32	6	55	-55	-104	-94	-86	0	-17	-342

U.S. government and private publications. All revisions of the data, concepts, and coverage made by the source agencies are incorporated into the databank. Each individual series in the bank is consistent, continuous, and up to date. All series which have seasonal fluctuations have been adjusted either by the source agency or by Citibank.

The database is maintained by what Citibank calls "a group of well trained and experienced data experts." It is available through a number of commercial computer time-sharing services as well as on a computer tape supplied directly by the bank. An alternative source for many of the series is Business Conditions Digest, a monthly publication of the U.S. Department of Commerce (1983). A computer tape containing this data can be purchased at moderate cost from the Bureau of Economic Analysis.

DYNAMIC DETRENDING

Recall the univariate ordinary least squares regression of Y on X

$$(55) \quad Y_i = a_1 X_i + a_2 + e_i$$

where the objective is to minimize the sum of the squared errors

$$(56) \quad \sum_i e_i^2 = \sum_i (Y_i - a_1 X_i - a_2)^2$$

This gives rise to the normal equations

$$\begin{aligned} \underline{Y} &= a_1 \underline{X} + a_2 \\ (57) \quad \sum_i X_i Y_i &= a_1 \sum_i X_i + a_2 \sum_i X_i^2 \end{aligned}$$

where \underline{X} and \underline{Y} are, respectively, the means of X and Y . These equations are solved simultaneously for a_1 and a_2 .

The usual geometric interpretation of this formulation is shown in Figure 19 where the regression line is located so as to minimize the sum of the squared vertical distances from the points to the line. Different insight is obtained by considering the problem in deviation form. Define

$$(58) \quad Y' = Y - \underline{Y} \quad \text{and} \quad X' = X - \underline{X}$$

Then the coefficients of the regression equation of Y' on X'

$$(59) \quad Y' = a'_1 X' + a'_2$$

turn out to be $a'_1 = a_1$ and $a'_2 = 0$. In other words, through an axis translation, deviations of Y from its mean are explained by deviations of X from its mean (see Figure 20). So viewed, the original parameter a_1 is the one of interest and the original constant a_2 is nothing more than a mean adjustment factor with no "explanatory force." All of this, of course, is standard, first chapter regression theory (Maddala, 1977).

Suppose now that Y and X are values of two given time series so that the sequence of values has meaning. Suppose, too, that the mean of Y does not change much over the period, but that the mean of X increases. In other words, the X -series trends upward, but the Y -series is

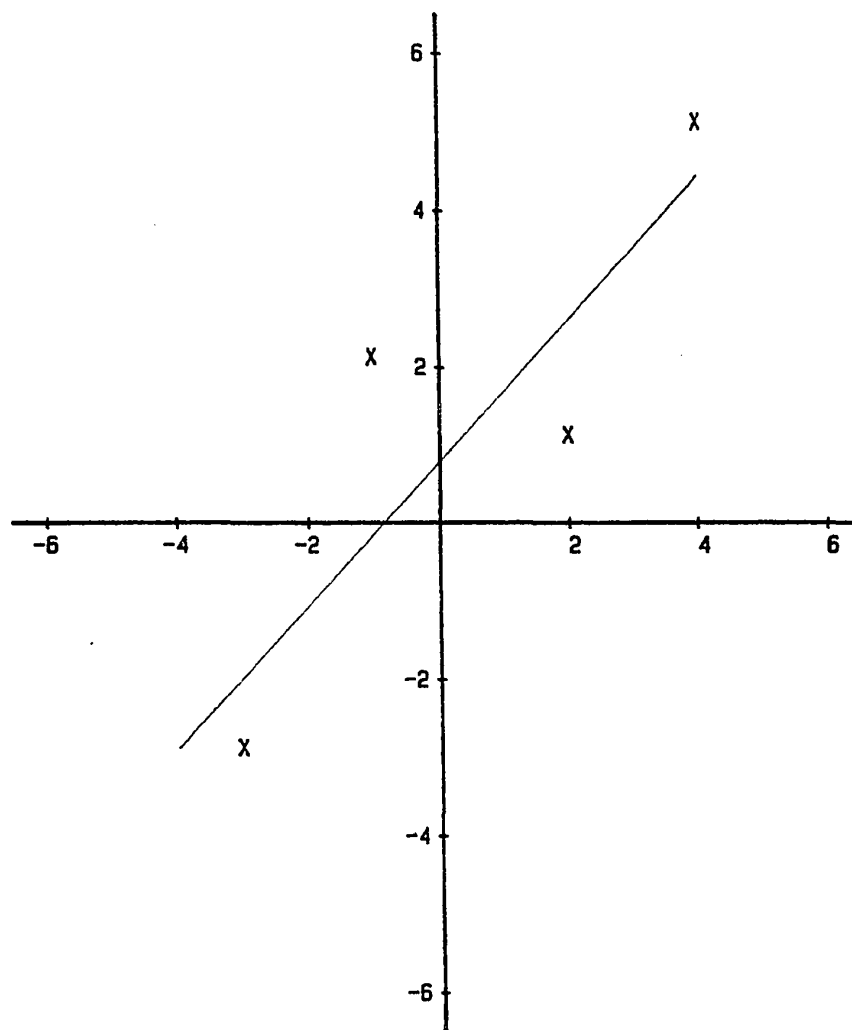


Figure 19. Ordinary least squares regression.

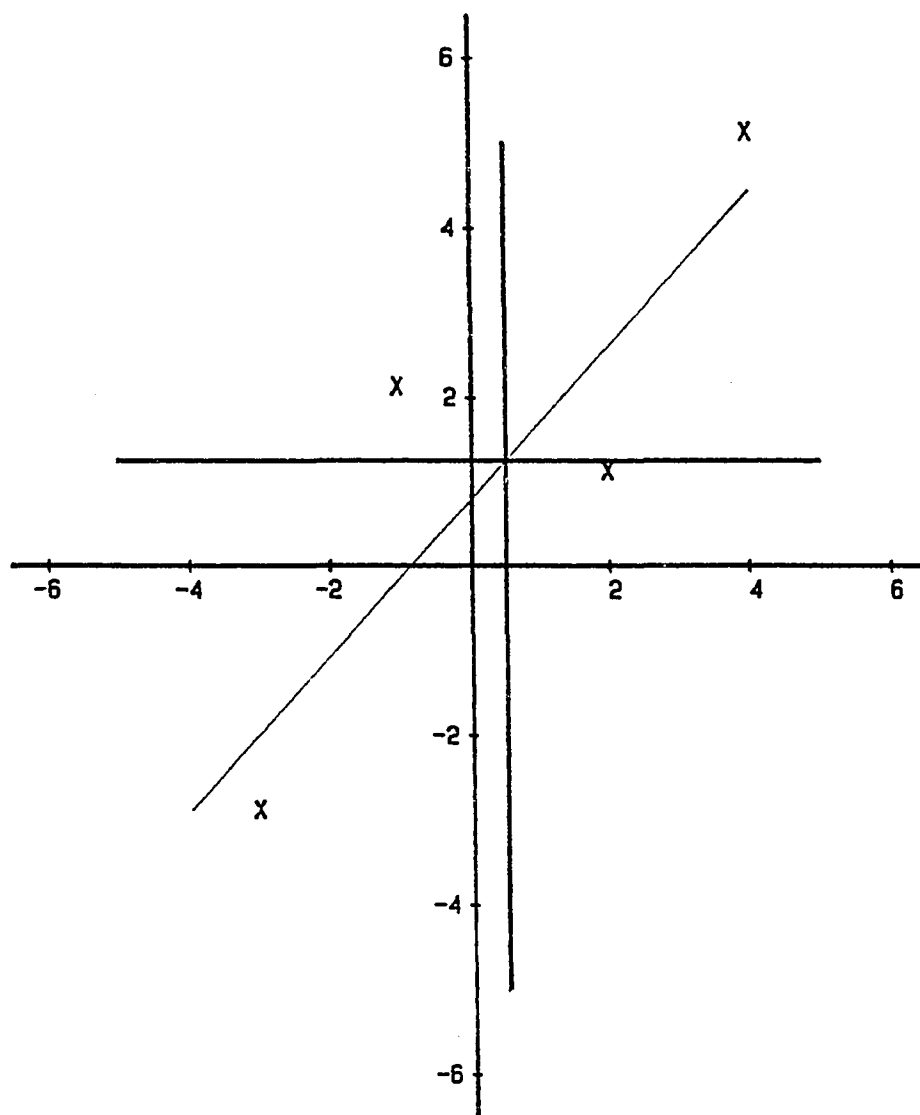


Figure 20. Axis translation for regression in deviations-from-means form.

flat. In such a situation one might postulate that deviations of Y from its mean are explained by deviations of X from its trend, not from its mean. This encourages an examination of methods for determining the trend in a series, since once the trend is found the detrended series is immediately available by subtraction.

Of the many methods of determining the trend of a series, a commonly employed approach is to smooth the series using a filter. Filters are simply linear transformations or moving averages of the series. In general, there is no reason why filters cannot be two-sided, using both past and future values of the series. But to maintain the integrity of the information set, only past values can be used for forecasting purposes; consequently one-sided, left-handed filters must be employed. Finite-memory filters use only a fixed number of historical values; infinite-memory filters take into account all past values of the series. While infinite memory filters contain more information, their performance may deteriorate over time if there are changes in trend. A useful compromise which ultimately captures changes in trend yet uses all past information is an infinite-memory filter that gives greater weight to recent data and lesser weight to older data. A method having these attributes can be called dynamic detrending.

An algorithm satisfying these requirements is the

technique of weighted recursive least squares developed above. Applying the sequential method of equations (42), (43), (44), and (45) gives a sequence of two-element coefficient vectors $a_{<j-1>}$, which may be used to project the trend of the series one month ahead

$$(60) \quad a_{1<j-1>}t(j) + a_{2<j-1>}$$

The detrended values, $x_*(j)$, are deviations from the projected trend, i.e.,

$$(61) \quad x_*(j) = x(j) - (a_{1<j-1>}t(j) + a_{2<j-1>})$$

All series used in this study (except the interest rate series itself) were detrended by equation (61) using a weight matrix U with diagonal elements equal to increasing integral powers of (1.0035); in other words, each succeeding value of the series is given approximately .7% more weight than the previous one. A somewhat arbitrary choice, this provides for weights which increase year-over-year by about 8.7%. and it has the effect of maintaining a better balance between new and old points. In unweighted recursive regressions, each additional data point effectively receives less weight. For example, when there are 19 data points, the next contributes about 5%, but when there are 99 points the next contributes only about 1%. Use of the .7% weighting factor assures that all new terms contribute at least 2% to the regression regardless many data points are being used. Figure 21 compares the relative weights assigned to new points in the weighted

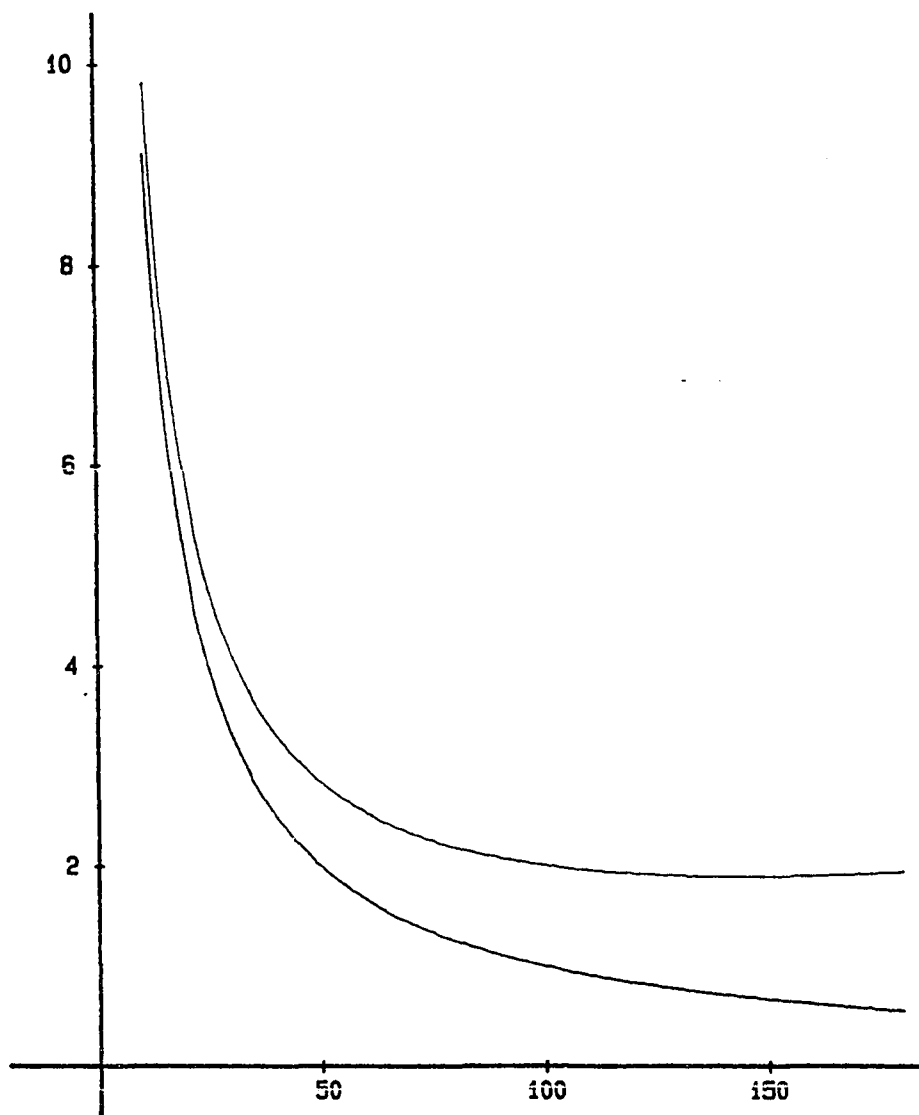


Figure 21. Relative weights assigned to new points in weighted and unweighted regressions (in percent).

and unweighted recursive regressions.

Starting values are always a problem when dealing with infinite memory filters (Makridakis and Wheelwright, 1978). A good solution is simply to obtain extra data at the beginning of the series, detrend that extended series, then drop the early values. In this study the detrending algorithm was applied to values of the series for the years 1967 through 1982. Any detrended values not needed to estimate the model for January, 1968, were dropped.

ECONOMIC THEORY

There is no single, generally accepted theory of interest rates. In fact, agreement cannot even be reached on what interest is. Boehm-Bawerk (1922, 1923) and Conrad (1963) each outline scores of competing interest rate theories; a reasonable synthesis remains elusive. The models developed here follow from classical economic theory.

Take it as given that the rate of interest, representing in some sense the price paid for the use of credit (or the reward received for lending money), reflects the interaction between the supply of credit and the demand for it. Supply comes from a willingness by income earners or other wealth holders to forego current consumption in favor of greater expected consumption in the future; thus the ratio of expected future consumption to foregone current

consumption, sometimes called the rate of time preference, is closely related to the rate of interest.

Demand for credit comes from the choice by some consumers to borrow for current consumption, from governments' need to finance deficits, and from the decision of business firms to invest in inventory and plant and equipment. Business' demand for credit relates interest rates to the marginal efficiency of real investment. Time preference and the marginal efficiency of investment are generally viewed as aspects of the "real" sector of the economy. They determine the "real" interest rate.

Market interest rates, however, are monetary or financial phenomena. Interest rates quoted daily in the business press are "nominal" returns on financial, not real, assets. The relationship of nominal rates to real rates involves inflation. Lenders, seeking to protect and enhance the real purchasing power of the funds they have lent, will demand a premium over and above the real rate of interest if they anticipate inflation will erode the value of the dollars that are repaid. Borrowers, expecting to repay the funds in depreciated currency, will be prepared to pay the higher nominal interest rate. These inflation-related considerations are summarized in an equation named after Irving Fisher (1930), who gave the classical exposition of these ideas:

(62)

$$n = r + p$$

where n is the nominal or market rate of interest, r is the real rate, and p represents the anticipated rate of inflation. Humphrey (1983) discusses the Fisher equation and its historical antecedents in detail.

Although this discussion of the theory of interest rates is very brief, it is sufficient to carry the weight of most of the empirical models developed below. There are, of course, many other factors that affect interest rates. Three short articles in the Review of the Federal Reserve Bank of St. Louis clearly present the basic principles from a monetarist perspective (Bowsher, 1980; Santoni and Stone, 1981a and 1981b). The papers by Hamburger and Silber (1969) and Feldstein and Eckstein (1970) have been mentioned earlier and are recommended. Pring's (1981) book and the treatment by the Conference Board (1973) look at the question from a forecaster's viewpoint.

Another point needs emphasis: it is one thing to say that the factors discussed affect interest rates contemporaneously; it is quite another to say that current values of these variables are useful in forecasting future interest rates. A judicious choice of the specific series used to represent these factors may, however, overcome this objection.

FORECASTING MODELS

The Investment Model.

An important and highly variable element of credit demand is borrowing by business for capital investment. One leading indicator of future capital spending is construction contracts awarded for commercial and industrial buildings, a copyrighted series published by the F. W. Dodge Division of McGraw-Hill Information Systems Company, used here by permission. Since this series represents current commitments by business to spend for future capital expenditures, it possesses the required element of futurity.

The first model forecasts the direction of change in interest rates, DCR_{t+1} , using only this variable (denoted by the symbol CC). Because of delays in collection and publication, values of the series cannot be incorporated into the information set except with a two-month or longer lag. (Lags can be measured either from the month that the forecast is made or from the month the forecast is for. The former usage is employed here. Thus in the following equation there is a two-month lag for data to enter the information set or a three-month lag for the information to enter the forecast. In the format of equation (49) the lag is $k = 2$ and the forecast horizon is $h = 1$.)

The forecasting model is

$$(63) \quad DCR_{t+1} = \text{sgn} (w_1 CC_{t-2} + w_2)$$

The reported and detrended construction contract series are shown in Figure 22, the net profit in Figure 23, the profit/loss ratio in Figure 24, and the month-by-month profits in Table VIII. The coefficients (tabulated in Appendix E) have the anticipated sign.

Summary statistics for this model are:

Net Profit:	1195 basis points
Profit/Loss Ratio:	1.71
Number of Hits:	101
Rate of Return:	.943 percent
Worst Computed Loss:	265 basis points

The Price-Investment Model.

The second model considers not only investment commitments but also inflation. This time investment commitments are measured by a different series, contracts and orders for plant and equipment, lagged two months (from McGraw-Hill and Bureau of the Census). This series, denoted CO, also represents current commitments for future capital expenditures and so can be expected to lead the actual credit demand.

The second variable in this model is inflation (denoted IN), represented by the Commodity Research Bureau's index of spot prices on 22 commodities (a copyrighted series used by permission). The commodity price series is available daily so that no lag is required. Several of the references cited above provide evidence that

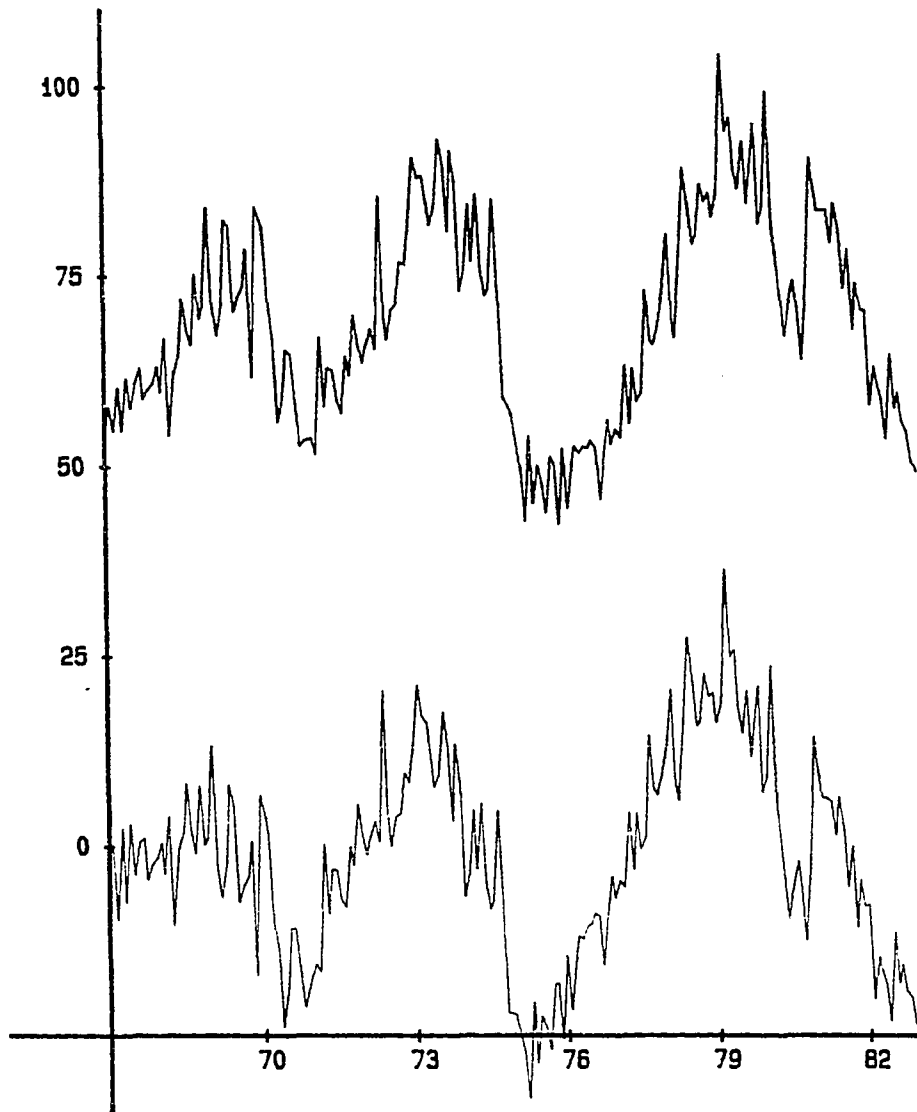


Figure 22. Construction contracts for commercial and industrial buildings -- reported and detrended series. (This is a copyrighted series used by permission of F.W. Dodge Division, McGraw-Hill Information Systems Company.)

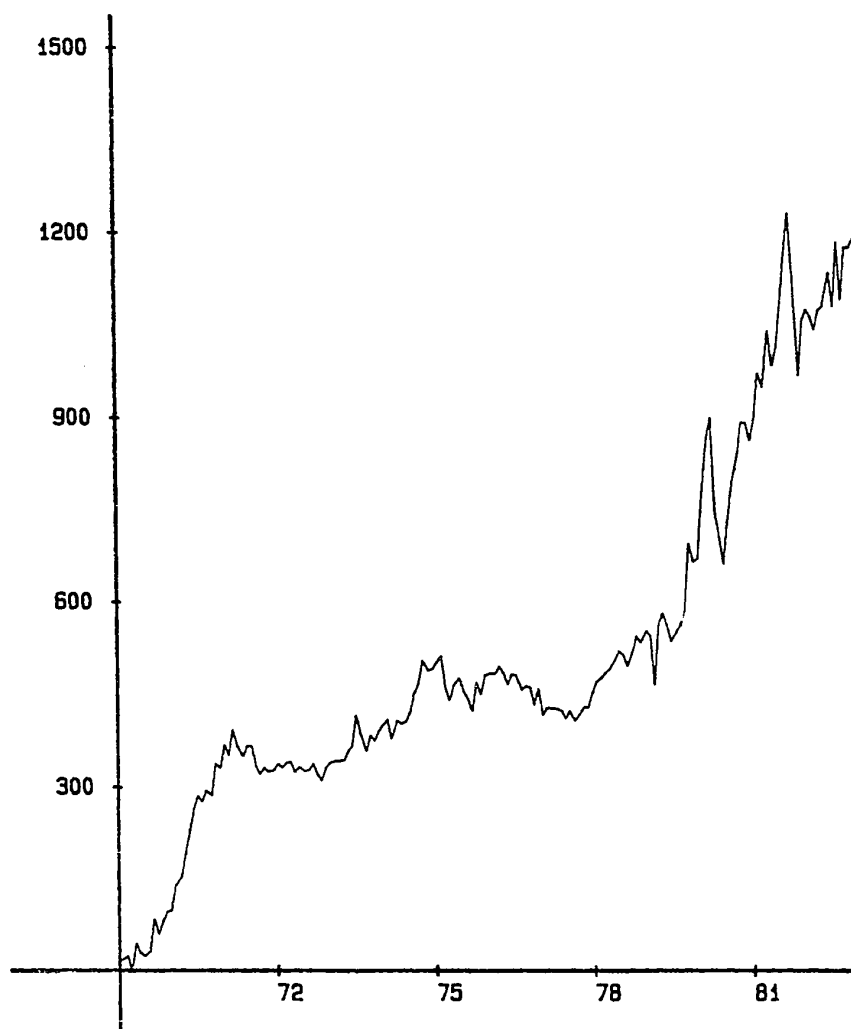


Figure 23. Cumulative net profit -- investment model.



Figure 24. Profit/loss ratio — investment model.

TABLE VIII

MONTH BY MONTH PROFIT OR LOSS -- INVESTMENT MODEL

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Tot
1969	16	3	5	-24	45	-15	-5	9	50	-24	22	14	96
1970	2	41	13	39	39	35	20	-9	19	-9	51	-6	235
1971	38	-17	41	-27	-17	17	1	-34	-12	10	-8	4	-4
1972	10	-6	8	1	-16	9	-7	1	10	-16	-11	21	4
1973	8	4	-1	2	14	8	50	-28	-30	25	-8	16	60
1974	10	9	-32	30	-6	4	16	31	15	37	-16	3	101
1975	11	10	-52	-20	24	11	-22	-11	-19	46	-19	30	-11
1976	4	-2	12	-9	-20	17	-2	-23	6	-3	-28	27	-21
1977	-44	12	-1	0	-4	-11	11	-15	8	15	-1	23	-7
1978	18	5	10	6	13	17	-8	-16	20	27	-10	19	101
1979	-9	-78	95	19	-18	-26	16	12	21	109	-30	6	117
1980	100	96	32	-152	-46	-38	72	62	42	54	0	-29	193
1981	37	72	-23	92	-56	28	72	84	63	-103	-162	91	195
1982	17	-12	-20	32	6	55	-55	104	-94	86	0	17	136

expectations about future rates of inflation are determined in part by past inflation; hence historical inflation may aid in forecasting the p term in equation (62).

Moreover, this particular measure of inflation tends to lead the more widely followed indicators like the Consumer Price Index and the GNP implicit price deflator.

The forecasting model using these variables is

$$(64) \quad DCR_{t+1} = \text{sgn} (w_1 CO_{t-2} + w_2 IN_t + w_3)$$

The reported and detrended variables are shown in Figures 25 and 26, the net profit in Figure 27, the profit/loss ratio in Figure 28, and the month-by-month profits in Table IX. The coefficients are tabulated in Appendix E and have the anticipated sign most of the time.

Summary statistics for this model are:

Net Profit:	1219 basis points
Profit/Loss Ratio:	1.73
Number of Hits:	101
Rate of Return:	.957 percent
Worst Computed Loss:	373 basis points

The Labor Model.

For many years the conventional wisdom had it that there is a trade-off between the rate of unemployment and the rate of inflation: by permitting (or causing) the rate of unemployment to rise, policymakers believed they could reduce the rate of inflation. The supposed trade-off is graphically presented as the Phillips curve, which is a

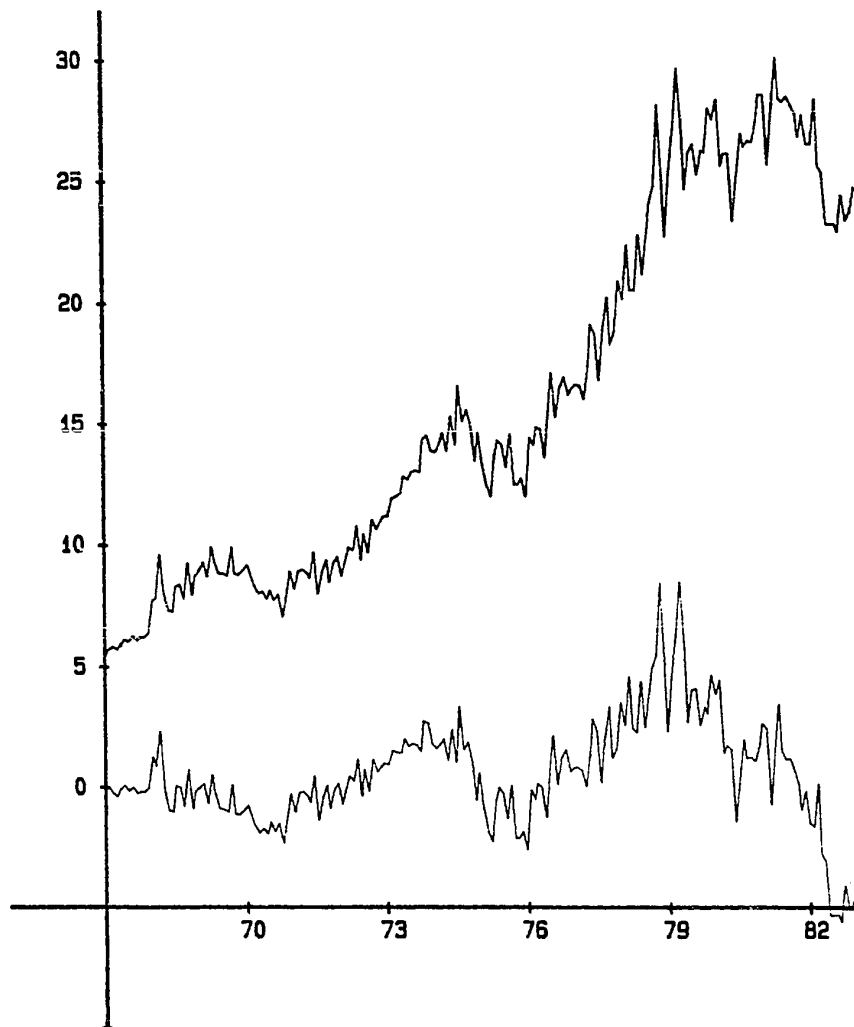


Figure 25. Contracts and orders for plant and equipment
-- reported and detrended series.

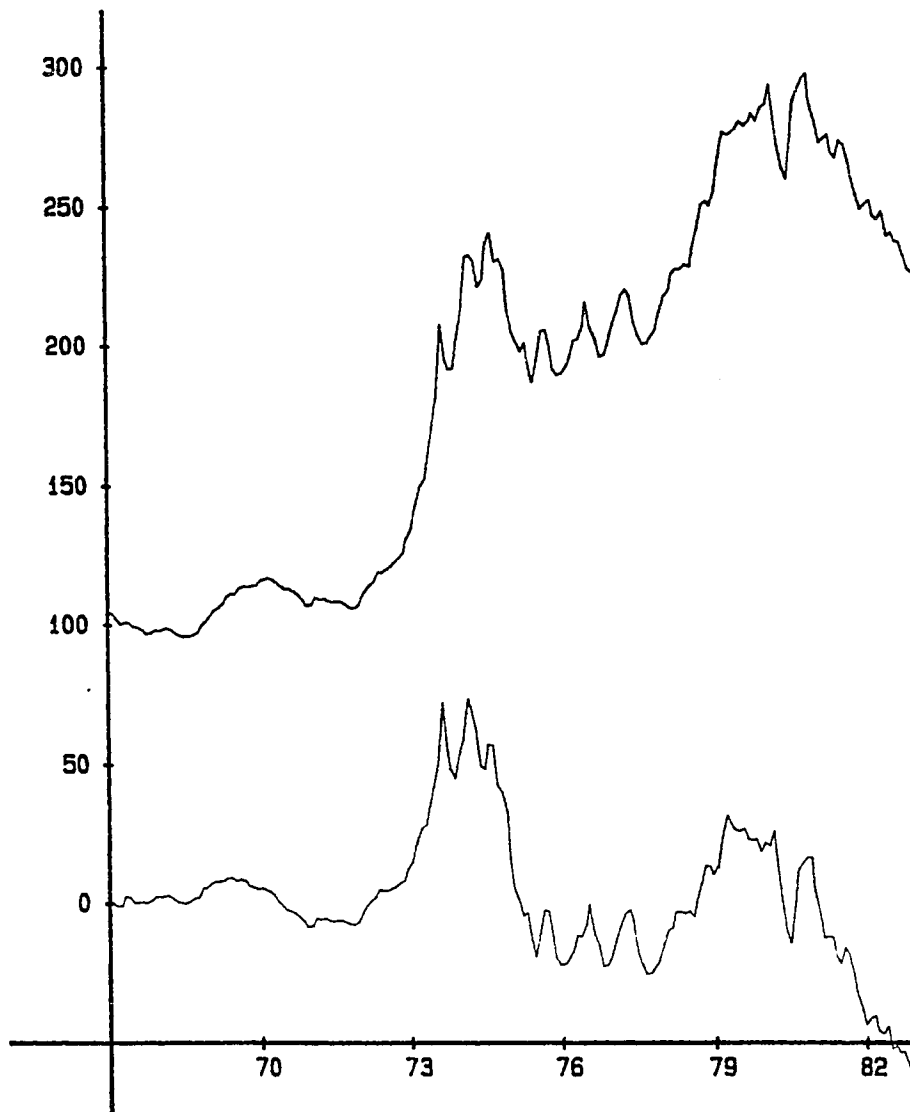


Figure 26. Commodity Research Bureau index of spot prices on 22 commodities — reported and detrended series. {This is a copyrighted series used by permission of the Commodity Research Bureau, Inc.}

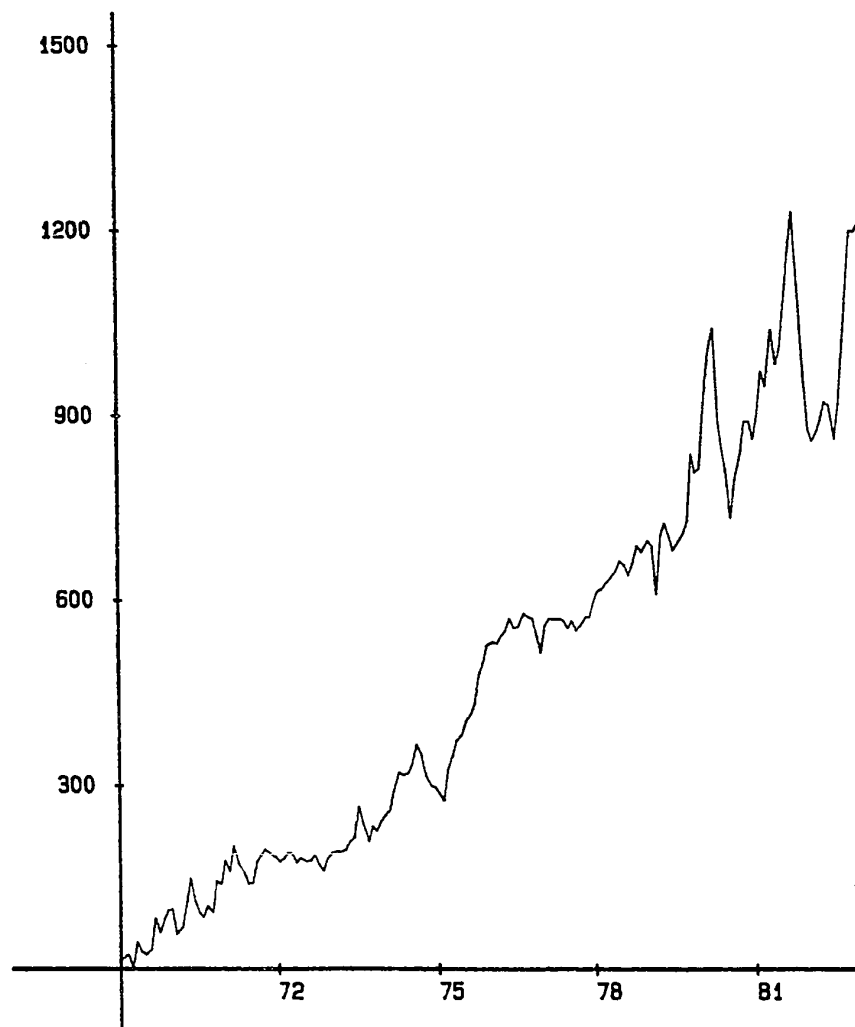


Figure 27. Cumulative net profit -- price-investment model.

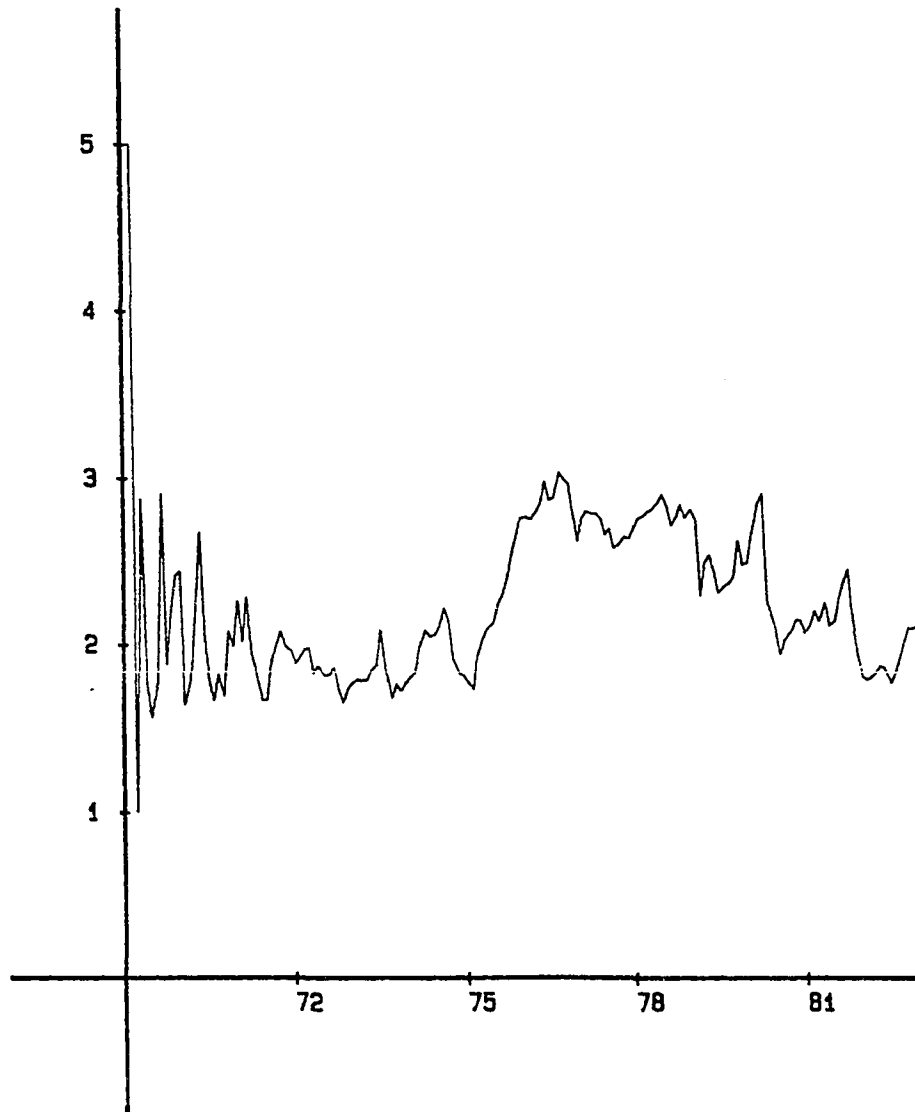


Figure 28. Profit/loss ratio -- price-investment model.

TABLE IX

MONTH BY MONTH PROFIT OR LOSS -- PRICE-INVESTMENT MODEL

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Tot
1969	16	3	5	-24	45	-15	-5	9	50	-24	22	14	96
1970	2	-41	13	39	39	-35	-20	-9	19	-9	51	-6	43
1971	38	-17	41	-27	-17	-17	1	34	12	10	-8	-4	46
1972	-10	6	8	1	-16	9	-7	1	10	-16	-11	21	-4
1973	8	4	-1	2	14	8	50	-28	-30	25	-8	16	60
1974	10	9	32	30	-6	4	16	31	-15	-37	-16	-3	55
1975	-11	-10	52	20	24	11	22	11	19	46	19	30	233
1976	4	-2	12	9	20	-17	2	23	-6	-3	-28	-27	-13
1977	44	12	-1	0	-4	-11	11	-15	8	15	-1	23	81
1978	18	5	10	6	13	17	-9	-16	20	27	-10	19	101
1979	-9	-78	95	19	-18	-26	16	12	21	109	-30	6	117
1980	100	96	32	-152	-46	-38	-72	62	42	54	0	-29	49
1981	37	72	-23	92	-56	28	72	84	63	-103	-162	-91	13
1982	-17	12	20	32	-6	-55	55	104	94	86	0	17	342

plot of the rate of inflation on the vertical axis versus the rate of unemployment on the horizontal axis. The transmission mechanism from the unemployment rate to the inflation rate is supposed to be through aggregate demand: if more people were unemployed, it is argued, aggregate demand would fall and the upward pressure on prices would lessen. About the time the term stagflation was coined in the 1970's, the Phillips curve became a Phillips cloud.

Geoffrey Moore (1979) of the National Bureau of Economic Research has attempted to revitalize the theory, although in modified form. He claims that the deficiency is in the variable used to measure aggregate demand, not in the assumed transmission mechanism itself. The problem, Moore argues, has been the shifting relationship between employment and the unemployment in recent years. In the 1950's and early 1960's the working population was increasing at a relatively slow pace, so that only about two new jobs needed to be created per year in order to reduce unemployment by one job. By the late 1970's, however, the working-age population was expanding more rapidly so that six additional jobs were required to take one person off the unemployment rolls. Consequently, Moore suggests, the Phillips curve should be redrawn with the percentage of the working-age population, not the unemployment rate, on the horizontal axis. The result is an upward sloping curve associating higher employment

ratios with higher inflation rates.

All of this suggests that the employment ratio, the ratio of the employed civilian population to the total civilian population (both counting only persons 16 years and older), might serve as a leading indicator of inflation and hence of interest rates. Letting the ratio be represented by the symbol ER , the forecasting equation is

$$(65) \quad DCR_{t+1} = \text{sgn} (w_1 ER_{t-2} + w_2)$$

The reported and detrended employment ratio is shown in Figure 29, the net profit in Figure 30, the profit/loss ratio in Figure 31, and the month-by-month profits in Table X. The sign of the coefficients (tabulated in Appendix E) is in accordance with the theory.

Summary statistics are:

Net Profit:	1193 basis points
Profit/Loss Ratio:	1.71
Number of Hits:	95
Rate of Return:	.932 percent
Worst Computed Loss:	265 basis points

An Atheoretical Model.

The three previous models have had the support of economic theory; this one does not. Nonetheless it rests solidly on a long tradition of atheoretical economics going back probably to the birth of the science, enhanced by the efforts of Burns and Mitchell in the 1930's, and continued currently in such works as Sargent and Sims'

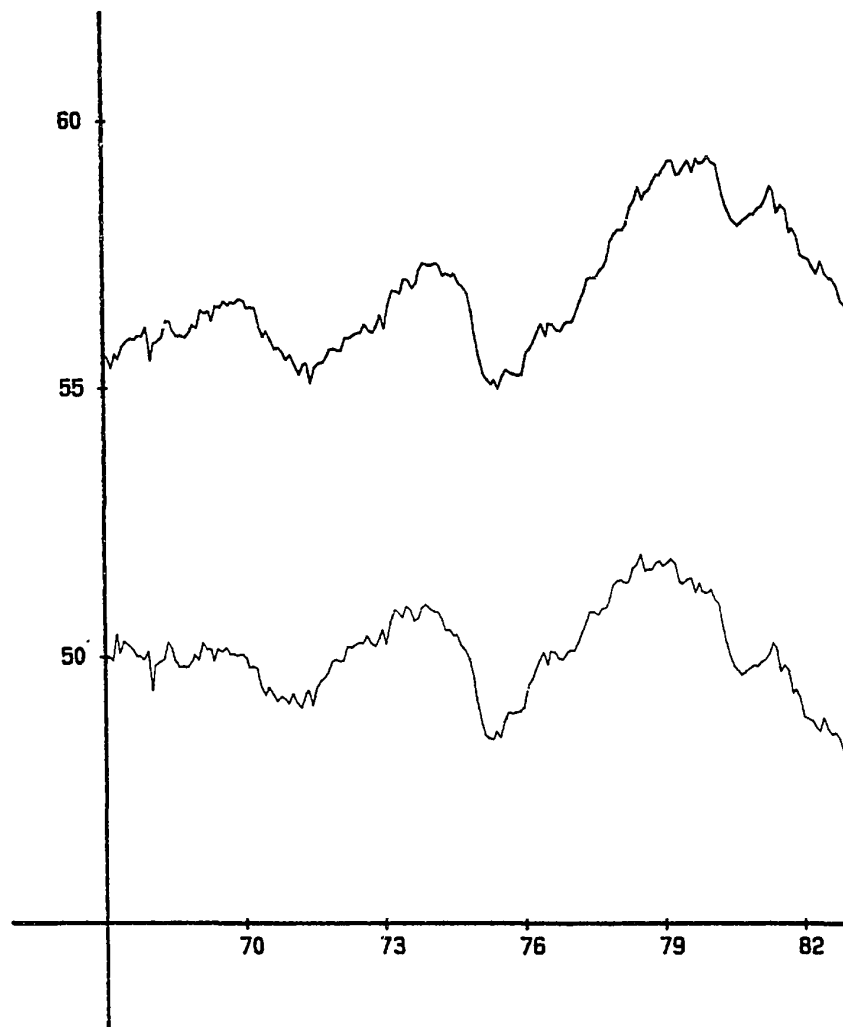


Figure 29. Civilian employment ratio—reported and detrended series (rescaled).

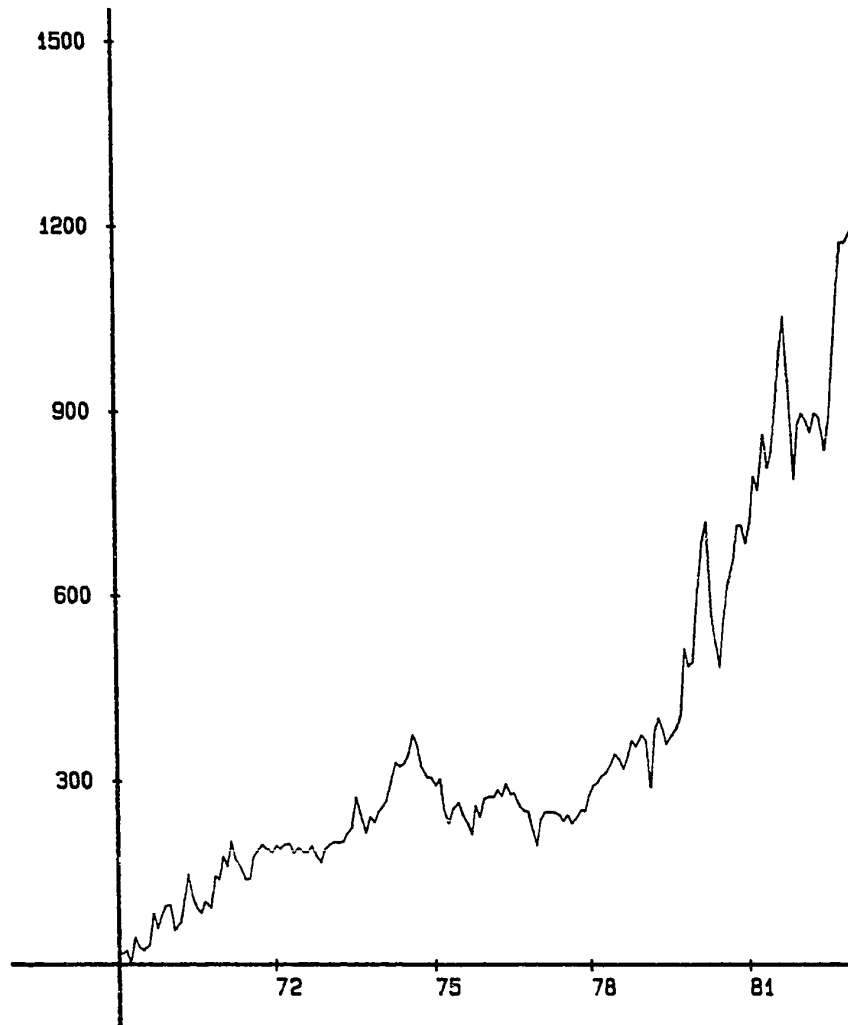


Figure 30. Cumulative net profit--labor model.

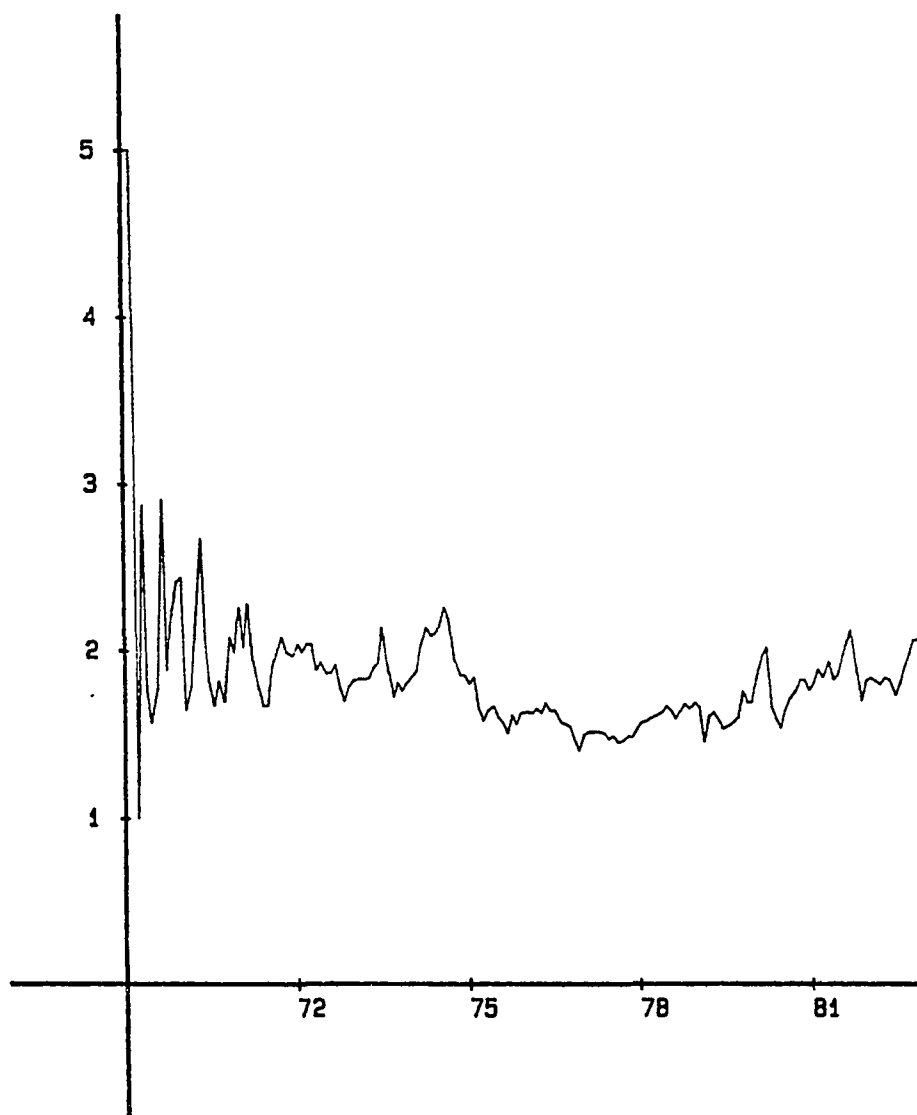


Figure 31. Profit/loss ratio--labor model.

TABLE X

MONTH BY MONTH PROFIT OR LOSS -- LABOR MODEL

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Tot
1969	16	3	5	-24	45	-15	-5	9	50	-24	22	14	96
1970	2	-41	13	39	39	-35	-20	-9	19	-9	51	-6	43
1971	38	-17	41	-27	-17	-17	1	34	12	10	-8	-4	46
1972	10	-6	8	1	-16	9	-7	1	10	-16	-11	21	4
1973	8	4	-1	2	14	8	50	-28	-30	25	-8	16	60
1974	10	9	32	30	-6	4	16	31	-15	-37	-16	-3	55
1975	-11	10	-52	-20	24	11	-22	-11	-19	46	-19	30	-33
1976	4	-2	12	-9	20	-17	2	-23	-6	-3	-28	-27	-77
1977	44	12	-1	0	-4	-11	11	-15	8	15	-1	23	81
1978	18	5	10	6	13	17	-8	-16	20	27	-10	19	101
1979	-9	-78	95	19	-18	-26	16	12	21	109	-30	6	117
1980	100	96	32	-152	-46	-38	72	62	42	54	0	-29	193
1981	37	72	-23	92	-56	28	72	84	63	-103	-162	91	195
1982	17	-12	-20	32	-6	-55	55	104	94	86	0	17	312

(1977) "Business Cycle Modeling Without Pretending to Have Too Much a Priori Economic Theory." Economic empiricism needs no apology--more often than not, theories were developed to explain observed regularities.

Three of the most watched economic series were discussed briefly in Chapter III: the leading, coincident and lagging indexes of economic activity. The first difference of the ratio of the coincident index to the lagging index (denoted DRCL), is combined with construction contract series (CC) to produce a model superior to the three developed above.

$$(66) \quad DCR_{t+1} = \text{sgn} (w_1 DRCL_{t-1} + w_2 CC_{t-2} + w_3)$$

The reported and detrended ratio of coincident to lagging indicators is shown in Figure 32, the net profit in Figure 33, the profit/loss ratio in Figure 34, and the month-by-month profits in Table XI. The coefficients are tabulated in Appendix E.

Summary statistics for this model are:

Net Profit:	1477 basis points
Profit/Loss Ratio:	1.96
Number of Hits:	102
Rate of Return:	1.024 percent
Worst Computed Loss:	190 basis points

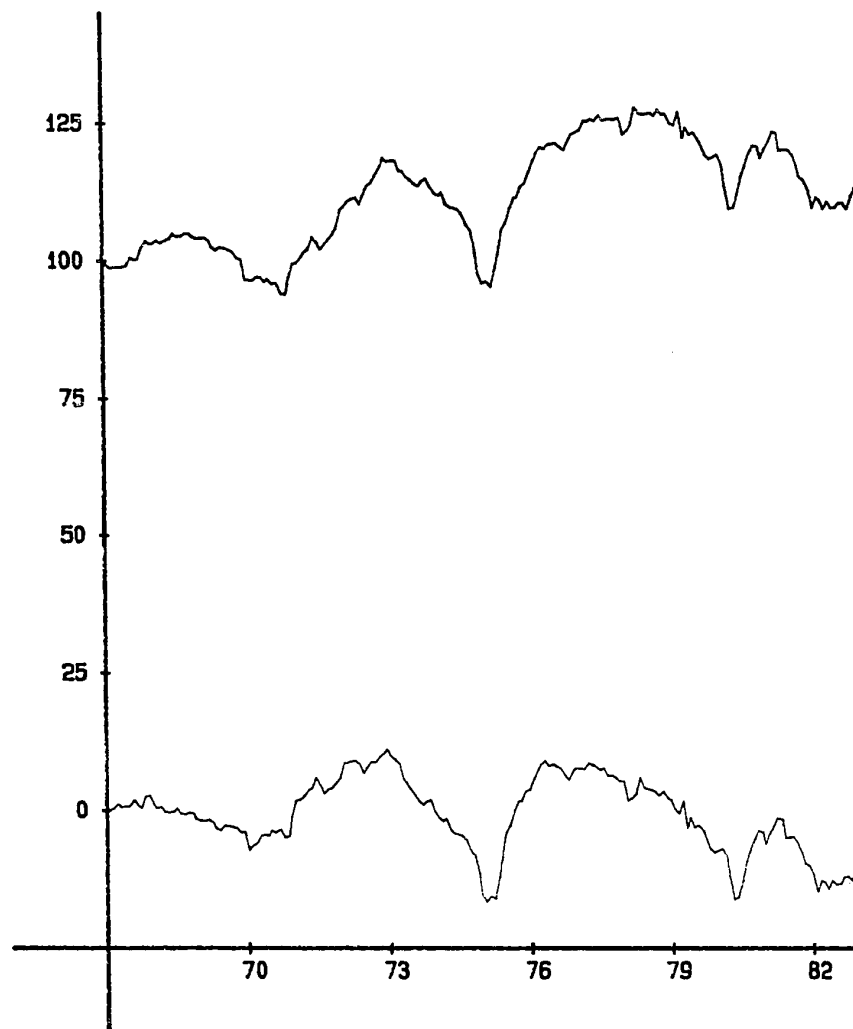


Figure 32. Ratio of coincident to lagging indicators--reported and detrended series.

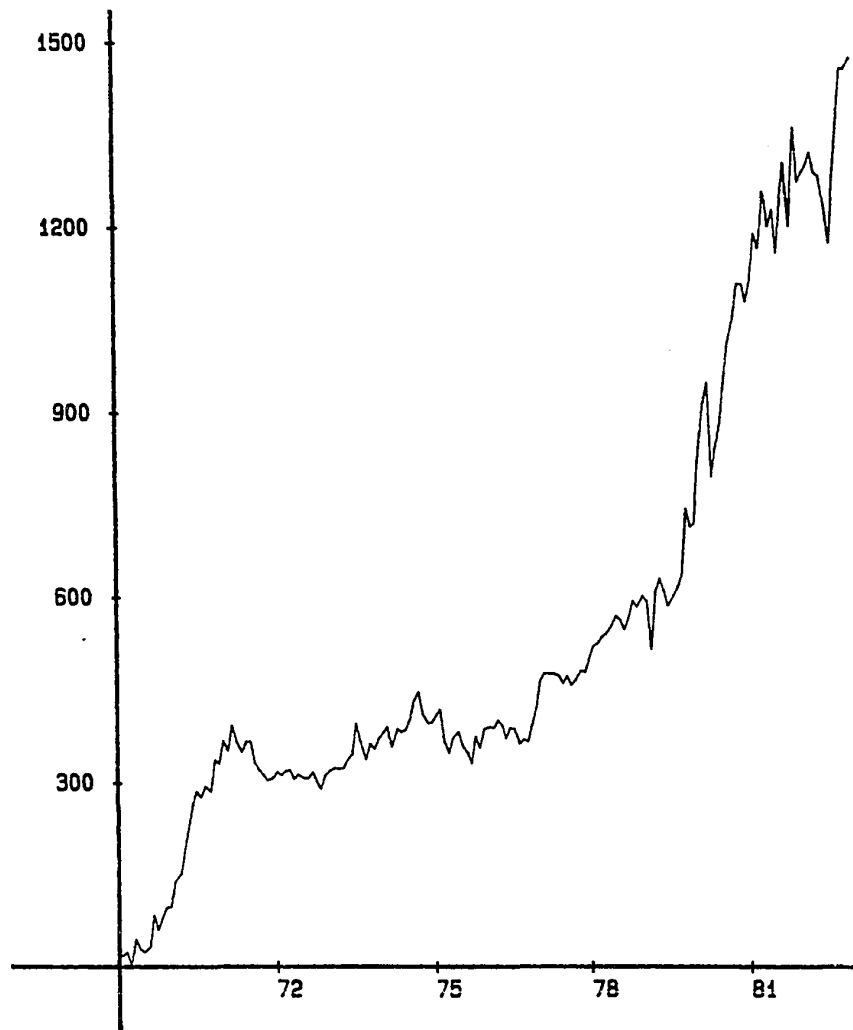


Figure 33. Cumulative net profit--atheoretical model.

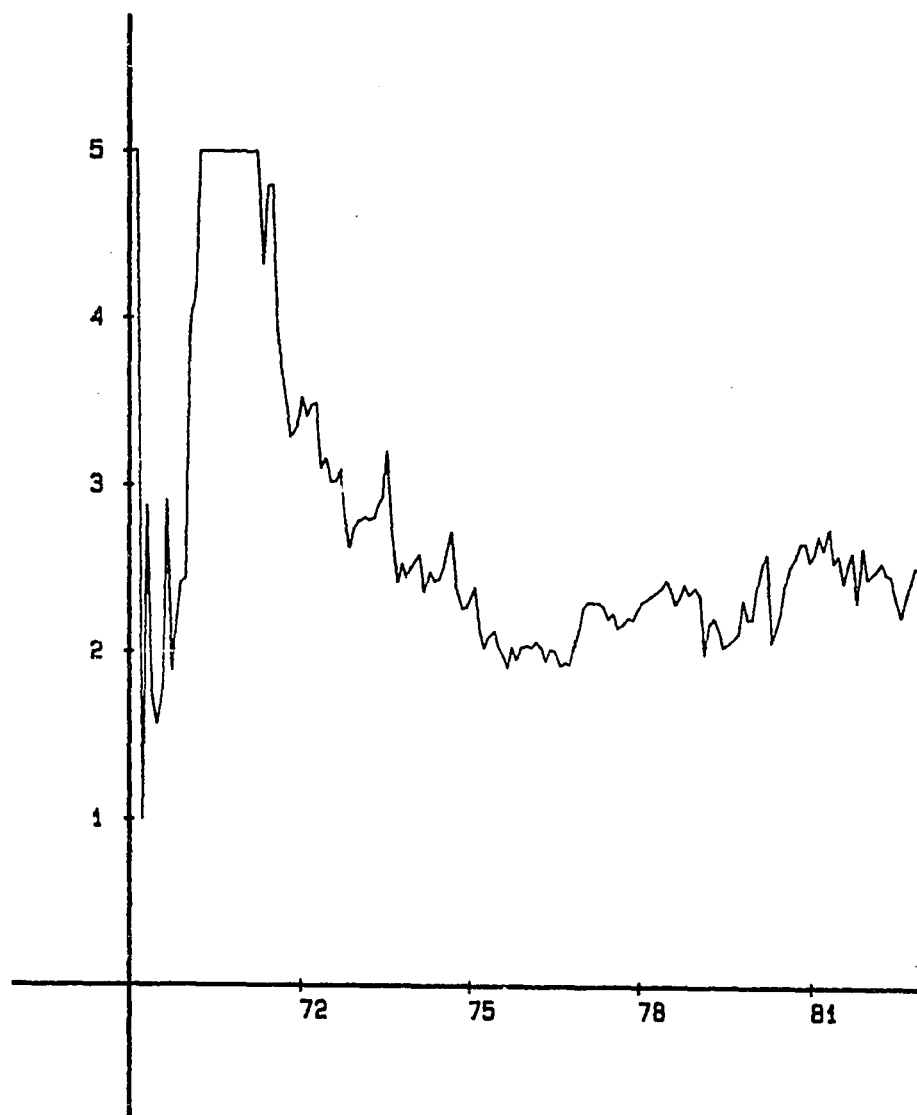


Figure 34. Profit/loss ratio--atheoretical model.

TABLE XI

MONTH BY MONTH PROFIT OR LOSS -- ATHEORTICAL MODEL

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Tot
1969	16	3	5	-24	45	-15	-5	9	50	-24	22	14	96
1970	2	41	13	39	39	35	20	-9	19	-9	51	-6	235
1971	38	-17	41	-27	-17	17	1	-34	-12	-10	-3	4	-24
1972	10	-6	8	1	-16	9	-7	1	10	-16	-11	21	4
1973	8	4	-1	2	14	8	50	-28	-30	25	-8	16	60
1974	10	9	-32	30	-6	4	16	31	15	-37	-16	3	27
1975	11	10	-52	-20	24	11	-22	-11	-19	46	-19	30	-11
1976	4	-2	12	-9	-20	17	-2	-23	6	-3	28	27	35
1977	44	12	-1	0	-4	-11	11	-15	8	15	-1	23	81
1978	18	5	10	6	13	17	-8	-16	20	27	-10	19	101
1979	-9	-78	95	19	-18	-26	16	12	21	109	-30	6	117
1980	100	96	32	-152	46	38	72	62	42	54	0	-29	361
1981	37	72	-23	92	-56	28	-72	84	63	-103	162	-91	193
1982	17	12	20	-32	-6	-55	-55	104	94	86	0	17	202

EVALUATING THE FORECASTING MODELS

How well do the models developed here perform in comparison with their naive competitors, the random walk model and the simple Bayesian rule? Figures 35 through 39 are graphs of cumulative relative frequency functions from the Monte Carlo simulation described earlier for each of the five variables: the net profit, the profit/loss ratio, the number of hits, the rate of return, and the worst computed loss. The results achieved by each of the decision rules are indicated on the frequency graphs by B (for the Bayesian rule), I (for the investment model), P (for the price-investment model), L (for the labor model) and A (for the atheoretical model). When one symbol is placed above another on the graphs, that indicates that both models achieved approximately equal results.

One measure of a model's ability to forecast is the level of confidence at which one can reject the following null hypotheses in favor of their alternatives.

H_0 : Net profit = 0

H_1 : Net profit > 0

H_0 : Number of hits = 82.5

H_1 : Number of hits > 82.5

H_0 : Rate of return = .616

H_1 : Rate of return > .616

Only the last of these, the one related to the rate of return, is a proper test of the efficient market theory; under certain circumstances the three other hypotheses tests may contain biases and might improperly reject the EMT. For example, the number of hits achieved by a given

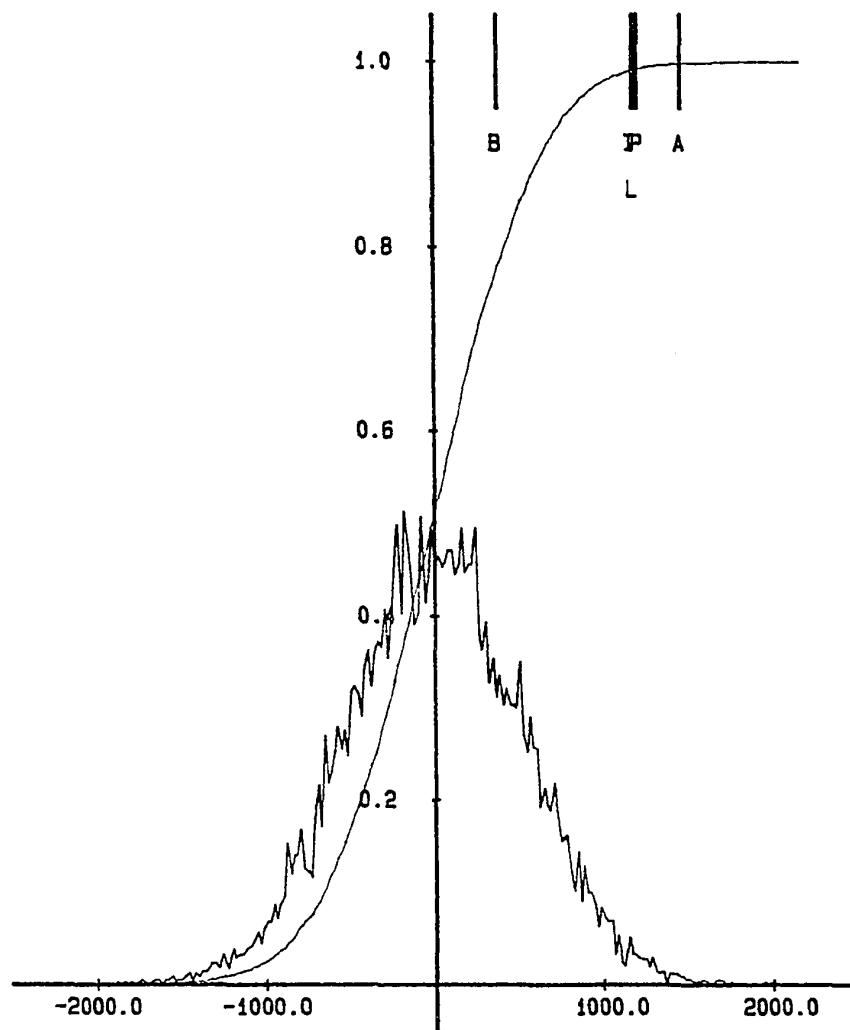


Figure 35. Relative frequency {rescaled} and cumulative relative frequency for the net profit in the Monte Carlo simulation, comparing the models.

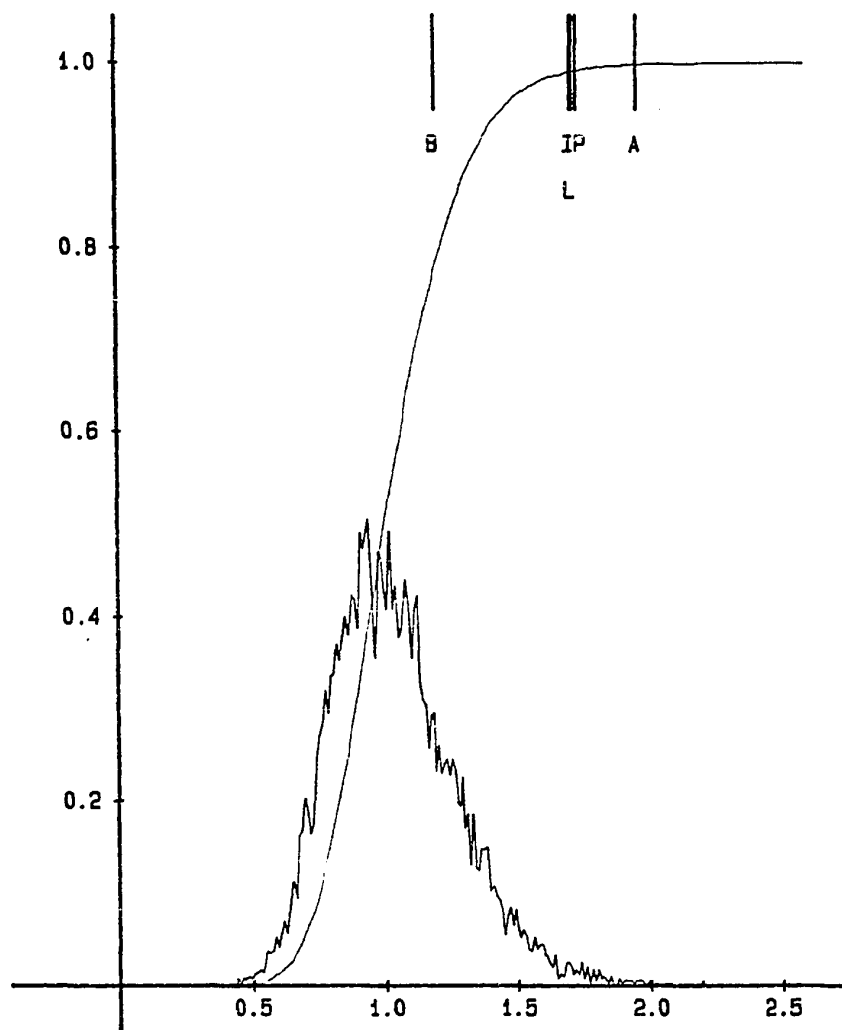


Figure 36. Relative frequency (rescaled) and cumulative relative frequency for the profit/loss ratio in the Monte Carlo simulation, comparing the models.

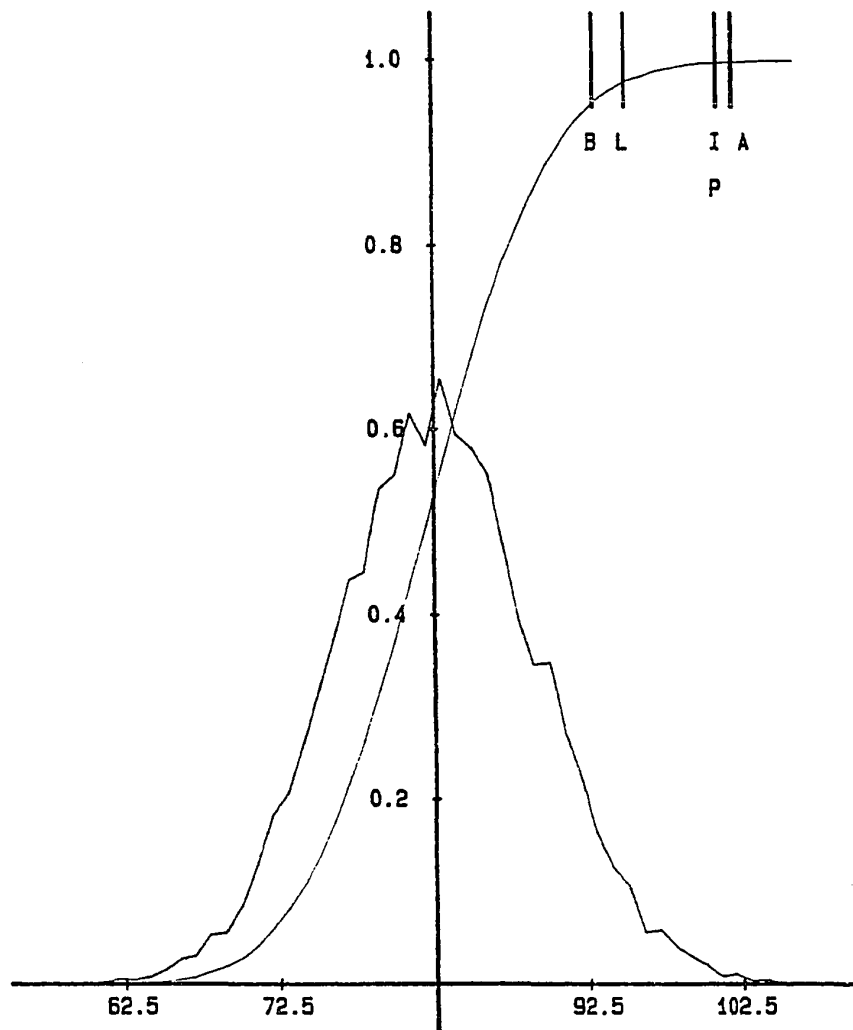


Figure 37. Relative frequency (rescaled) and cumulative relative frequency for the number of hits in the Monte Carlo simulation, comparing the models.

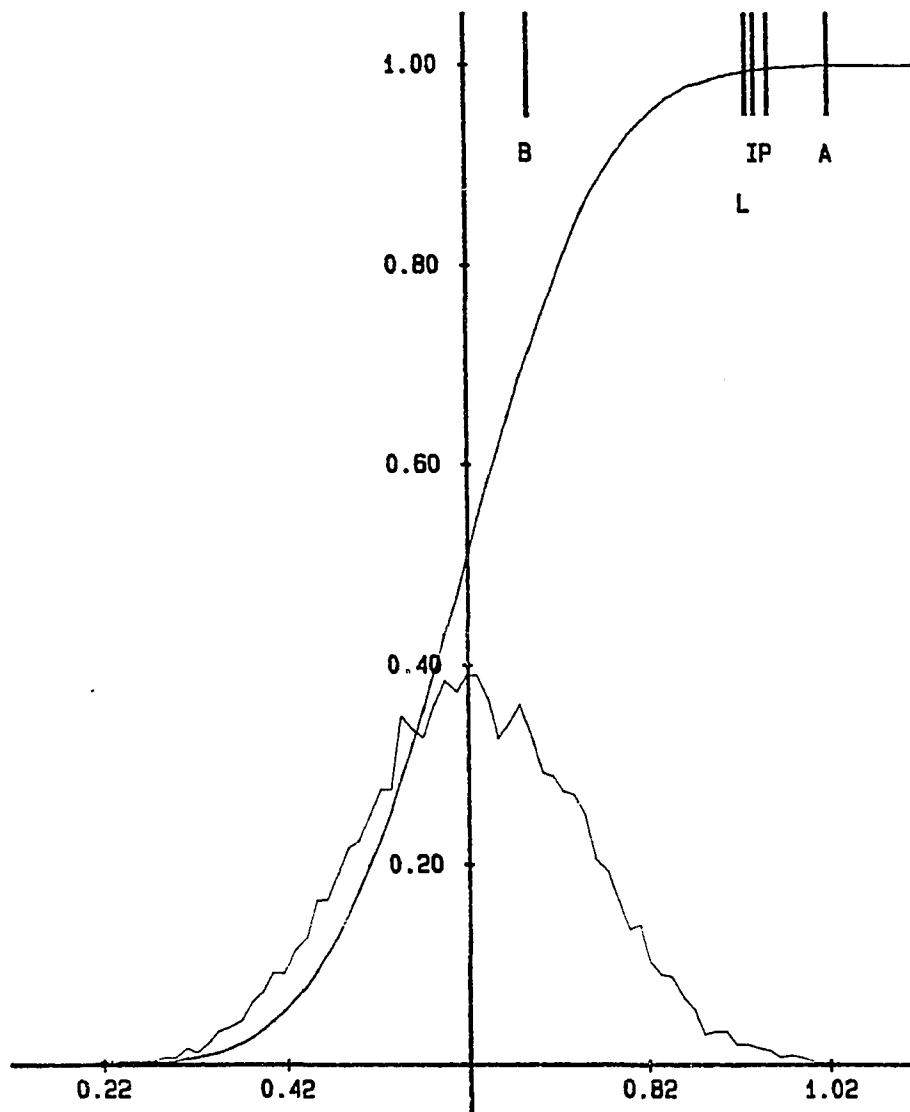


Figure 38. Relative frequency (rescaled) and cumulative relative frequency for the rate of return in the Monte Carlo simulation, comparing the models.

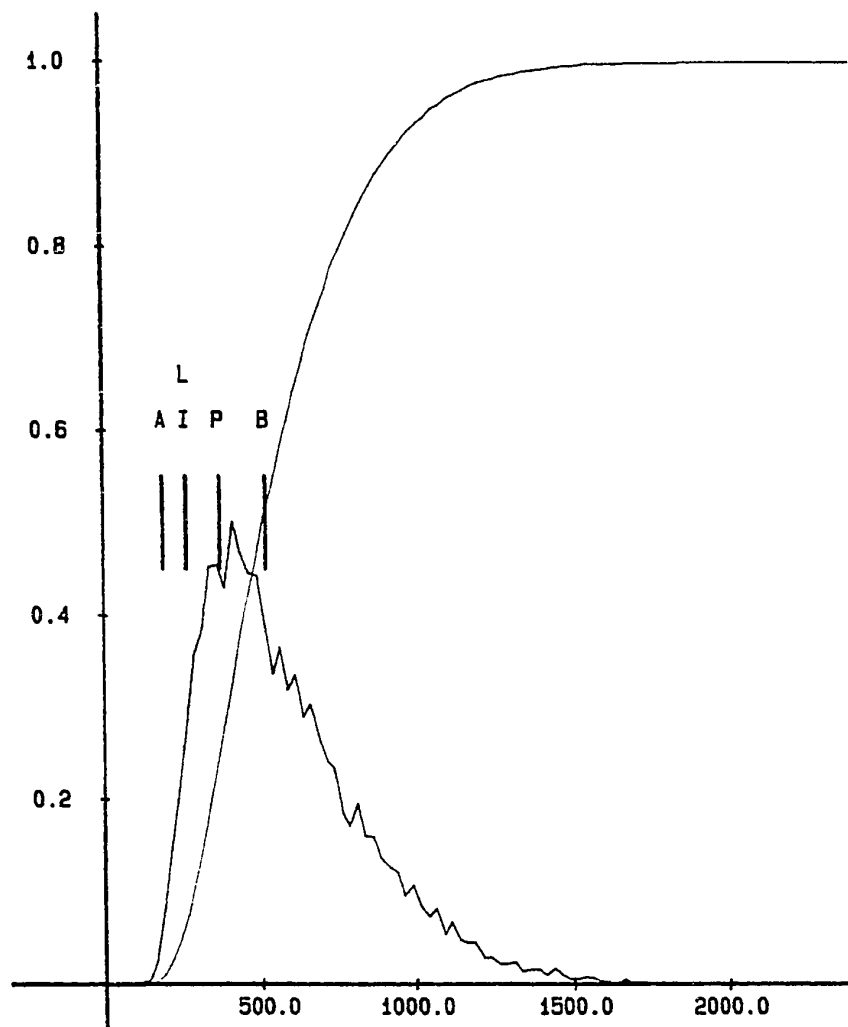


Figure 39. Relative frequency {rescaled} and cumulative relative frequency for the worst computed loss in the Monte Carlo simulation, comparing the models.

model could be significantly above that expected by chance without the rate of return achieving a significantly higher level. This could occur if the model were successful in calling the direction of change when the amount of change was small, but calling it incorrectly when the amount was large. Similar, but more subtle, biases could affect the net profit and profit/loss ratio measures. It is shown in Appendix B that two other possible sources of bias, transaction costs and differential risk, do not alter the conclusion.

Table XII summarizes the hypothesis tests. Because the Monte Carlo simulation produced distributions with slightly larger means, fatter tails and larger standard deviations than expected under the normality assumption, statistics from the simulation were used to test the hypotheses. This is conservative in that it slightly reduces the probability of rejecting the null hypothesis. Also shown in the table are the number of simulation trials that produced results less than that of the rule being tested. All null hypotheses are rejected at high confidence levels.

An evaluation of forecasting methods on the basis of historical data must be viewed with merciless skepticism since there are many opportunities for faulty reasoning. Chicanery need not be involved as the errors may be unconscious. "A universal human problem," says Watzlowick

TABLE XII

HYPOTHESIS TEST STATISTICS

H_0 : Net Profit = 3.85 vs. H_1 : Net Profit > 3.85

Rule or Model	Net Profit	Number of Simulations Less Than	Standard Deviations from Mean	Confidence Level
Bayesian	389	7724	.75	.774
Investment	1195	9900	2.32	.990
Price-Investment	1219	9912	2.37	.991
Labor	1193	9899	2.32	.990
Atheoretical	1477	9980	2.87	.998

H_0 : Number of Hits = 82.57 vs. H_1 : Number of Hits > 82.57

Rule or Model	Number of Hits	Number of Simulations Less Than	Standard Deviations from Mean	Confidence Level
Bayesian	93	9374	1.61	.946
Investment	101	9973	2.84	.998
Price-Investment	101	9973	2.84	.998
Labor	95	9665	1.92	.972
Atheoretical	102	9981	3.00	.998

H_0 : Rate of Return = .621 vs. H_1 : Rate of Return > .621

Rule or Model	Rate of Return	Number of Simulations Less Than	Standard Deviations from Mean	Confidence Level
Bayesian	.690	7073	.55	.710
Investment	.943	9957	2.59	.995
Price-Investment	.957	9969	2.70	.996
Labor	.932	9943	2.50	.994
Atheoretical	1.024	9992	3.24	.999

(1976), is that

once we have arrived at a solution--and in the process of getting there have paid a fairly high price ... --our investment in this solution becomes so great that we may prefer to distort reality to fit our solution rather than to sacrifice the solution.

The L-Method Test.

This study has avoided the obvious trap of testing on the training data, i.e., directly evaluating the classifiers on the basis of the data used to develop them. There is, however, a related but more subtle form of data mining. Sometimes called training on the testing data, "it arises when a classifier undergoes a series of refinements guided by the repeated testing on the same test data." (Duda and Hart, 1973)

In principle, the solution to either problem is simple: obtain a new set of testing data. But when the samples are limited, as here, there is a real dilemma. If data are reserved for testing, the classifier will not be well designed. If all data are used for design, one cannot have confidence about the classifier's performance on an independent set. Lachenbruch and Mickey (1968) and Fukunaga and Kessell (1971) provide an ingenious escape, called the method of leaving one out or simply the L method. This method gives an almost unbiased estimate of the performance of the classifier while utilizing the available data as effectively as possible.

The method involves developing not one classifier but n of them (where n is the number of data points in the sample). The first classifier is designed in the usual way using a subset of $n-1$ data points. It is then tested to see whether it correctly classifies the one omitted data point. Similarly, the second classifier is designed using a different subset of $n-1$ points and tested on the one left out. This process continues until all n distinct subsets of $n-1$ points have been used in the design process. The aggregate result of the n one-at-a-time classifications of the points left out closely approximates the performance to be expected of the final classifier on an independent sample from the same population.

The method utilizes the data set effectively but at the considerable cost of designing n different classifiers. Its major advantage, however, is psychological. It gives the designer considerable confidence that he did not "distort reality to fit his solution."

The L method was used to estimate the performance of the models developed here, with the results shown in Table XIII.

TABLE XIII

COMPARISON OF MODEL PERFORMANCE WITH THE L-METHOD TEST

Model	Actual Performance		L-Method Test	
	Net Profit	Hits	Net Profit	Hits
Investment	1195	101	1219	98
Price-Investment	1219	101	853	97
Labor	1193	95	1099	99
Atheoretical	1477	102	1095	97

That the L-method test ever shows superior performance is at first surprising but is explained by recalling that the classifiers actually used for forecasting were predicated only on the data available prior to making the forecast. The classifier used to make the first forecast, for example, was based on only twelve data points. The L-method test, on the other hand, is based on all the points (with, of course, one left out).

The results of the L-method test are comforting, but one should exercise caution before assuming that the future performance of the models would be as favorable. Even if the economy were statistically stable, the results above are only estimates and, like all statistical estimates, are surrounded by a region of uncertainty (favorable and unfavorable). Another important caveat is that mutations or shocks to the economy could render the models unreliable. Finally, there is the question (considered more fully below) whether the success of the models might lead

to their widespread use, thereby rendering them ineffective.

ANALYSIS AND DISCUSSION

Two prominent spikes are apparent in all of the graphs of the cumulative net profit: one spike appears in early 1980, the other in late 1981. None of the models were able to forecast well during these periods. Why?

The 1980 failure is easy to explain. It was caused by a major discontinuity or shock to the monetary system. On March 14, 1980, President Carter announced a series of unprecedented monetary and credit controls. Responding to the president's action the

Board of Governors imposed reserve requirements and special deposit requirements on certain types of consumer credit and managed liabilities of commercial banks, a surcharge of 3 percent on frequent borrowers from the discount window, a special deposit requirement on money market funds, and a voluntary restraint program for the growth of total loans of commercial banks.... This program was later viewed as having played a greater role than had been anticipated by affecting the demand for credit and the flow of funds between financial institutions (Gilbert and Trebing, 1981).

With the demand for credit falling, interest rates dropped precipitously. But by July all controls had been removed or were scheduled to expire and interest rates resumed their upward course.

The 1981 failure is not so easily explained. Perhaps the most that can be said is that during the

immediately preceding months interest rates had risen rapidly to levels not reached in modern U.S. history (Board of Governors of the Federal Reserve System, 1983). Maybe yields had overshot their equilibrium. Since the models had forecast correctly the earlier rise and benefited thereby, the subsequent loss was less serious than otherwise (unless, of course, the forecasting model were being implemented just then).

The point of this analysis is not to justify or excuse the errors made by the forecasting models, rather it is to emphasize that no forecasting system is likely to perform well all the time, particularly in the presence of unprecedented economic shocks.

The empirical evidence having been presented, it is time to reconsider the theory it rejects, the efficient market theory. Why is it that this study has found highly significant evidence that the market for long-term government bonds is inefficient, whereas many other investigators have concluded differently?

The simple explanation is that the information sets, forecasting methods, and statistical hypotheses used in this research differ from those used by other investigators.

Tests of the efficient market theory must be made with respect to a specified information set which, in turn, must be manipulated some way into a statistically

testable hypothesis. Consequently, by employing different information sets, different manipulation techniques, and different statistical hypotheses, two investigators can reach contradictory conclusions.

Others have found the same type of divergence. Hafer (1983), for example, found that forecasters attempting to estimate changes in the money supply were efficient in using one information set but inefficient in their use of another. Shiller (1979) found inefficiency in the long-term bond market when he used a new type of statistical hypothesis called the "variance bounds" test.

Most reported tests of the efficient market theory have employed more or less "standard" information sets and have tested more or less "standard" statistical hypotheses. Most conclusions, therefore, have been "standard," i.e., markets are efficient.

Indeed, the author's first forecasting efforts only confirmed all this--when the usual methods (Box-Jenkins ARIMA models and ordinary least squares regression, for example) were applied to the usual variables (past interest rate changes and the money supply, for example), the usual conclusion emerged--markets are efficient.

The preceding "standardization" argument only explains why many tests of the efficient market theory agree. To explain what they agree on requires another step. This insight is gained by remembering that there is a

duality between price forecasting and the efficient market theory; they are but different sides of the same coin. The variables chosen by market participants, on the basis of conventional wisdom, to help them forecast future interest rates are just those same obvious ones that the economists frequently pick when they test the efficient market theory. The same point can be made with respect to the methods used to transform the variables and methods used to forecast.

What happens is this: since the limited number of variables and methods being considered are poured over by many market participants, their usefulness as forecasting tools is rapidly diminished. Since "everyone" thinks these variables and methods are important, they become, in the self-defeating nature of things, unimportant. It is not surprising, then, that when economists conduct tests using the same set of variables and methods, they find the markets to be highly efficient. The result of all this is a self-defeating prophecy on the part of the market participants which becomes a self-fulfilling prophecy for investigators who expect to find efficient markets.

The author found evidence to reject the EMT only when he redefined the problem (from forecasting the amount of change in interest rates to forecasting the direction of change) and changed the forecasting technique (from regression to pattern recognition). Obviously, an estimate of the amount of change developed using ordinary least

squares is also an estimate of the direction of change, but it is not immediately clear why the regression estimate should be inferior to an estimate of the direction of change developed using pattern recognition. The superiority of PR to OLS in drawing distinctions is discussed, and demonstrated by example, in Appendix B.

An important question remains: Is it to be expected that if these models were to become widely used in interest rate forecasting, would they, too, begin to fail? Catch 22? The following answer incorporates several of the points made by Felsen (1975a).

The self-defeating-forecast argument applies as much to the models developed here as to any other temporarily successful price forecasting technique: Markets tend to be efficient and tend to eliminate any possibility of excess profits if the methods of generating them become known. Thus if only a few well known institutional investors were to compile superior investment records on the basis of these models, their use might quickly proliferate with a resulting deterioration in performance. Possibly inhibiting this process is the relative complexity of the models (with their weighting and recursive-updating features) and the general lack of familiarity of most investment managers with the approach.

However, even if the current models were to fail, the algorithms might continue to forecast well using different

information sets. In other words, the existing models might cease to work, but other successful ones might be developed using the same technique. The development of new models clearly requires higher levels of skill than the maintenance of existing ones. Even more demanding is the development of new pattern recognition techniques and algorithms. Felsen speculates that by staying on the leading edge of pattern recognition development, one might continue to earn superior returns.

A parallel can be drawn to technological advance: A firm can obtain a business advantage by being the first to develop and exploit a new scientific technique; to maintain the advantage, however, requires the firm to continue to make new advances in the process since other firms will be quick to copy the superior technology. Not many firms are capable of continuous innovation.

COMBINING FORECASTS

It is well established that an appropriate combination of two forecasts prepared using different information sets and/or different forecasting methods will frequently outperform either of the component forecasts. Bates and Granger (1969) have presented several methods for computing the coefficients, k_i , in the formula

$$(67) \quad C = k_1 F_1 + k_2 F_2$$

where C is the composite forecast and F_1 and F_2

are its components. Pearson (1981) has proposed a method for computing k_i in cases when there are more than two forecasts; Bunn (1979) and Reinmuth and Guerts (1979) consider alternative procedures. All of these, however, attend only to forecasts of a continuous variable, not the binary type developed here. Several possible rules for generating a single forecast from binary components are considered below.

It is obvious that if the two or more forecasts agree in predicting an increase or decrease in rates, the problem of choosing among them disappears. When the binary forecasts disagree, the problem is one of deciding which to follow.

One reasonable approach is to choose that model which has demonstrated the best recent performance, measured, say, by a moving sum of net profits. Let the moving sum of net profits (MSNP) be

$$(68) \quad \text{MSNP}_t(j) = \sum_i v_i \text{sgn}(w'x(i))$$

where the summation runs over the most recent j terms and the other symbols are as previously defined. Then the composite forecast at any time t is taken to be the one with the larger MSNP_t (ties for any given month can be broken by recomputing MSNP with a larger value of j , the number of terms in the summation). The initial value of j could be fixed arbitrarily; but after some experience with the models has been accumulated, the value of j could

be chosen as that which would have maximized past profits. Summary results from forecasting by this rule (called the moving sum model) using various values of j are given in Table XIV.

If there were an odd number of forecasts to be combined, a one might follow the decision given by the majority. The results of forecasting by this majority model are shown in Table XV.

If the forecaster were already holding a portfolio of long-term bonds, he might feel it appropriate to hedge this portfolio only when all of the forecasting models agreed that interest rates were about to increase; otherwise he would maintain an unhedged position. The results of this hedge model are given in Table XVI.

A speculator, on the other hand, might wish to expose himself to risk only when all of his models agreed on the future direction of interest rates. This permits a new kind of strategy since up to this point it has always been required that the forecasting models come forth with some estimate of the direction of change, i.e., "no opinion" has not been allowed. With that restriction relaxed, the classifier is not forced to render forecasts when it "thinks" it inappropriate to do so. Clearly this is a more advanced form of behavior and more closely resembles the range of response exhibited by human decisionmakers. Of course, the decision not to make a forecast is not the same

TABLE XIV

SUMMARY STATISTICS FOR THE MOVING-SUM MODELS

Model*			Net Profit	Profit/Loss Ratio	Hits	Worst Computed Loss
I	P	1	1399	1.89	106	265
I	P	2	1531	2.02	106	265
I	P	3	1315	1.81	105	273
I	P	6	1511	2.00	106	265
I	L	1	1185	1.71	101	265
I	L	2	1365	1.86	102	265
I	L	3	1181	1.70	101	265
I	L	6	1139	1.67	99	265
I	A	1	1065	1.61	101	363
I	A	2	1153	1.68	102	363
I	A	3	1137	1.67	102	379
I	A	6	1165	1.69	102	356
P	L	1	1113	1.65	99	360
P	L	2	1073	1.62	98	400
P	L	3	1073	1.62	98	400
P	L	6	1073	1.62	98	400
P	A	1	1475	1.96	105	373
P	A	2	1455	1.94	106	356
P	A	3	1367	1.86	105	400
P	A	6	1449	1.94	107	400
L	A	1	1175	1.70	97	424
L	A	2	1167	1.69	98	424
L	A	3	1169	1.69	100	440
L	A	6	1267	1.77	101	400

*I = Investment Model
P = Price-Investment Model
L = Labor Model
A = Atheoretical Model

1 = 1 Month
2 = 2 Months
3 = 3 Months
6 = 6 Months

TABLE XV

SUMMARY STATISTICS FOR THE MAJORITY-RULE MODELS

Models Included*	Net Profit	Profit/Loss Ratio	Number of Hits	Worst Computed Loss
I P L	1193	1.71	95	265
I P A	1157	1.68	101	391
I L A	1275	1.78	100	273
P L A	1075	1.62	96	356

*I = Investment Model
P = Price-Investment Model
L = Labor Model
A = Atheoretical Model

TABLE XVI

SUMMARY STATISTICS FOR THE HEDGE-RULE MODELS

Models Included*	Net Profit	Profit/Loss Ratio	Number of Hits	Worst Computed Loss
I P	1163	1.69	101	373
I L	1421	1.91	101	265
I A	1547	2.03	104	152
P L	935	1.52	95	373
P A	1581	2.07	104	152
L A	1721	2.22	105	152
I P L	1163	1.69	101	373
I P A	1567	2.05	104	152
I L A	1707	2.20	105	152
P L A	1581	2.07	104	152
I P L A	1567	2.05	104	152

*I = Investment Model
P = Price-Investment Model
L = Labor Model
A = Atheoretical Model

as a forecast of no change in interest rates.

When the decision rule permits the forecaster not to take a market position, the profit/loss ratio may increase even while the sum of the actual profits is decreasing.

One such rule is very simple: If the forecasts of all the models agree, that is the forecast; if they do not agree, there is no forecast. The results of this rule, called the veto model are shown in Table XVII. The profit/loss ratios in that table reflect the fact that these composite models did not make a forecast at every opportunity.

Figures 40 through 53 show the net profit and profit/loss ratio for selected composite models.

When several forecasts are available, the possible strategies become more varied and sophisticated than those explored so far. It turns out, for example, that some models seem relatively more successful in forecasting increases in rates while others do better when forecasting decreases. A composite rule taking these differences into account might give superior performance.

At some point analysis must give way to conclusions and action. This seems like a good point to end the discussion of new decision rules and assemble together the conclusions reached and see what they imply.

TABLE XVII
SUMMARY STATISTICS FOR THE VETO-RULE MODEL

Models Included*		Net Profit	Profit/Loss Ratio	Worst Computed Loss		
I	P	1207	2.03	265		
I	L	1194	1.86	265		
I	A	1336	2.09	152		
P	L	1206	1.82	294		
P	A	1348	2.24	233		
L	A	1335	2.14	152		
I	P	L	1207	2.03	265	
I	P	A	1367	2.53	152	
I	L	A	1295	2.20	152	
P	L	A	1407	2.42	164	
I	P	L	A	1367	2.53	152

*I = Investment Model
P = Price-Investment Model
L = Labor Model
A = Atheoretical Model

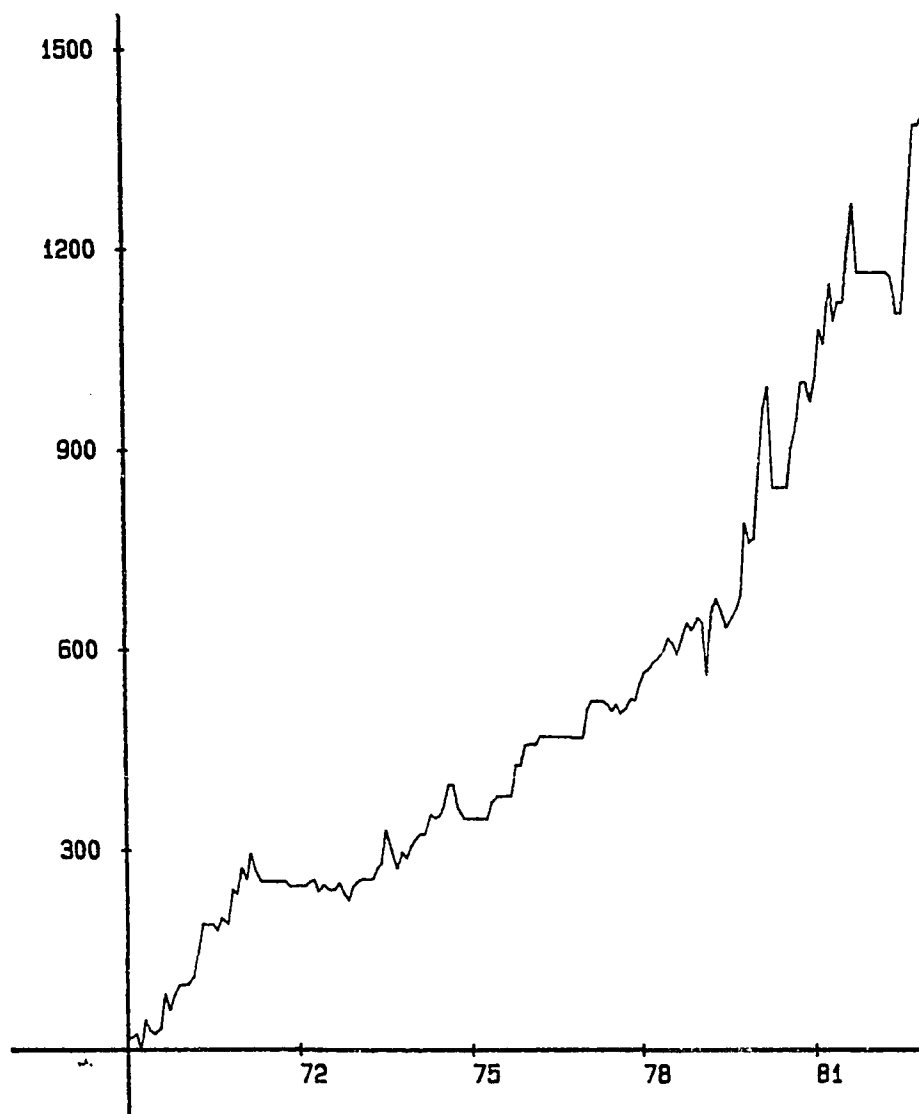


Figure 40. Cumulative net profit--investment model, labor model and atheoretical model combined according to the majority rule.

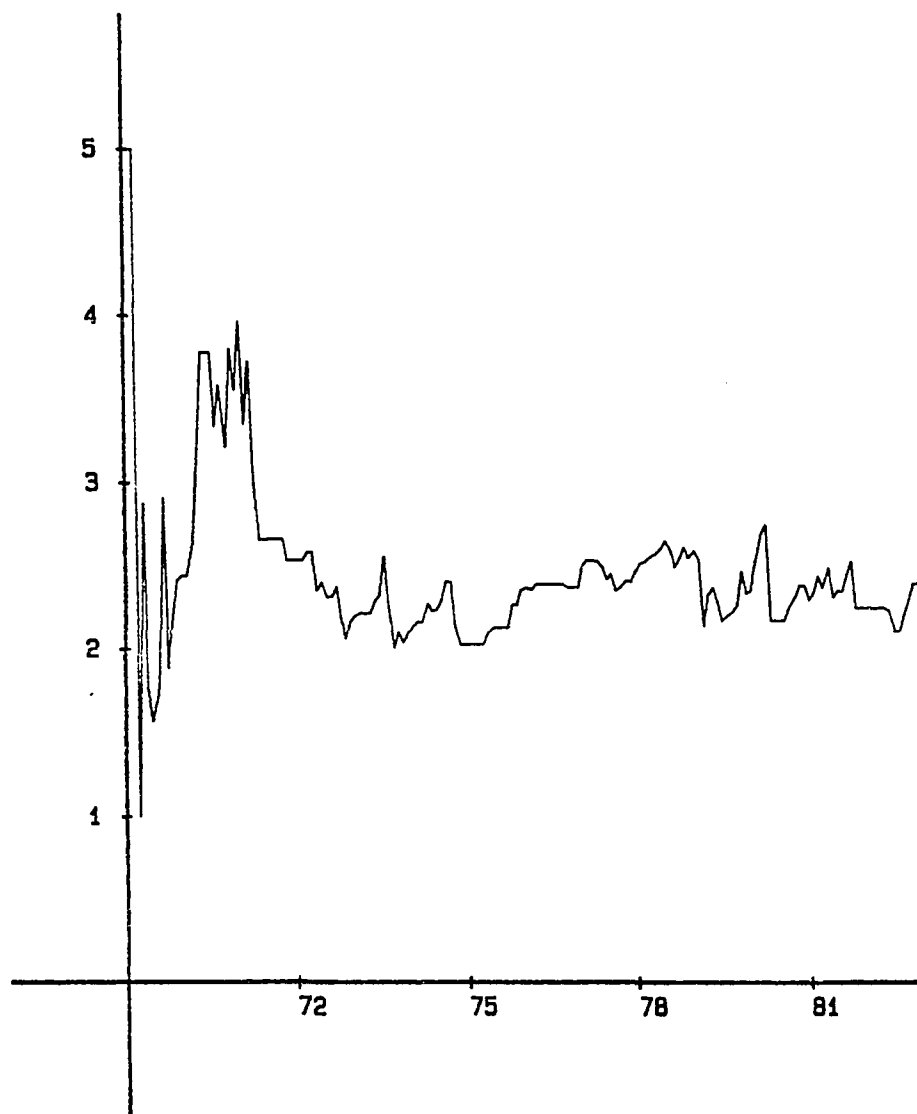


Figure 41. Profit/loss ratio—investment model, labor model and atheoretical model combined according to the majority rule.

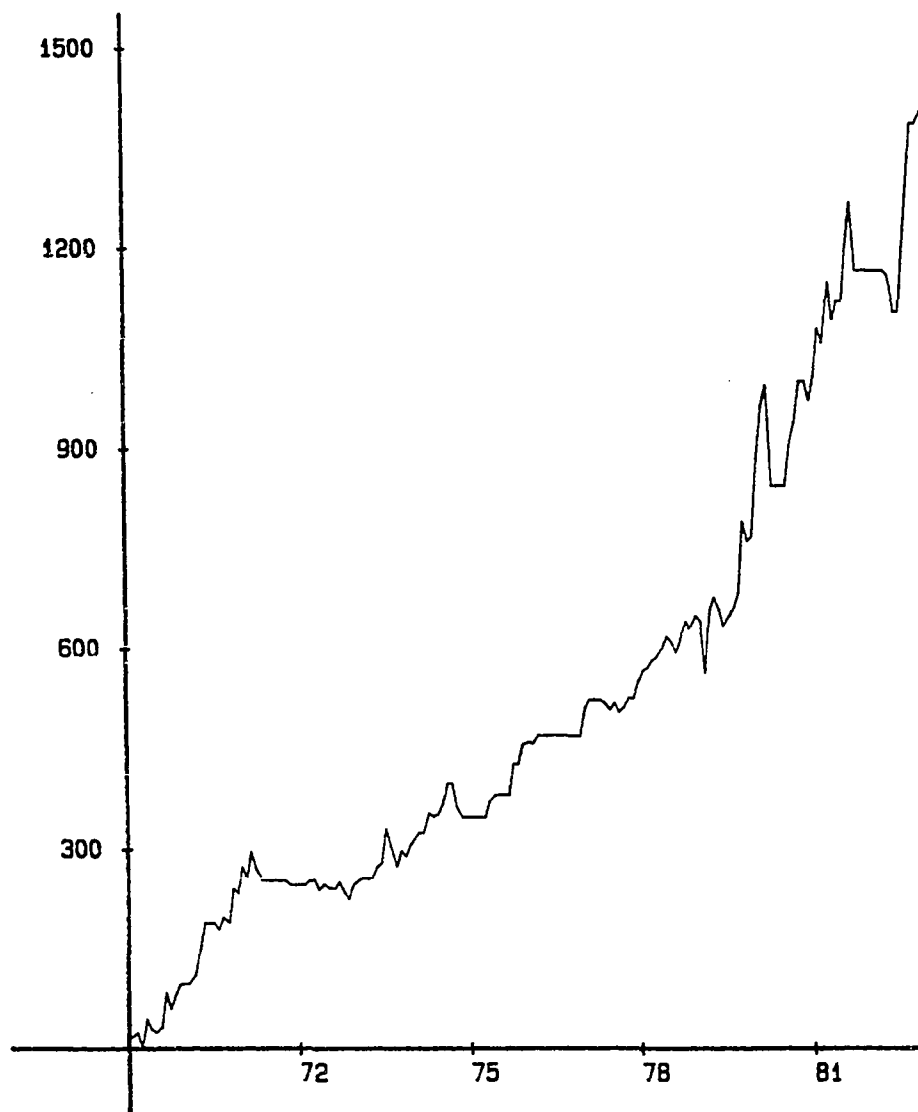


Figure 42. Cumulative net profit--investment model and labor model combined according to the hedge rule.

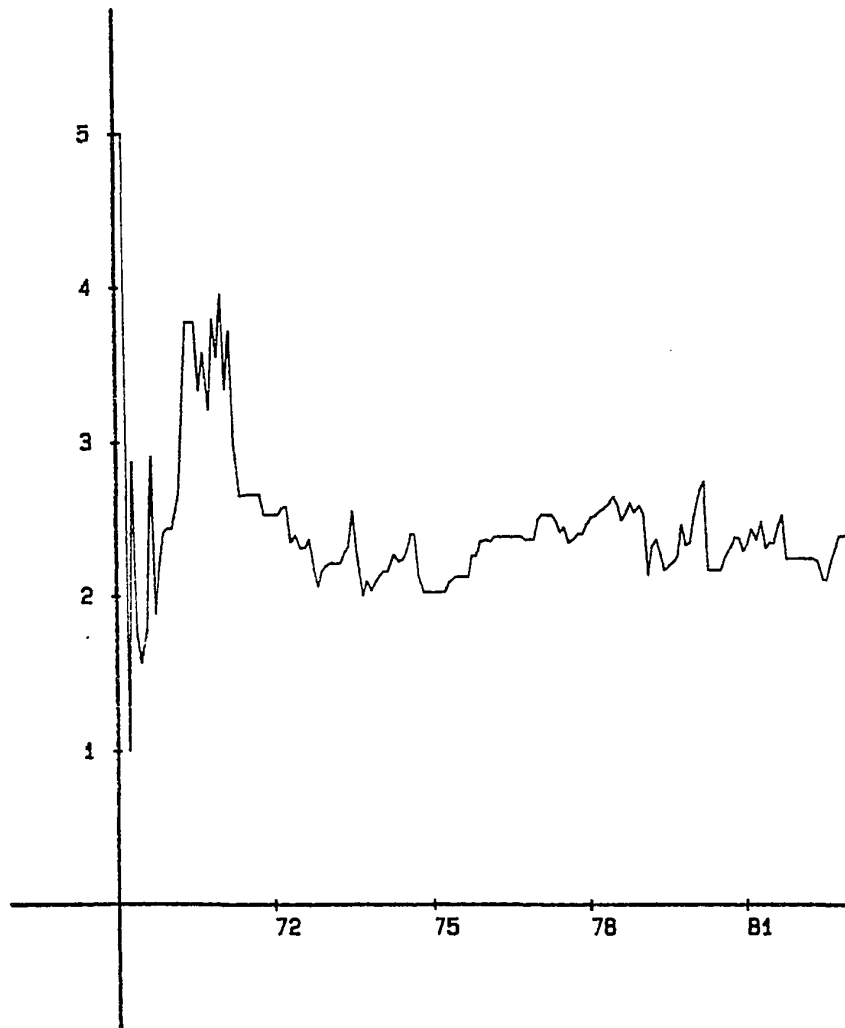


Figure 43. Profit/loss ratio—investment model and labor model combined according to the hedge rule.

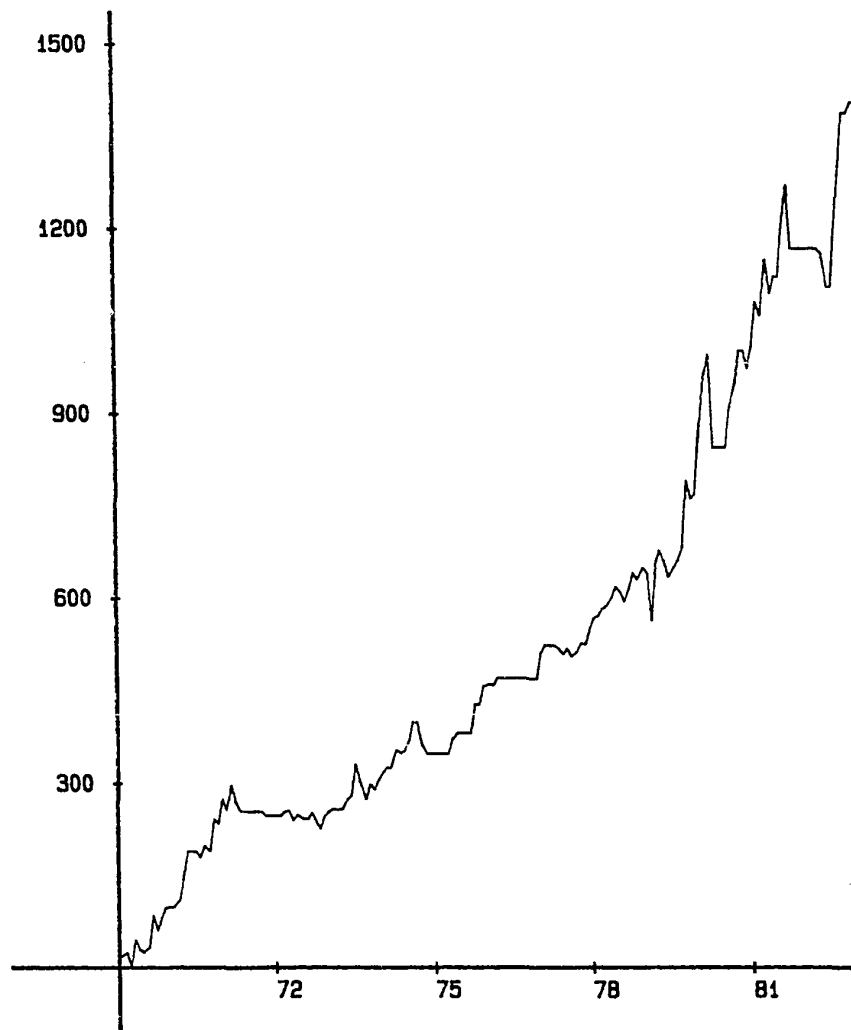


Figure 44. Cumulative net profit—price-investment model and labor model combined according to the hedge rule.

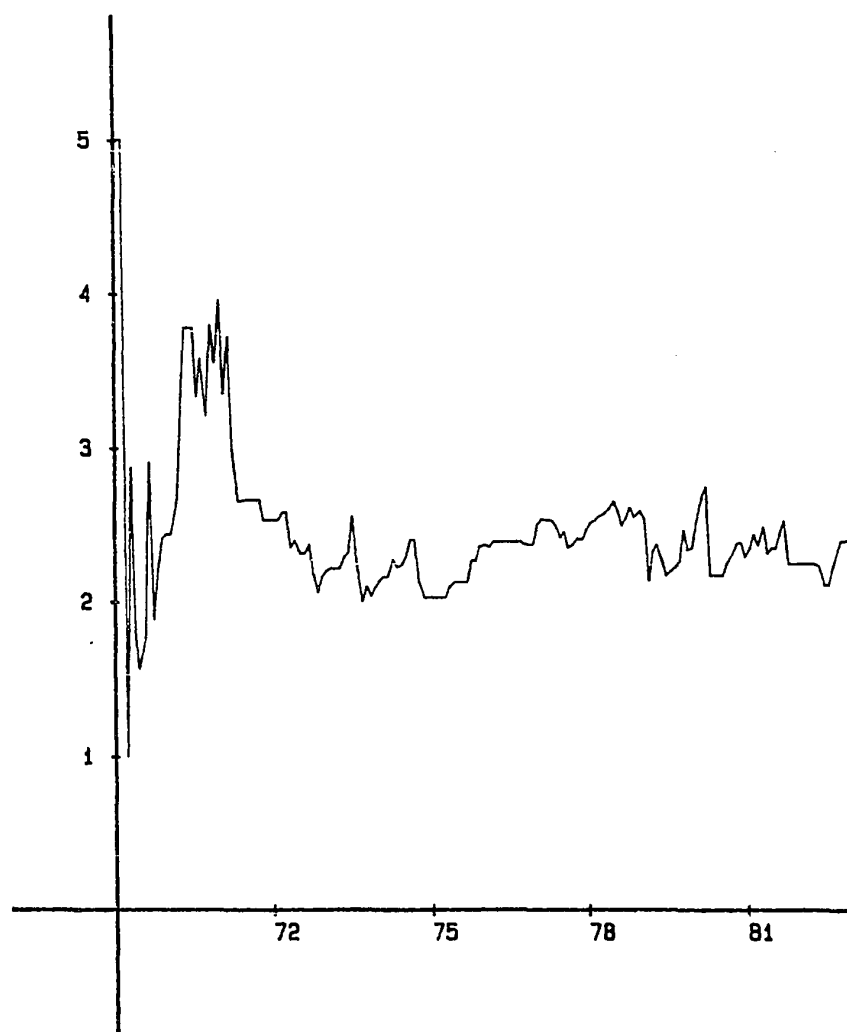


Figure 45. Profit/loss ratio--price-investment model and labor model combined according to the hedge rule.

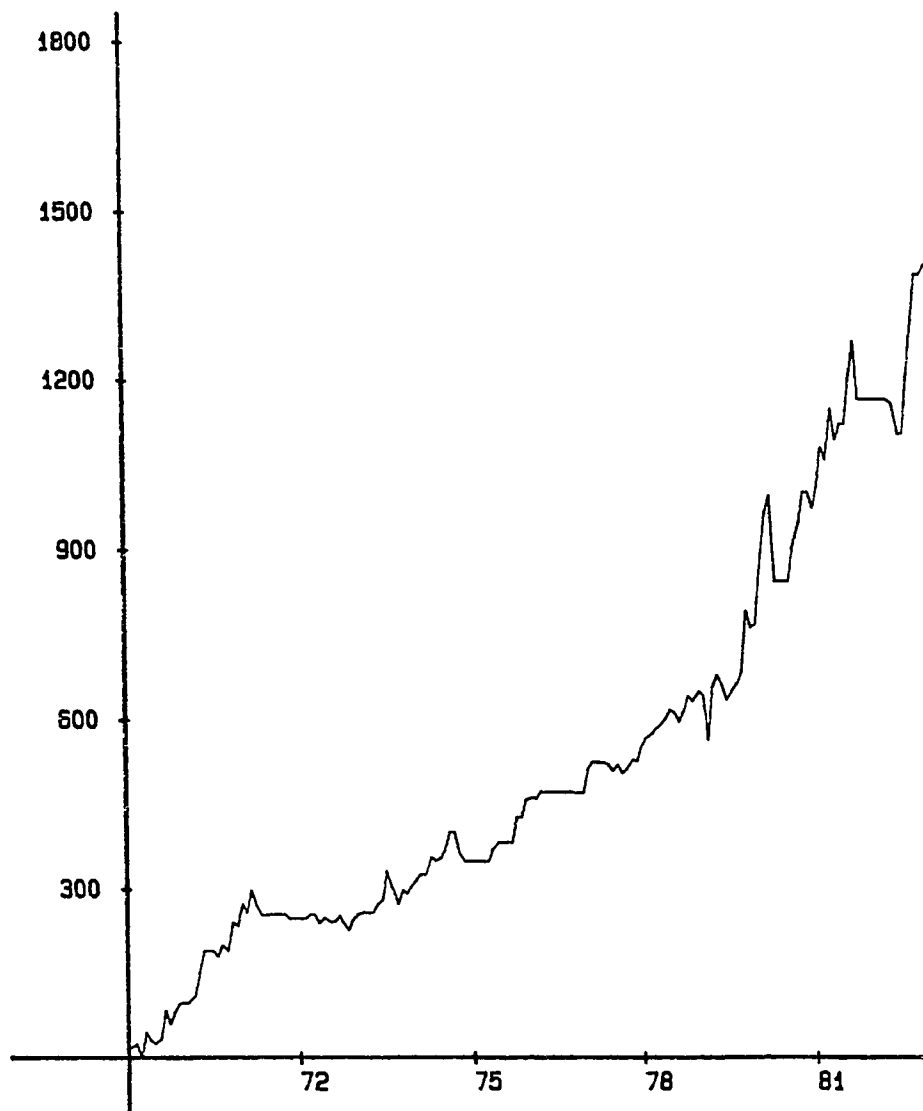


Figure 46. Cumulative net profit--labor model and atheoretical model combined according to the hedge rule.

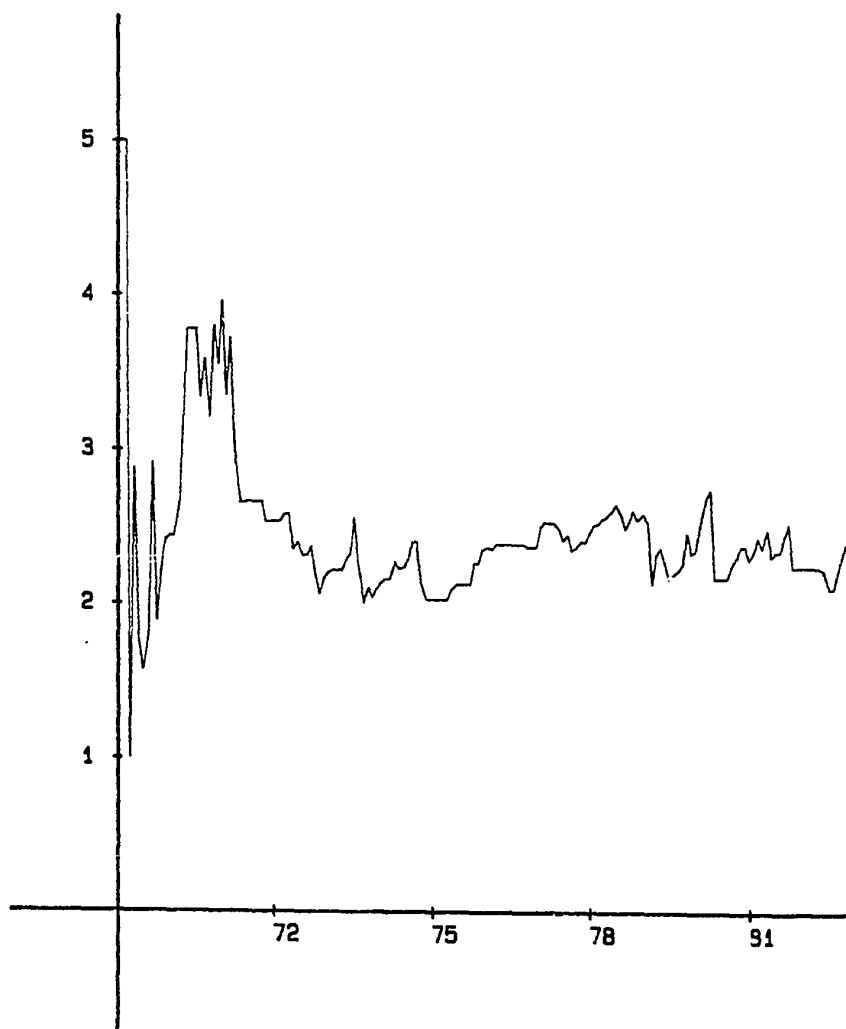


Figure 47. Profit/loss ratio--labor model and atheoretical model combined according to the hedge rule.

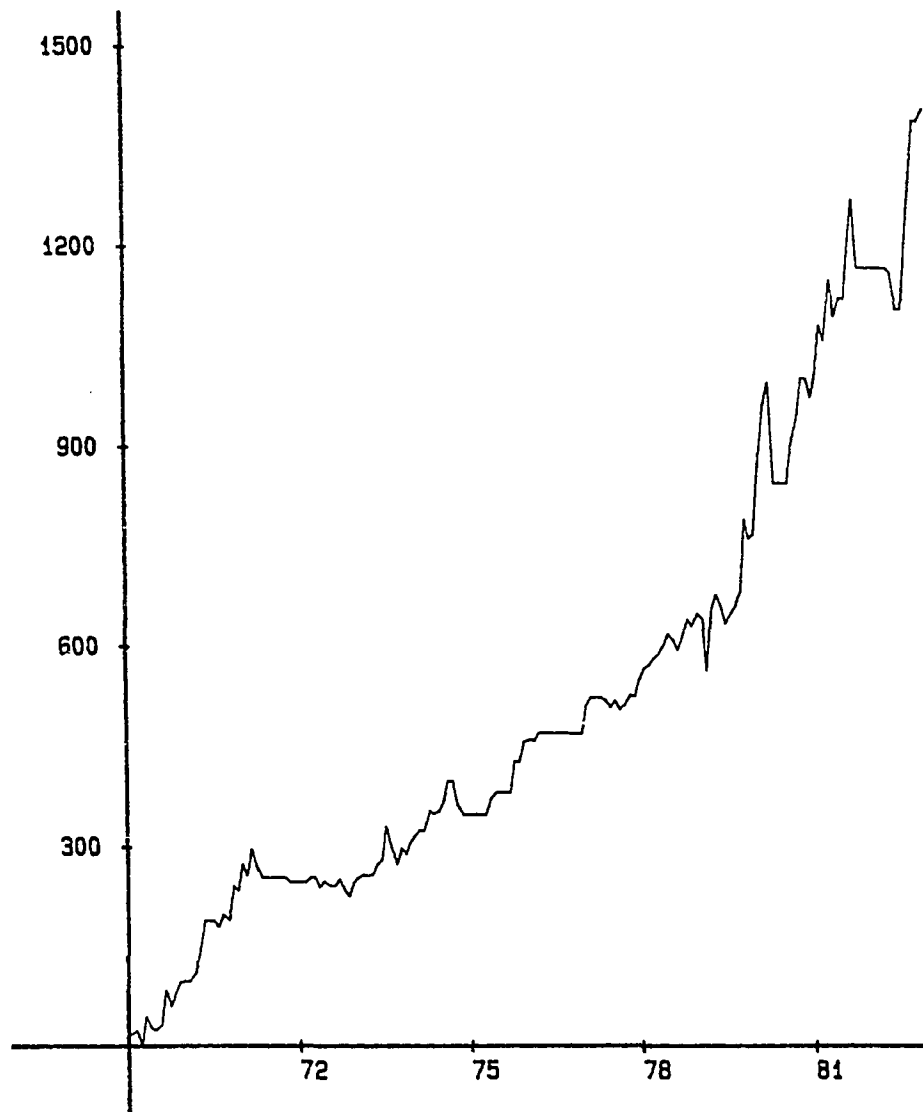


Figure 48. Cumulative net profit--investment model, price-investment model and labor model combined according to the veto rule.

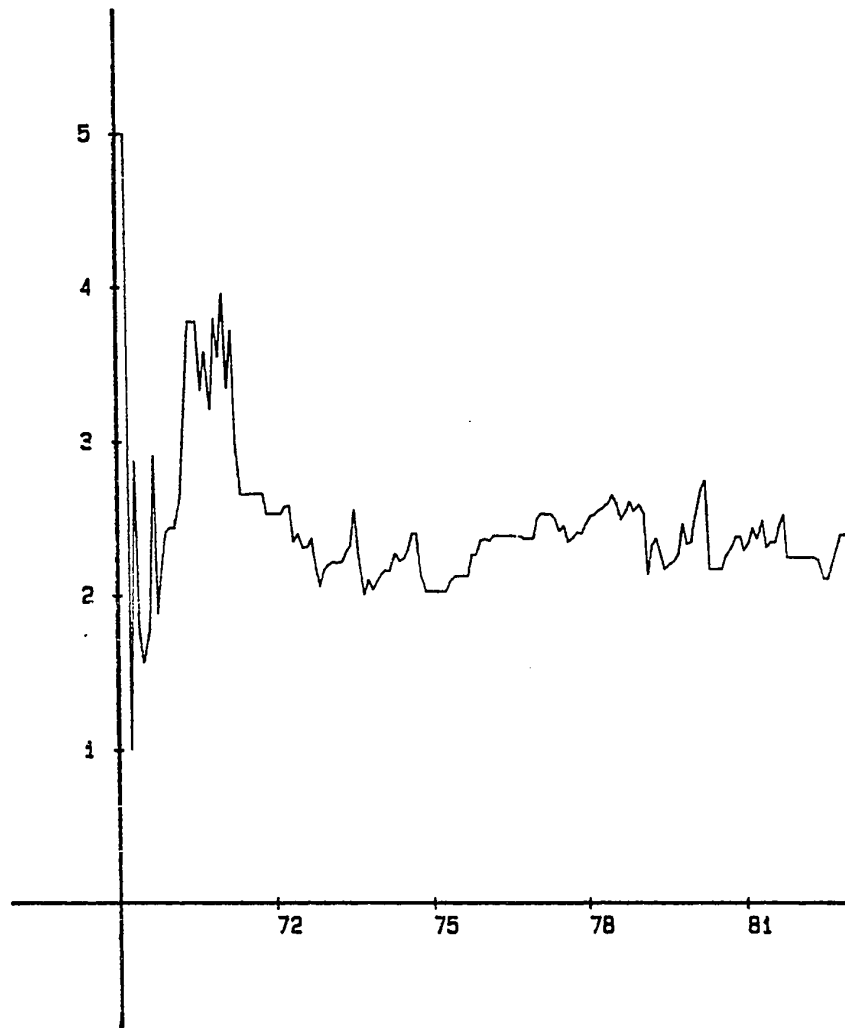


Figure 49. Profit/loss ratio--investment model, price-investment model and labor model combined according to the veto rule.

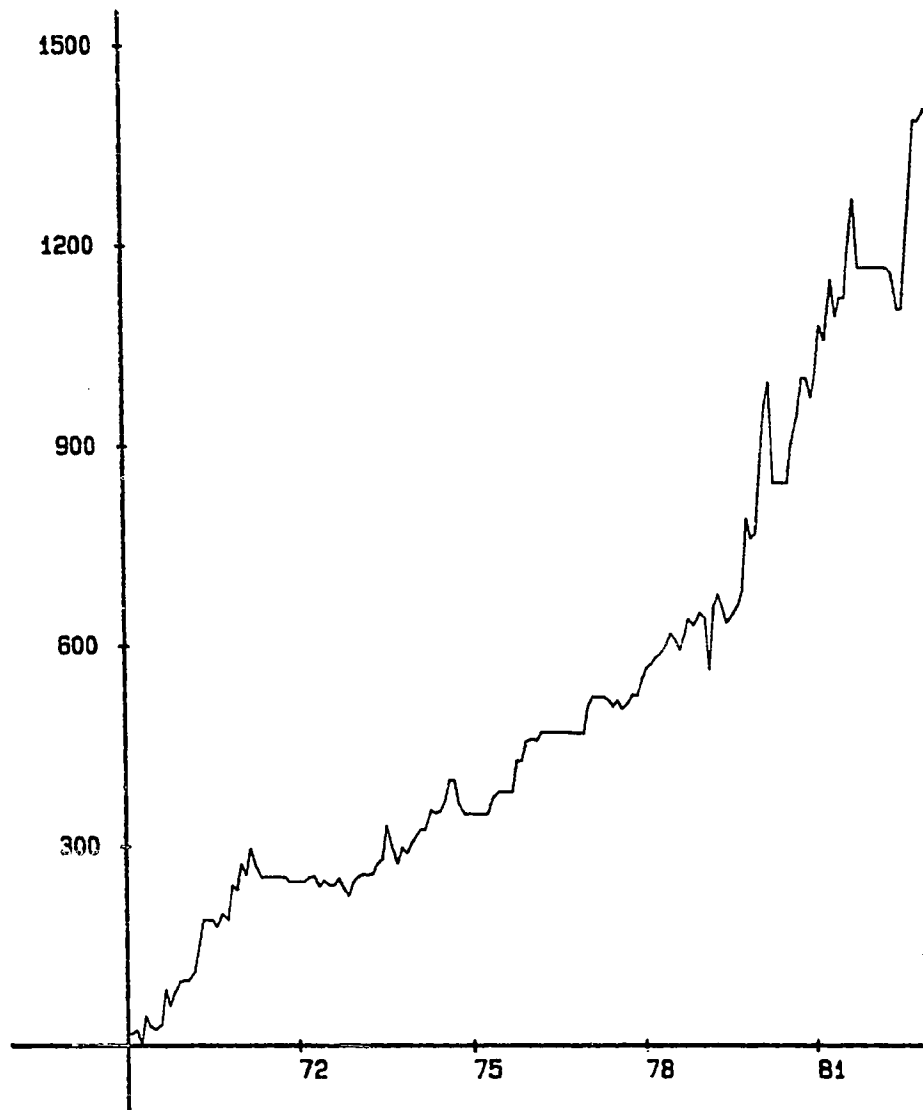


Figure 50. Cumulative net profit--investment model, price-investment model and atheoretical model combined according to the veto rule.

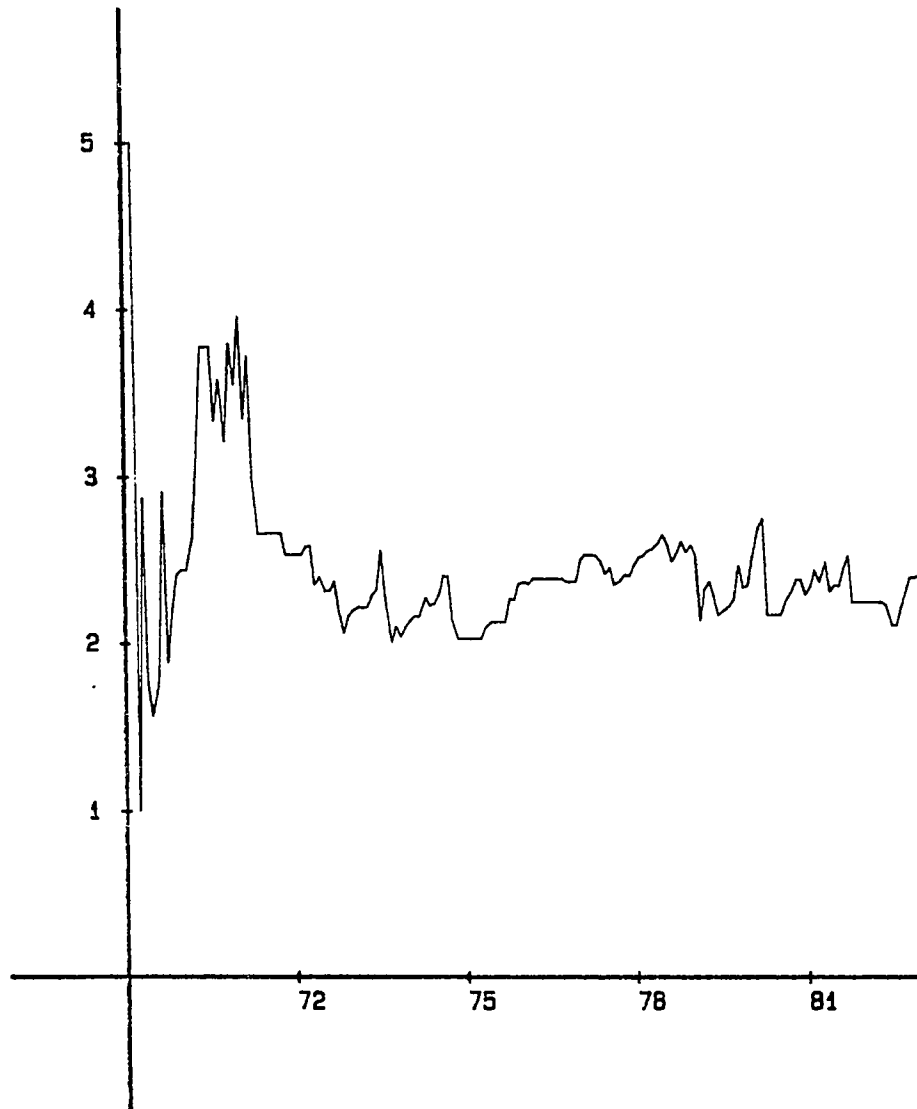


Figure 51. Profit/loss ratio--investment model, price-investment model and atheoretical model combined according to the veto rule.

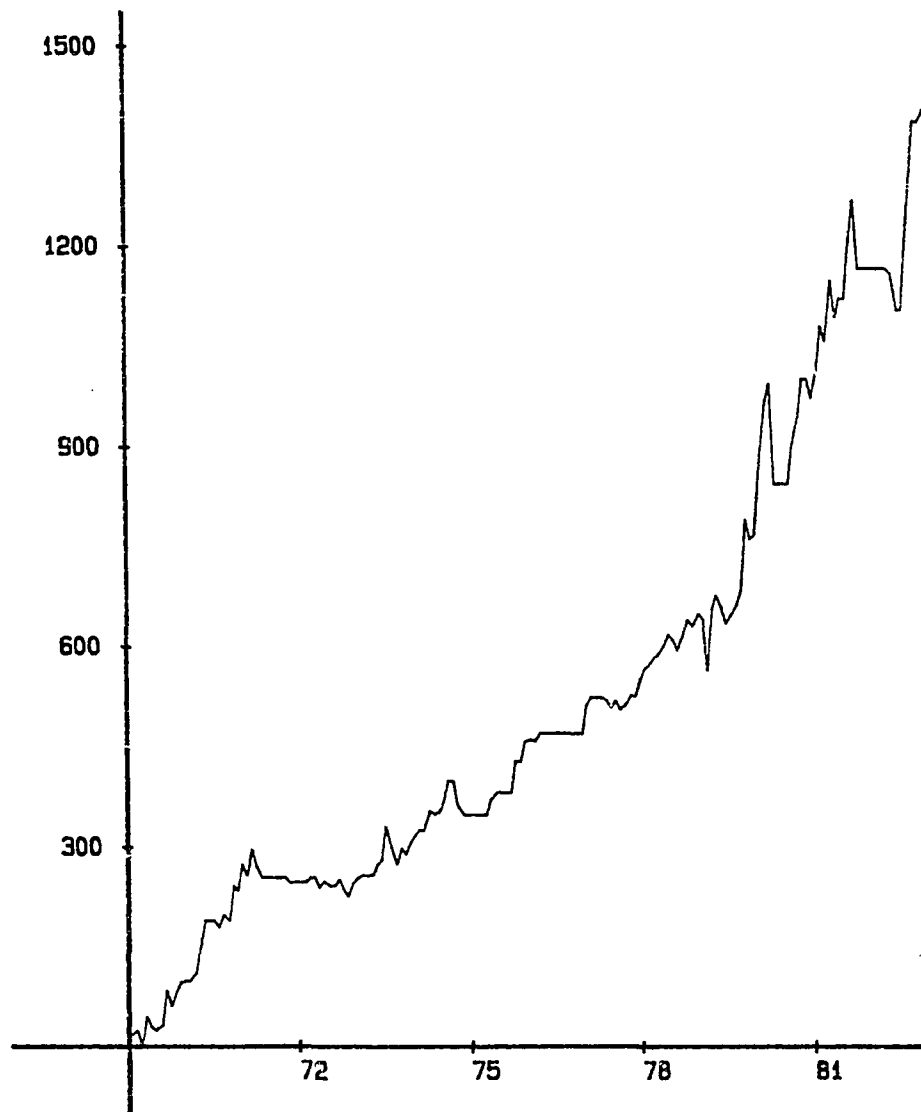


Figure 52. Cumulative net profit--price-investment model, labor model and atheoretical model combined according to the veto rule.

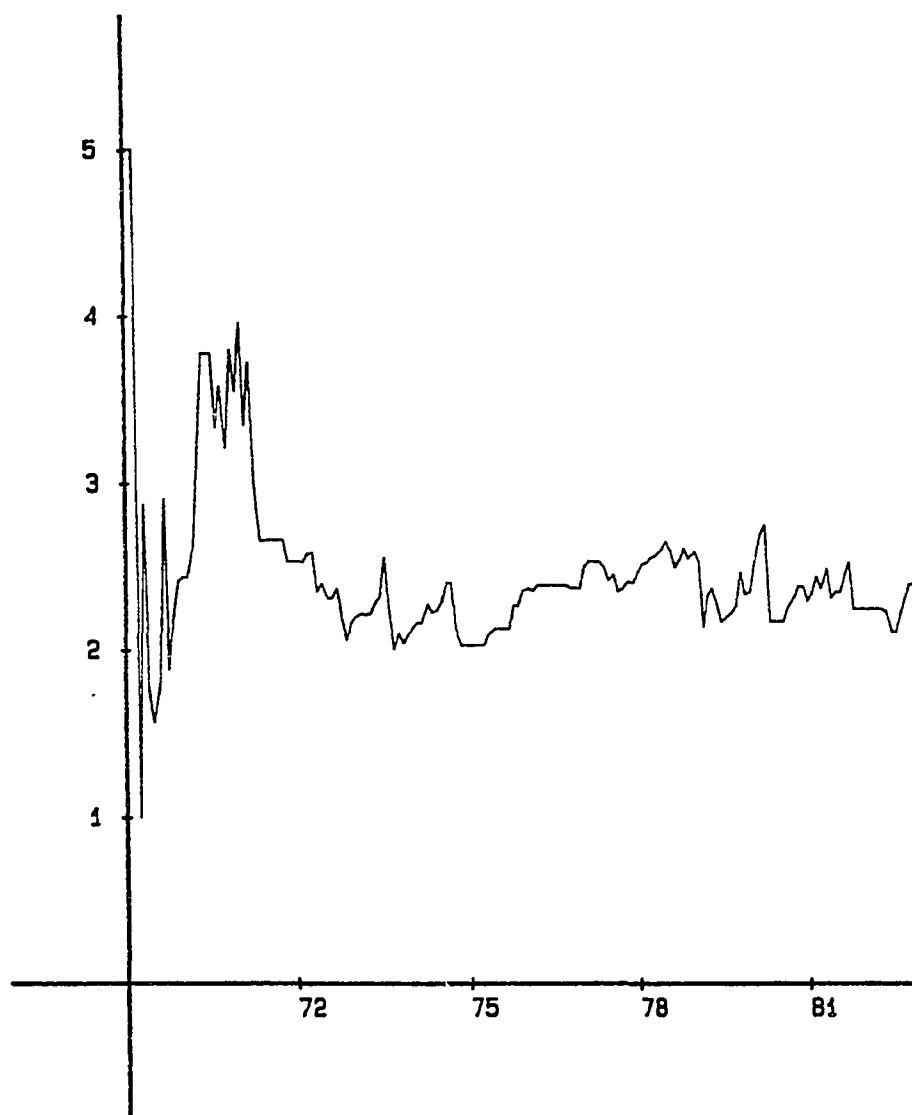


Figure 53. Profit/loss ratio—price-investment model, labor model and atheoretical model combined according to the veto rule.

CHAPTER V

SUMMARY, CONCLUSIONS, IMPLICATIONS, AND DIRECTIONS FOR SUBSEQUENT RESEARCH

This final chapter looks backward and forward: back to summarize what has been done and to gather in one place the major conclusions that have been drawn; ahead to the implications of the work, and beyond to the directions subsequent research might take.

SUMMARY

The introductory chapter explains the motivation for wanting to forecast interest rates, showing the growing importance of interest as an element of national income and describing the usefulness of an interest rate forecast to portfolio managers, corporate borrowers, speculators and policymakers.

Chapter II looks at the interest rate forecasting problem from several perspectives. The major theoretical impediment to forecasting--the efficient market theory--is discussed at length. The nature of the economic environment in which forecasting takes place is considered. Different views of the securities markets and different

approaches to price forecasting are discussed. The chapter concludes with a nontechnical overview of pattern recognition, a new approach to forecasting.

Chapter III starts from basic principles and develops the subject of pattern recognition just far enough to derive the classifiers needed to forecast the direction of change in interest rates. The important results are recursively updated weighted-least-squared-error algorithm and the hyperbolic tangent algorithm.

Chapter IV reports the substance of the research project. First, it provides unambiguous notions of what a forecast is and which interest rate is to be forecast. It then proposes a reasonable forecasting objective and suggests standards by which forecasting success can be judged. A brief discussion of the economic theory underlying the choice of variables is followed by the development of a method of detrending those variables so as to better capture information useful in forecasting interest rates. Four specific forecasting models are then presented and their performance is evaluated in comparison with a random walk simulation and with a Bayesian alternative. These forecasts are then combined in several different ways to yield composites which are superior in performance to their components.

CONCLUSIONS

The major conclusions emerging from this research are these.

The nation's economic system is complex and evolving and it sometimes exhibits irrational behavior. In such an environment, forecasting interest rates is not likely to be an easy task.

In most fields hard work combined with creativity and a superior technology usually outperforms lesser effort combined with poorer technology. Interest rate forecasting is not an exception.

Pattern recognition is an interdisciplinary endeavor and represents a relatively new and potentially powerful intellectual technology. When systems are statistically unstable, pattern recognition appears to offer a superior forecasting method. This is so because the algorithms permit the forecasting system to learn from its errors and adapt to changing conditions.

The potential for self-delusion in economic forecasting is great. A clear understanding of the available information set avoids "forecasting the past." The method called "leaving-one-out" protects against the trap of data mining.

The efficient market theory, while probably useful to economists as a first approximation, suffers serious

problems. It is contradicted by findings from fields as varied as physiology and sociology. More important, the empirical economic evidence presented here requires one to reject the theory at the highest levels of confidence. Since the market for long-term U.S. government bonds is one of the broadest, deepest and best informed in the world, failure to find confirmation here seems fatal to the theory in its "operational form."

The reason this study rejects the efficient market theory, while some others have not, is the use here of different information sets, different forecasting techniques, and different statistical hypotheses.

Four simple models indicate that one should be able to forecast the direction of change in interest rates correctly about 60% of the time. If the results are dollar weighted and the forecaster is allowed more sophisticated strategies (including the right to refuse to forecast when uncertainty is greatest), then performance exceeding 70% may be possible. These percentages compare favorably with the 45% historically achieved by a buy-and-hold strategy, the 50% expected from a random decision rule, or the 55% that the Bayesian rule would have scored.

Another conclusion, one supported by the whole of the research rather than just the portion described here, is that the greatest profit/loss ratios are achieved by combining many individual models together. For risk-adverse

speculators, variations on the veto rule will prove most attractive; for those who must always have an opinion on the future course on interest rates, variations on the majority rule or the hedge rule are the methods of choice.

IMPLICATIONS

There are implications from this research for each of the audiences identified in the first chapter.

Portfolio managers who use the forecasting methods developed here can expect to make significant improvements in their investment performance. The case of a financial institution with a hypothetical \$1 million fixed-income portfolio that was always invested in twenty-year constant-maturity U.S. government bonds is considered in Appendix D. The appendix shows that over the fourteen years covered by this research, that portfolio would have yielded approximately \$1.2 million in interest but would have incurred capital losses of nearly \$300,000, for a net gain of about \$900,000. On the other hand, had the lender followed a hedging strategy indicated by any of the forecasting models developed here, the interest income would have been about the same, but the capital loss would have been replaced by a profit of more than \$350,000; so the net gain would have been \$1.55 million, about 70% greater than that generated by the naive strategy. This represents a major increase in total investment return.

It is easy for the owners of financial assets to buy and sell them in securities markets or to hedge them in the futures markets. It is much more difficult and expensive for corporations to issue and redeem their debt instruments. Nevertheless, those corporations that frequently raise large amounts of borrowed capital, utilities for example, could expect to make meaningful reductions in their interest costs if they were to time their bond offerings in accordance with the forecasting models.

There are two lessons here for speculators. The most important is that it may possible to make large profits following the buy- and sell-signals generated by the pattern recognition techniques. But the relatively large amount of the worst computed loss should lead speculators to understand that important losses are probably inevitable from time to time, so careful management of speculative capital is essential if ruin is to be avoided. In other words, since a "fail-safe" strategy is probably impossible to achieve, speculators should adopt a strategy that is "safe-fail" (Holling, 1976, 1977).

Economic policymakers, more than most lenders, borrowers, or speculators, need point estimates of the economic variables they follow. Moreover, they want their models to make predictions conditional on certain policy action being implemented. For these reasons pattern

recognition models that give unconditional direction-of-change forecasts do not really conform to the needs of economic policymakers. The forecasts may still be useful in reducing uncertainty and may lead to better policy actions.

Economic theorists need to reconsider the efficient market theory in view of the evidence--psychological and sociological as well as economic--presented here. If markets are seriously inefficient, then the whole structure of those economic theories built around the concept of rational expectations may be faulty. Modifying these theories will be difficult, says Lester Thurow (1983), because "it involves abandoning a beautiful sailing ship--the price-auction model--that happens to be torn apart and sinking in a riptide" of contrary evidence. Teweles, Harlow, and Stone (1974) use a different metaphor: "The greatest tragedy in all history is the murder of a beautiful theory by a gang of brutal facts."

The demonstrated effectiveness of pattern recognition as a forecasting method should encourage its use by other economic forecasters, particularly those concerned with the so-called turning point problem.

All should be aware that economic shocks, mutations of the economy, or the self-defeating-forecast phenomenon could render the models developed here less effective in the future.

DIRECTIONS FOR SUBSEQUENT RESEARCH

This work is not the last word on the application of pattern recognition techniques to interest rate forecasting; it is only the first chapter. Much remains that could be done.

The range of variables entering the pattern classifier could be extended to include other economic series and combinations and transformations thereof. It might be more fruitful, however, to include variables suggested by some of the other perspectives discussed in Chapter II (this is not, of course, equivalent to adopting those other perspectives). The variable being forecast could also be changed. For example, different forecasting horizons might be considered, or the forecast might relate to different financial instruments, say Treasury bills or municipal bonds.

Other criterion functions would be more appropriate in circumstances where the utility function of the decision maker was neither linear nor symmetric. Different numerical methods for optimizing the criterion functions could be considered. For example, it seems likely that the criterion function J_L might be expressed and solved as a linear programming problem, albeit one involving an extremely large simplex matrix.

Preliminary computations indicate that it may be

possible to exploit the fact that some classification models excel when forecasting rate increases while other models perform better when forecasting decreases. This suggests yet other methods of combining forecasts.

Graphic presentations can reveal much in a data structure that would otherwise go undetected. Competent researchers in statistics (Anscombe, 1973) and pattern recognition (Friedman and Tukey, 1974; Stanley, Nienow, and Lendaris, 1969) have discredited the idea that "performing intricate calculations is virtuous, whereas actually looking at the data is cheating" (Anscombe). When the graphs are on a computer screen under user control, their power is multiplied. Future research should certainly take full advantage of whatever graphics capabilities are available.

Nonlinear classifiers frequently perform better than linear ones. With the aid of interactive computer graphics, it may be feasible to tackle the greater complexity these nonlinear decision functions entail.

The adoption of new intellectual tools depends on their ready availability and ease of use. When that use requires computations that only a computer can feasibly perform, "packages" are needed. There seems to be a need for a well-thought-out set of interactive computer programs for use in pattern recognition. A serious difficulty, of course, is that many problems require that

the algorithms be specially designed, rather than called from a standard set of routines.

All of these advances might make it easier to answer "the most crucial question facing the world economy: Where will interest rates go?"

REFERENCES

- Allison, G. (1971). *Essence of Decision: Explaining the Cuban Missile Crisis*. Boston: Little, Brown.
- American Council of Life Insurance (1983). *Life Insurance Fact Book*. Washington.
- Andersen, D. (1977). *Mathematical Models and Decision Making in Bureaucracies: A Case Study Told from Three Points of View*. Ph.D. Dissertation. Massachusetts Institute of Technology.
- Anderson, O.D. (ed.)(1979). *Forecasting*. Amsterdam: North-Holland.
- Anscombe, F. (1973). *Graphs in statistical analysis*. *The American Statistician* 27 (February) 17-21.
- _____. (1981). *Computing in Statistical Science through APL*. New York: Springer-Verlag.
- Aronson, D. (1980). *Cybernetics: the computer trading approach of the 1980s*. *Commodities* (January) 38-43, 69.
- Ashby, W.R. (1964). *An Introduction to Cybernetics*. London: Methuen.
- Balasko, Y. (1978). *The behavior of economic equilibria: a catastrophe theory approach*. *Behavioral Science* 23 (September) 375-82.
- Barnes, R. (1979). *Taming the Pits: A Technical Approach to Commodity Trading*. New York: John Wiley.
- Baruch, B. (1980). *Quoted in The Battle for Financial Security* by R. Bridwell. New York: Times Books.
- Batchelor, B. (1978). *Pattern Recognition*. New York: Plenum Press.

- Bates, J. and Granger, C.W.J. (1969). The combination of forecasts. *Operational Research Quarterly* 20 (December) 451-68.
- Bateson, G. (1972). *Steps to an Ecology of Mind*. New York: Ballentine Books.
- Beishon, J. and Peters, G. (ed.) (1976). *Systems Behavior*. London: Harper & Row.
- Bellman, R. (1957). *Dynamic Programming*. Princeton: Princeton University Press.
- _____ (1977). Large systems. In *Futures Research: New Directions* edited by H.A. Linstone and W.H.C. Simmonds. Reading: Addison-Wesley. 100-3.
- Berger, J. (1980). *Statistical Decision Theory*. New York: Springer-Verlag.
- Berger, P. (1966). *The Social Construction of Reality*. New York: Doubleday.
- Bernstein, J. (1982). *The Handbook of Commodity Cycles: A Window on Time*. New York: John Wiley.
- Beveridge, G. and Schechter, R. (1970). *Optimization: Theory and Practice*. New York: McGraw-Hill.
- Blin, J. (1973). *Patterns and Configurations in Economic Science*. Dordrecht, Holland: Reidel.
- Board of Governors of the Federal Reserve System (1981). *Seasonal Adjustment of the Monetary Aggregates: Report of the Committee of Experts on Seasonal Adjustment Techniques*. Washington.
- _____ (various dates). *Federal Reserve Statistical Release G.13*. Washington.
- _____ (various dates). *Federal Reserve Statistical Release H.15*. Washington.
- _____ (1983). *1983 Historical Chart Book*. Washington.
- Boehm-Bawerk, E. (1922). *Capital and Interest*. Translated by W. Smart. London: Macmillian.
- _____ (1923). *The Positive Theory of Capital*. Translated by W. Smart. New York: G.E. Stechert.

- Bonini, C. and Freeland, J. (1979). Forecasting by smoothed regression: development and application to predicting customer utility bills. In Forecasting edited by S. Makridakis and S. Wheelwright. Amsterdam: North-Holland. 279-96.
- Bowsher, N. (1980). Rise and fall of interest rates. Review of the Federal Reserve Bank of St. Louis (August-September) 16-23.
- Box, G. and Jenkins, G. (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
- Branch, B. (1978). Testing the unbiased expectations theory of interest rates. The Financial Review 13 (Fall) 51-66.
- Breuer, J. (1958). Introduction to the Theory of Sets. Translated by H. Fehr. Englewood Cliffs: Prentice Hall.
- Bridwell, R. (1980). The Battle for Financial Security. New York: Times Books.
- Brown, B. and Maital, S. (1981). What do economists know?: An empirical study of experts' expectations. Econometrica 49 (March) 491-504.
- Brown, G. Spencer (1957) Probability and Scientific Inference. New York: Logmans, Green.
- Bunn, D. (1979). Composition of estimates for decision making. Technological Forecasting and Social Change 13 (June) 157-67.
- Business Week (1983). The outlook for interest rates. (September 26) 136-41.
- Cacoullos, T (ed.) (1973). Discriminant Analysis and Applications. New York: Academic Press.
- Cagan, P. (1972). The recent cyclical movements of interest rates in historical perspective. Business Economics 7 (January) 43-52.
- Carlozzi, N. (1983). Exchange rate volatility: is intervention the answer? Business Review of the Federal Reserve Bank of Philadelphia (November-December) 3-10.

- Chaing, A. (1974). *Fundamental Methods of Mathematical Economics*. New York: McGraw-Hill.
- Chow, B. and Brophy, D. (1978). The U.S. Treasury bill market and hypotheses regarding the term structure of interest rates. *The Financial Review* 13 (Fall) 36-50.
- Churchman, C.W. (1971). *The Design of an Inquiring System*. New York: Basic Books.
- _____ (1977). A philosophy for complexity. In *Futures Research: New Directions* edited by H.A. Linstone and W.H.C. Simmonds. Reading: Addison Wesley.
- Citibank, NA (1983). CITIBASE: Citibank economic database. Machine-readable magnetic data file. New York.
- Cleary, J. (1982). *The Professional Forecaster*. Belmont, California: Lifetime Learning Publications.
- The Conference Board (1973). *A Guide to Forecasting Interest Rates*. New York: The Conference Board.
- Conrad, J. (1963). *An Introduction to the Theory of Interest*. Berkeley: University of California Press.
- Cooley, W. and Lohnes, P. (1971). *Multivariate Data Analysis*. New York: John Wiley.
- Cooper, L. and Steinberg, D. (1970). *Introduction to Methods of Optimization*. Philadelphia: W.B. Saunders.
- Cootner, P. (ed.) (1964). *The Random Character of Stock Market Prices*. Cambridge: M.I.T. Press.
- Coxeter, H. (1961). *Introduction to Geometry*. New York: John Wiley.
- Cramer, R. and Miller, R. (1976). Dynamic modeling of multivariate time series for use in bank analysis. *Journal of Money, Credit and Banking* 8 (February) 85-96.
- Dale, C. and Workman, R. (1980). The arc sine law and the treasury bill futures market. *Financial Analysts Journal* 36 (November-December) 71-4.

- Davidson, P. (1978). Money and the Real World. New York: John Wiley.
- Dreman, David (1984). The world according to Keynes. Forbes 133 (January 30) 136.
- Duda, R. and Hart, P. (1973). Pattern Classification and Scene Analysis. New York: John Wiley.
- Fahremeier, L. (1976). Recursive parameter estimation in linear statistical models. In Compstat 1976: Proceedings in Computational Statistic edited by J. Gordesch and P. Naeve. Wein: Physica-Verlag. 139-45.
- Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. Journal of Finance 25 (May) 383-47.
- Federal Reserve Bank of Minneapolis (1978). Rational expectations--fresh ideas that challenge some established views of policy making. In 1977 Annual Report. Minneapolis.
- Feller, W. (1957). An Introduction to Probability Theory and Its Applications. New York: John Wiley.
- Felsen, J. (1975a). Artificial intelligence techniques applied to reduction of uncertainty in decision analysis through learning. Operational Research Quarterly 25 (October) 581-98.
- _____ (1975b). Cybernetic Approach to Stock Market Analysis. Hicksville, New York: Exposition Press.
- Feldstein, M. and Eckstein, O. (1970). The fundamental determinants of the interest rate. Review of Economics and Statistics 52 (November) 363-75.
- Fihr, I. (1930). The Theory of Interest. New York: Macmillian.
- Fogler, H. (1974). A pattern recognition model for forecasting. Management Science 20 (April) 1178-89.

- Forrester, J. (1976). Understanding the counterintuitive behavior of social systems. In *Systems Behavior* edited by J. Beishon and G. Peters. London: Harper & Row. 223-40.
- Fraser, D. (1977). On the accuracy and usefulness of interest rate forecasts. *Business Economics* 12 (September) 38-44.
- Friedman, J. and Tukey, J. (1974). A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers* C-23 (September) 881-890.
- Friedman, M. (1969). Factors affecting the level of interest rates. *Proceedings of the 1968 Conference on Savings and Residential Financing* sponsored by the United States Savings and Loan League. Chicago: The League. 11-27. Reprinted in *Current Issues in Monetary Theory and Policy* edited by T. Havrilesky and J. Boorman (1976). 362-79.
- Fu, K. (1976). *Digital Pattern Recognition*. New York: Springer-Verlag.
- Fukunaga, K. and Kessell, D. (1971). Estimation of classification error. *IEEE Transaction on Computers* C-20 (December) 1521-7.
- The Futures Discount Group (1983). *Statement of fees*. Chicago.
- Gardner, M. (1968). On the meaning of randomness and some ways of achieving it. *Scientific American* 219 (July) 116-21.
- Gay, G. and Kolb, R. (eds.) (1982). *Interest Rate Futures: Concepts and Issues*. Richmond, Virginia: Robert F. Dame.
- Georgescu-Roegen, N. (1975). Bio-economic aspects of entropy. In *Entropy and Information in Science and Philosophy* edited by L. Kubat and J. Zeman. New York: Elsevier. 125-42.
- Gerbino, K. (1982). *The secret of wall street*. Beverly Hills: Gerbino & Co.
- Gibson, W. and Kaufman, G. (1968). The sensitivity of interest rates to changes in money and income. *Journal of Political Economy* 76 (May-June) 472-8.

- Gilbert, R. and Trebing, M. (1981). The FOMC in 1980: a year of reserve targeting. Review of the Federal Reserve Bank of St. Louis 63 (August-September) 2-16.
- Gilman, L. and Rose, A. (1976). APL--An Interactive Approach. New York: John Wiley.
- Gordesch, J. and Naeve, P. (eds.) (1976). Compstat 1976: Proceedings in Computational Statistics. Wein: Physica-Verlag.
- Graham, B. and Dodd, D. (1962). Security Analysis. New York: McGraw-Hill.
- Granger, C.W.J. (1969). Prediction with a generalized cost of error function. Operational Research Quarterly 20 (June) 199-207.
- Granger, C.W.J. and Newbold, P. (1975). Economic forecasting: the atheist's viewpoint. In Modelling the Economy edited by G. Renton. London: Heinemann Educational Books. 131-47.
- _____. (1977). Forecasting Economic Time Series. New York: Academic Press.
- Green, P. and Carroll, J.D. (1976). Mathematical Tools for Applied Multivariate Analysis. New York: Academic Press.
- Hafer, R. (1983). Weekly money supply forecasts: effects of the October 1979 change in the monetary control procedures. Review of the Federal Reserve Bank of St. Louis 65 (April) 26-32.
- Haitovsky, Y., Treyz, S. and Su, V. (1974). Forecasts with Quarterly Macroeconomic Models. New York: Columbia University Press.
- Hamburger, M. and Silber, W. (1969). An empirical study of interest rate determination. Review of Economics and Statistics 51 (August) 369-73.
- Hare, V.C. (1967). Systems Analysis: A Diagnostic Approach. New York: Harcourt Brace Jovanovitch.
- Harris, R. (1975). A Primer of Multivariate Statistics. New York: Academic Press.

- Harrison, P. and Stevens, C. (1971). A Bayesian approach to short-term forecasting. *Operational Research Quarterly* 22 (December) 341-62.
- _____ (1976). Bayesian forecasting. *Journal of the Royal Statistical Society* 38B:205-47.
- Harvey, A. (1981). *The Econometric Analysis of Time Series*. New York: John Wiley.
- Havrilesky, T. and Boorman, J. (eds.) (1976). *Current Issues in Monetary Theory and Policy*. Arlington Heights, Illinois: AHM Publishing.
- Heilbroner, R. (1967). *The Worldly Philosophers*. New York: Simon and Schuster.
- Helzer, G. (1983). *Applied Linear Algebra with APL*. Boston: Little, Brown.
- Henderson, J. and Quandt, R. (1980). *Microeconomic Theory: A Mathematical Approach*. New York: McGraw-Hill.
- Henshel, R. and Kennedy, L. (1973). Self-altering prophecies: consequences for the feasibility of social prediction. *General Systems 18 Yearbook of the Society for General Systems Research*. 119-26.
- Highleyman, W. (1962). Linear decision functions with application to pattern recognition. *Proceeding of the IRE* 50 (June) 1501-14.
- Holling, C. (1976). Resilience and stability of ecosystems. In *Evolution and Consciousness: Human Systems in Transition* edited by E. Jantsch and C. Waddington. Reading: Addison-Wesley. 73-92.
- _____ (1977). The curious behavior of complex systems: lessons from ecology. In *Futures Research: New Directions* edited by H.A. Linstone and W.H.C. Simmonds. Reading: Addison-Wesley. 114-29.
- Holt, C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages. *Research Memorandum No. 52*. Office of Naval Research.
- Holt, T. (1982). *The Holt Investment Advisory*. (November 24). Westport, Connecticut: T.J. Holt & Co.

- Horan, L. (1978). Forecasting long-term interest rates--a new method. *Business Economics* 13 (September) 5-9.
- Horn, F. and Farah, V. (1979). *Trading in Commodity Futures*. New York: New York Institute of Finance.
- Huber, P. (1977). *Robust Statistical Procedures*. Philadelphia: Society for Industrial and Applied Mathematics.
- Humphrey, T. (1983). The early history of the real/nominal interest rate relationship. *Review of the Federal Reserve Bank of St. Louis* 69 (May-June) 2-10.
- Hunt, L. (1973). Alternative econometric models for the yield on long-term corporate bonds. *Business Economics* 8 (September) 31-8.
- _____ (1976). *Dynamics of Forecasting Financial Cycles*. Greenwich, Connecticut: JAI Press.
- Iverson, K. (1962). *A Programming Language*. New York: John Wiley.
- Jantsch, E. (1976). Evolution: self-realization through self-transcendence. In *Evolution and Consciousness: Human Systems in Transition* edited by E. Jantsch and C. Waddington. Reading: Addison-Wesley. 37-72.
- Johnson, F. and Harrison, P. (1980). An application of forecasting in the alcoholic drinks industry. *Journal of the Operational Research Society* 31 (July) 699-709.
- Johnson, L. and Montgomery, D. (1979). Forecasting with exponential smoothing and related methods. In *Forecasting* edited by S. Makridakis and S. Wheelwright. Amsterdam: North-Holland. 31-44.
- Kahneman, D. and Tversky, A. (1979). Intuitive prediction: biases and corrective procedures. In *Forecasting* edited by S. Makridakis and S. Wheelwright. Amsterdam: North-Holland. 313-27.
- Kaufman, P. (1978). *Commodity Trading Systems and Methods*. New York: John Wiley.
- _____ (ed.) (1980). *Technical Analysis in Commodities*. New York: John Wiley.
- Kennedy, W. and Gentle, J. (1980). *Statistical Computing*. New York: Marcel Dekker.

- Kendall, M. (1979). Where do we go from here? In Forecasting edited by O.D. Anderson. Amsterdam: North-Holland. 273-79.
- Keynes, J. (1930). A Treatise on Money. London: Macmillan.
- _____ (1936). The General Theory of Employment, Interest, and Money. London: Macmillan.
- _____ (1967). Quoted by Adam Smith (pseudonym) in The Money Game. New York: Random House.
- _____ (1983). Quoted by L. Minard in The original contrarian. Forbes 132 (September 26) 42-52.
- Keys, J. (pseudonym) (1972). Only Two Can Play This Game. New York: Julian Press.
- Kindleberger, C. (1978). Manias, Panics, and Crashes. New York: Basic Books.
- Klein, F. and Prestbo, J. (1974). News and Markets. Chicago: Henry Regnery.
- Koopmans, T. (1947). Measurement without theory. Review of Economics and Statistics 29 (August) 161-72.
- Kubat, L. and Zeman, J. (1975). Entropy and Information in Science and Philosophy. New York: Elsevier.
- Kuhn, T. (1970). The Structure of Scientific Revolutions. Chicago: University of Chicago Press.
- Kullback, S. (1959). Information Theory and Statistics. New York: John Wiley.
- Lachenbruch, P. and Mickey, R. (1968). Estimation of error rates in discriminant analysis. Technometrics 10 (February) 1-11.
- Lang, R. and Rasche, R. (1978). A comparison of yields on futures contracts and implied forward rates. Review of the Federal Reserve Bank of St. Louis (December) 21-30.
- Laszlo, E. (1973). The World System. New York: George Braziller.
- Levy, R. (1971). The predictive significance of five-point chart patterns. The Journal of Business 44 (July) 316-23.

- Lilly, J. (1967). Programming and Metaprogramming in the Human Biocomputer. New York: Julian Press.
- _____. (1975). Simulations of God: The Science of Belief. New York: Bantam Books.
- Linstone, H.A. (1984). Multiple Perspectives for Decision Making: Bridging the Gap between Analysis and Action. New York: Elsevier (in preparation).
- _____. et. al. (1981). The Multiple Perspectives Concept. Futures Research Institute, Portland State University. Portland.
- Linstone, H.A. and Simmonds, W.H.C. (eds.) (1977). Futures Research: New Directions. Reading: Addison-Wesley.
- Liversidge, A. (1983). Money machine. Omni (June) 38, 117.
- Loeb, G. (1965). The Battle for Investment Survival. New York: Simon and Schuster.
- Loosigian, A. (1980). Interest Rate Futures. Princeton: Dow Jones.
- McCracken, J. (1976). Interest rate forecasting ... and other popular delusions. Bankers Magazine 129 (Winter) 71-7.
- MacKay, C. (1932). Extraordinary Popular Delusions and the Madness of Crowds. Wells, Vermont: Fraser Publishing.
- McNees, S. (1973). The predictive accuracy of econometric forecasts. New England Economic Review (September-October) 3-27.
- Maddala, G. (1977). Econometrics. New York: McGraw-Hill.
- Makridakis, S. and Wheelwright, S. (1978). Forecasting Methods and Applications. New York: John Wiley.
- _____, (eds.) (1979). Forecasting. Amsterdam: North-Holland.
- Martino, J. (1972). Technological Forecasting for Decisionmaking. New York: Elsevier.
- Mastrapasqua, F. (1982). Credit Market Comments. (November 24). New York: Smith Barney Harris Upjohn.

- Meisel, W. (1968). Least-squares methods in abstract pattern recognition. *Information Sciences* 1 (December) 43-54.
- Miller, J. (1978). *Living Systems*. New York: McGraw-Hill.
- Miller, R. and Hickman, J. (1973). Time series analysis and forecasting. *Transactions of the Society of Actuaries* 25:267-302.
- Minard, L. (1983). The original contrarian. *Forbes* 132 (September 26) 42-52.
- Mitroff, I. (1977). On errors of the third kind. In *Futures Research: New Directions* edited by H.A. Linstone and W.H.C. Simmonds. Reading: Addison-Wesley. 45-52.
- Moore, G. (1979). Inflation's turn. In *Forecasting* edited by S. Makridakis and S. Wheelwright. Amsterdam: North Holland. 213-25.
- Morgenstern, O. (1963) *On the Accuracy of Economic Statistics*. Princeton: Princeton University Press.
- Mucciardi, A. and Gose, E. (1971) A comparison of seven techniques for choosing subsets of pattern recognition properties. *IEEE Transactions on Computers* C-20 (September) 1023-31.
- Mullineaux, D. (1981). Efficient markets, interest rates, and monetary policy. *Business Review of the Federal Reserve Bank of Philadelphia* (May-June) 3-10.
- Neave, E. and Shaftel, T. (1978). An accelerated technique for ridge following using conjugate directions. In *Computers and Mathematical Programming* edited by W. White. National Bureau of Standards Special Publication No. 502. Washington: U.S. Department of Commerce. 338-353.
- Nelson, C. (1973). *Applied Time Series Analysis*. San Francisco: Holden Day.
- Nelson, R. (1982). *An Evolutionary Theory of Economic Change*. Cambridge: Belknap.
- Oregonian (1983). 1984 interest rates? Ask Ouija board. (December 14) D-13.

- Pearson, J. (1981). Interest rate forecasting models. Unpublished manuscript. Economic Research Workshop, Portland State University. Portland.
- Peterson, D. and Mattson, R. (1966). A method of finding linear discriminant functions for a class of performance criteria. IEEE Transactions on Information Theory IT-12 (July) 380-7.
- Pindyck, R. and Rubinfeld, D. (1981). Econometric Models and Economic Forecasts. New York: McGraw-Hill.
- Poole, W. (1976). Rational expectations in the macro model. Brookings Papers on Economic Activity 2:463-505.
- _____ (1978). Using T-bill futures to gauge interest-rate expectations. Economic Review of the Federal Reserve Bank of San Francisco 3 (Spring) 7, 14-5.
- Powell, M. (1965). An efficient method for finding the minimum of a function of several variables without derivatives. The Computer Journal 7:155-62.
- Powers, M. and Vogel, D. (1981) Inside the Financial Futures Markets. New York: John Wiley.
- Prell, M. (1973). How well do the experts forecast interest rates? Monthly Review of the Federal Reserve Bank of Kansas City (September-October) 3-8.
- Pring, M. (1980). Technical Analysis Explained. New York: McGraw-Hill.
- _____ (1981). How to Forecast Interest Rates. New York: McGraw-Hill.
- Puglisi, D. (1978). Is the futures market for Treasury bills efficient? The Journal of Portfolio Management 4 (Winter) 64-7.
- Reinmuth, J. and Wittink, D. (1974). Recursive models for forecasting seasonal processes. Journal of Financial and Quantitative Analysis 4 (September) 659-84.
- Reinmuth, J. and Geurts, M. (1979). A multideterministic approach to forecasting. In Forecasting edited by S. Makridakis and S. Wheelwright. Amsterdam: North-Holland. 203-12.

- Renton, G. (1975). *Modelling the Economy*. London: Heinemann Educational Books.
- Rider P. and Fischer C. (1962). *Mathematics of Investment*. New York: Holt, Rinehart and Wilson.
- Ridley, B. (1976). *Time, Space and Things*. Harmondsworth, Middlesex, England: Penguin Books.
- Rohatyn, F. (1984). A case for reindustrialization. *Forbes* 133 (January 30) 54-9.
- Rosenblatt, F. (1957). The perceptron: a perceiving and recognizing automaton. Cornell Aeronautical Laboratory Report 85-460-1.
- Santoni, G. and Stone, C. (1981a). Navigating through the interest rate morass: some basic principles. *Review of the Federal Reserve Bank of St. Louis* 63 (March) 11-18.
- _____ (1981b). What really happened to interest rates?: a longer-run analysis. *Review of the Federal Reserve Bank of St. Louis* 63 (November) 3-14.
- Sargent, T. and Sims, C. (1977). Business cycle modeling without pretending to have too much a priori economic theory. In *New Methods in Business Cycle Research: Proceedings from a Conference*. Minneapolis: Federal Reserve Bank of Minneapolis. 45-109.
- Schoeffler, S. (1955). *The Failures of Economics: A Diagnostic Study*. Cambridge: Harvard University Press.
- Schott, F. (1973). Forecasting long-term interest rates. *Business Economics* 8 (September) 46-53.
- _____ (1977). Interest rate forecasting in theory and practice. *Business Economics* 12 (September) 55-60.
- Seligman, D. (1983). Can you beat the stock market? *Fortune* (December 26) 82-96.
- Shannon, C. and Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

- Shiller, R. (1979). The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy* 87 (October) 1190-219.
- Shirk, G. (1980a). Cycles in interest rates. *Cycles* 31 (August) 145-150.
- _____ (1980b). Cycles in interest rates: part 2. *Cycles* 31 (September/October) 165-7.
- _____ (1980c). Cycles in interest rates: part 3. *Cycles* 31 (November) 194-7.
- _____ (1980d). Cycles in interest rates: part 4. *Cycles* 31 (December) 221-3.
- _____ (1981). Cycles in interest rates: part 5. *Cycles* 32 (April) 60-1.
- Shuman, J. and Rosenau, D. (1972). *The Kondratieff Wave*. New York: World Publishing.
- Simonoff, J. (1981). Application of statistical methodology to the evaluation of timing devices in commodity trading. *The Journal of Futures Markets* 1 (Winter) 649-56.
- Skolimowski, H. (1973). The twilight of physical descriptions and the ascent of normative models. In *World System* edited by E. Laszlo. New York: George Braziller.
- Smith, A. (pseudonym) (1967). *The Money Game*. New York: Random House.
- Smith, C. (1981). *The Mind of the Market*. Totawa, New Jersey: Rowman and Littlefield.
- Spencer-Brown, G. (1957) *Probability and Scientific Inference*. New York: Langmans, Green.
- _____ (1972). *Only Two Can Play This Game*. New York: Julian Press.
- Stanley, G., Nienow, W. and Lendaris, G. (1969). SARF: An interactive signature analysis research facility. AC Electronics - Defense Research Laboratories Publication No. TR69-15. Santa Barbara, California.
- Steinbruner, J. (1974). *The Cybernetic Theory of Decision*. Princeton: Princeton University Press.

- Stevens, S. (1968). Measurement, statistics and the schemapiric view. *Science* 161 (August) 849-56.
- Tatsuka, M. (1971). *Multivariate Analysis*. New York: John Wiley.
- Teweles, R., Harlow, C., and Stone, H. (1974). *The Commodity Futures Game: Who Wins? Who Loses? Why?* New York: McGraw-Hill.
- Theil, H. (1978). *Introduction to Econometrics*. Englewood Cliffs: Prentice-Hall.
- Thurow, L. (1983). *Dangerous Currents: The State of Economics*. New York: Random House.
- Tou, J. and Gonzalez, R. (1974). *Pattern Recognition Principles*. Reading: Addison-Wesley.
- Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185 (September) 1124-31.
- U.S. Department of Commerce, Bureau of Economic Analysis (1977). *Handbook of Cyclical Indicators*. Washington.
- _____ (1983). *Business Conditions Digest* (November) Washington.
- U.S. Treasury Department (no date). *Information about the treasury constant maturity yield series*. Mimeograph. Washington.
- Van de Geer, J. (1971). *Introduction to Multivariate Analysis for the Social Sciences*. San Francisco: W.H. Freeman.
- Vignola, A. and Dale, C. (1979). Is the futures market for Treasury bills efficient? *Journal of Portfolio Management* 5 (Winter) 78-81.
- Wall Street Journal (1980). Credit crest? Interest rates appear on verge of a decline after three-year rise. (April 18) 1.
- _____ (1982). Baffled borrowers: corporate treasurers find the rate outlook unusually confusing. (September 22) 1.

- Walpole, R. and Myers, R. (1978). Probability and Statistics for Engineers and Scientists. New York: Macmillan.
- Wantanabe, S. (1972). Frontiers in Pattern Recognition. New York: Academic Press.
- Warfield, J. (1981). Statement on the criteria for selecting research topics. General Systems Bulletin (Autumn) 2.
- Watzlawick, P. (1976). How Real Is Real? Confusion, Disinformation, Communication. New York: Random House.
- Webb, R. (1983). Why economic data should be used with care: the case of the suspiciously slow growth statistic. Economic Review of the Federal Reserve Bank of Richmond (July-August) 12-3.
- Weinberg, G. (1975). An Introduction to General Systems Thinking. New York: John Wiley.
- Winters, P. (1960). Forecasting sales by exponentially weighted moving averages. Management Science 6 (April) 324-42.
- Woodcock, A. and Davis, M. (1978). Catastrophe Theory. New York: Dutton.
- Working, H. (1960). Note on the correlation of first differences of averages in a random chain. Econometrica 28 (October) 916-8.
- Yau, S. and Garnett, J. (1972). Least-mean-squares approach to pattern recognition. In Frontiers of Pattern Recognition edited by S. Wantanabe. New York: Academic Press. 575-87.
- Young, T. and Calvert, T. (1974). Classification, Estimation and Pattern Recognition. New York: American Elsevier.
- Zadeh, L. (1972). A fuzzy-set-theoretic interpretation of linguistic hedges. Journal of Cybernetics 2 (July-September) 4-34.
- Zeeman, E. (1974). On the unstable behavior of stock exchanges. Journal of Mathematical Economics 1 (March) 39-49.

APPENDIX A

SUMMARY OF NOTATION

The principal symbols used in this report are listed below together with a brief explanation and the number of the equation near where they are first defined. Primes (') always indicate vector or matrix transposition.

<u>Eqn</u>	<u>Symbol and Explanation</u>
------------	-------------------------------

(1)	$x_0 = (x_1, x_2, \dots, x_n)'$
-----	---------------------------------

Feature vector

(7)	$x = (x_1, x_2, \dots, x_n, 1)'$
-----	----------------------------------

Augmented feature vector

(8)	$w = (w_1, w_2, \dots, w_n, w_{n+1})'$
-----	--

Augmented weight vector

(10)	$w(i)$
------	--------

Estimate of w after i iterations of the perceptron algorithm

(42)	$w_{<j>}$
------	-----------

Solution vector to problem P_j

(42)	P_j
------	-------

Pattern classification problem consisting of the first j features

(5)	$\hat{d}(x) = w'x$
-----	--------------------

Decision function

$$(4) \quad d(x) = w'x = 0$$

Partition boundary

$$(10) \quad y(i) = x(i) \text{ if } x(i) \text{ is in class } C_1 \\ = -x(i) \text{ if } x(i) \text{ is in class } C_2.$$

Modified feature vector

$$(13) \quad J_P(w) = .5 (|w'y| - w'y)$$

Perceptron criterion function

$$(19) \quad J_R(w) = .125 (|w'y| - w'y)^2$$

Relaxation criterion function

$$(26) \quad J_S(w) = \sum_i (a_i - w'y(i))^2$$

Minimum squared error criterion function

$$(30) \quad J_N(w) = \sum_i |v_i| - \sum_i |v_i| |\text{sgn}(v_i) - \text{sgn}(w'x(i))|$$

Net profit criterion function

$$(31) \quad J_L(w) = .5 \sum_i |v_i| |\text{sgn}(v_i) - \text{sgn}(w'x(i))|$$

Minimum loss criterion function

$$(33) \quad J_T(w) = .5 \sum_i |v_i| |\text{sgn}(v_i) - \tanh(sw'x(i))|$$

Hyperbolic tangent criterion function

$$(36) \quad J_W(w) = \sum_i (u_i a_i - u_i w'y(i))^2$$

Weighted least squared criterion function

$$(23) \quad a$$

A positive vector, sometimes equal to 1, the unit vector

$$(29) \quad b$$

A vector with elements equal to 1 if x is in class C_1 or -1 if x is in class C_2

$$(13) \quad 0$$

The zero vector

(28) $\mathbf{1}$

The unit vector

(30) v_i Change in the interest rate in period i (36) u_i The positive square root of the absolute value of v_i (29) \mathbf{X} A matrix with rows equal to \mathbf{x}' (22) \mathbf{X}^{-1} The inverse of \mathbf{X} , if defined(28) $\mathbf{X}^\# = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ The generalized inverse of \mathbf{X} (22) \mathbf{Y} A matrix with rows equal to \mathbf{y}' which rows equal $\mathbf{x}(i)$ if \mathbf{x} is in class C_1 or $-\mathbf{x}(i)$ if $\mathbf{x}(i)$ is in class C_2 (37) \mathbf{U} A diagonal matrix with nonzero elements equal to u_i (42) \mathbf{W} A matrix with rows equal to $\mathbf{w}'\langle j \rangle$ (14) $\text{grad } f(\mathbf{z}) = \frac{df(\mathbf{z})}{d\mathbf{z}}$

$$= \left(\frac{df}{dz_1}, \frac{df}{dz_2}, \dots, \frac{df}{dz_n} \right)'$$

The gradient of the function f (17) $\text{sgn}(\mathbf{w}'\mathbf{y}) = 1$ if $\mathbf{w}'\mathbf{y} > 0$ $= -1$ if $\mathbf{w}'\mathbf{y} \leq 0$

The sign function

APPENDIX B

TECHNICAL NOTES

1. Notes. To these technical notes the author has relegated material that, for any reason, might have tended to interrupt the flow of the main body of the report.

2. The mathematics of interest and the 10-for-1 rule. A bond is a promise to pay a stipulated amount of money on a specified future maturity date together with stipulated semiannual interest payments to the maturity date. This payment schedule is portrayed in the time diagram below.

_____i_____i_____i_____i_____i . . . _____i_____i_____i_____iP

where i represents the interest payments and P represents the principal payment at maturity.

The price or value of the bond is simply the discounted value of the future payments (Rider and Fischer, 1962):

$$\text{Bond value} = \begin{array}{c} \text{Discounted} \\ \text{value of} \\ \text{interest} \\ \text{payments} \end{array} + \begin{array}{c} \text{Discounted} \\ \text{value of} \\ \text{principal} \end{array}$$

For example, if a \$1000 twenty-year bond bearing 6% coupons (i.e., one paying interest at the rate of \$30 every six months) is discounted to return 10%, the value is

$$656.82 = 514.77 + 142.05$$

The 10-for-1 approximation is a rule of thumb for estimating the change in the present value of a long series of deferred payments when the rate used to discount them changes. For example, the twenty-year bond bearing 6% semiannual coupons is worth only \$598.84 when revalued to yield 11%. Thus the 1% change in the interest rate, from 10% to 11%, changed the value of the bond by \$57.98, or by about 9.7%

3. Commissions and transaction costs. The efficient market theory assumes zero commission, i.e., there are no costs to buy or sell in the market. Consequently, a challenge to the efficient market theory should not be upheld if commissions alone account for the "failure" of the theory.

Commission costs were not considered in the models derived in Chapter IV because the fees are small compared to the excess rates of returns generated by the models, and because incorporating them into the models would have added an unnecessary layer of complexity. This note demonstrates that transaction costs are minor and that their consideration does not invalidate any of the

conclusions reached in this study.

The typical bid-asked spread on actively traded long-term U.S. government bonds is \$1.25 per \$1000 face amount. Thus if a model generated a buy or sell signal every month, the rate of return (calculated on a monthly basis consistent with the the method used in Chapter IV) would have decreased by .125%. However, even the "most active" model generated fewer than 40 signals calling for a change in position; consequently, the average reduction in the rate of return would have been less than .030% ($.0125 \times 40 \times 1/168$). This .030% reduction is less than one-quarter of a standard deviation, so even the poorest forecasting model would continue to reject the efficient market theory at the 99% level of confidence.

It might even be argued that the discussion in the preceding paragraph is unnecessary and that the consideration of commissions should increase the confidence one has in rejecting the EMT. The argument runs as follows: If commissions are to be subtracted in computing the rates of return generated by the forecasting models, a similar subtraction should be made to the rates of return generated by the simulation. Since the simulation, on average, would call for 84 position changes, the reduction in the average simulation rate of return would exceed the reduction in the models' rates of return, and the confidence level at which one rejects the EMT would be higher.

4. Risk and the use of federal funds. This note considers the use of federal funds as the alternative investment medium employed when computing the rate of return, the measure used to test the efficient market theory. The preferred alternative would have been U.S. Treasury bills having one month remaining to maturity. Like the long-term bonds, these instruments are backed by the full faith and credit of the U.S. government, and the one-month maturity matches exactly the investment time horizon used in the study. The necessary data for this series was not available. The next most obvious alternative is three-month T-bills, but in order to properly compute capital gains and losses at the end of the month, the rate on T-bills with two months to maturity would be required--another series not available.

These data difficulties lead to the use of federal funds as the alternative to long-term bonds. Having a maturity of one day, Federal funds are the most liquid of all assets except cash (and note that the daily maturity makes the use of the monthly average yield appropriate in computing the rate of return). They are exchanged between commercial banks and certain other financial institutions. The closely related market for repurchase agreements effectively makes the federal funds market available to all large financial institutions. Federal funds entail some credit risk, i.e., there is some risk that the borrower

will default, but actual default is virtually unknown and so the risk must be considered minimal.

Now consider the question of risk more generally. Could it be that the models give evidence for the rejection of the efficient market theory because they incur greater risk? The answer is no, and in fact the risk borne by following the models is much less than the risk of the simulated random strategy. On average, the simulation had the investor committed to bonds half the time and to federal funds half the time; in contrast, the models had the investor committed to bonds less than one-third of the time. Now since the monthly variance of the return on bonds is more than 100 times as great as the variance of the return on federal funds, investors following the trading rules implied by the models accepted much less risk but achieved superior returns. Additional evidence that the models incurred less risk comes from the statistics on the worst computed loss. Hence any bias is on the side of the efficient market theory, not on the side of the models.

5. Profit discounting and accumulation. Many would argue that the unweighted sum of net profits is not an appropriate measure by which to judge or compare models. Some would say that the profits should be accumulated at some rate of interest, perhaps the rate on Treasury bills. Others would argue that the hypothetical profits from the distant past should be discounted because the economy has

probably changed and so it is the more recent profits that deserve greater weight in choosing between forecasting models. Both views have merit. The author has taken the easy decision to do neither. Provided with the month-by-month profit, the user is free to accumulate or discount at any rate thought appropriate.

6. Terminology. The terminological terrain gets treacherous from time to time even within a given field. For example, those pattern recognition techniques that make assumptions about underlying statistical distributions are frequently called Bayesian or statistical, while those that make no such assumptions are called deterministic. The distinction between the methods is better conveyed by the words parametric and nonparametric. So-called deterministic methods are frequently used on nondeterministic problems and often are just as Bayesian and just as statistical in spirit as the so-called Bayesian or statistical methods. It is, however, twenty years too late to complain about usage so firmly established in the literature.

When terminology becomes so jumbled and contradictory that even experts and authorities disagree on the meaning of words, that terminology is best avoided. For this reason the terms ex-post and ex-ante have not been used here. Compare, for example, the discussion of these words by Theil (1978) with that by Pindyck and Rubinfeld (1981).

A third type of terminology problem appears not through a poor choice of words nor because of disagreement by experts over their meaning, but because the same word is used in different senses in different fields. Linstone, for example, uses "technical" to refer to a perspective that is "rational" and "scientific." In the investment field, "technical" methods are generally considered the least "scientific."

7. Notation. Without good notation advanced mathematical thought is impossible. Without a positional number system, long division is very difficult (try dividing cmxxv by xxxvii without converting). Without Ricci's tensor notation, even Einstein could not have produced the general theory of relativity (Coxeter, 1961). And without Iverson's (1962) algorithmic notation and its implementation as a computer language, APL, (Gillman and Rose, 1976) the research here could not have been completed in reasonable time. Consider, for example, the following problem:

Given an n -vector of values of a dependent variable and an $(n \times 2)$ matrix of observations on two independent variables, obtain estimates of the ordinary least squares regression parameters. Also examine the residuals for heteroscedasticity related to the first independent variable. Do this by sorting the

observations on that variable and comparing the residuals (sorted in the same order).

A program to solve this problem, written in a language like BASIC or FORTRAN, would require pages of coding and would require several hours to develop and debug. In APL, the code takes five lines and three minutes. See Figure 54.

8. Error types. At least four types of errors can be distinguished. First are the two standard statistical types: predicting that interest rates will rise when they actually fall, and predicting that they will decrease when they actually increase. The probabilities of making these kinds of errors can be estimated from the data. The third kind of error is misspecifying the forecasting model (Kendall, 1979). The probability of making this kind of error is close to unity when one deals with systems as complex as the economy. The fourth type of error is solving the wrong problem (Mitroff, 1977).

Errors of the last two types usually receive insufficient attention. In the present instance, for example, the author's motivation to pursue this research came from a desire to find a method of improving investment performance. Initially this was translated to mean, "Find a method to estimate the future values of interest rates." Only when this goal seemed impossible to achieve, did it become apparent that the operational goal was to forecast

```

[0] Y REGRESS4ON X;XX;B;RES;S;SX1;SRIS;F
[1] ' ' ◊ 'REGRESSION COEFFICIENTS: ',,'F6.3' DFMT B+YBXX+1,X
[2] ' ' ◊ 'RESIDUALS : ',,'(F+ 'F6.2')DFMT RES+Y-XX+.xB
[3] ' ' ◊ 'SORTED FIRST VARIABLE : ',,F DFMT SX1+XI;1][S+4XI;1]]
[4] 'CORRESPONDING RESIDUALS: ',,F DFMT SRIS+RES[S]

```

```

      X
13.1  8.86
15.3  11.18
14.8  10.88
14.8  9.88
17.2  17.46
19.4  14.64
19.1  15.32
17.2  18.32
20.5  15.3

```

```

      Y
14.214 19.182 17.712 15.812 25.594 24.936 25.428 27.368 25.97

```

Y REGRESS4ON X

REGRESSION COEFFICIENTS: -4.163 0.704 1.044

RESIDUALS : -0.09 0.91 0.11 -0.75 -0.57 0.17 0.16 0.31 -0.26

SORTED FIRST VARIABLE : 13.10 14.80 14.80 15.30 17.20 17.20 19.10 19.40 20.50

CORRESPONDING RESIDUALS: -0.09 -0.75 0.11 0.91 -0.57 0.31 0.16 0.17 -0.26

Figure 54. An ordinary least regression program written in the notation of APL.

the direction of change of interest rates. Redefined, the problem seems to have a solution.

9. Ashby's law of requisite variety. The law of requisite variety is a mathematical theorem about the relationship of entries in a rectangular table or matrix (Ashby, 1964).

One effect of Ashby's law, says Hare (1967), is to direct our attention, as students of systems, to the study of two types of techniques and methods: those that increase the possible variety of actions available to us in a given situation, and those that we use to simplify, restrict, partition or otherwise cut the system variety as the need arises.

In the context of forecasting, the two types of action correspond to increasing the complexity of the forecasting system and to reducing the range of events or outcomes that must be forecast. By restricting the range of outcomes to two--interest rates increase or interest rates decrease--the forecasting problem becomes more manageable.

10. Other versions of the EMT. Chapter II described two versions of the efficient market theory: the "weak" version which states that price history alone is useless in forecasting price changes; and the "strong" form that no publicly available information can help in forecasting price changes. There are two other versions: the "very strong" version asserts that even inside, nonpublic information is useless; and the "very weak" version, which says only that profitable arbitrage is

impossible.

11. PR vs. OLS. It was argued in Chapter IV that this study rejects the efficient market theory while some others have not, because of the use here of different forecasting techniques and different information sets. This note demonstrates the superiority of pattern recognition over the usual application of ordinary least squares regression in making certain types of estimates.

As usually applied, regression is designed to make estimates of a continuous dependent variable based on information about a related independent variable, which is also usually measured continuously. The parameters of the estimating equation are computed in such a way as to obtain the "best fit", the term being defined, for ordinary least squares, as the minimum of the sum of the squared deviations. Mathematically, the problem is to find the set of coefficients w that gives the "best" solution to the set of inconsistent simultaneous linear equations

$$Xw = z$$

The OLS solution is

$$w = (X'X)^{-1}X' z$$

where z is the dependent variable and the rest of the notation is that previously established.

But it should be noted that there is nothing sacred about the least squares definition of best fit. Other definitions have been used, though not as widely. Perhaps

the best known alternative to OLS is the so-called MAD regression where the best fit is defined as the minimum of the absolute deviations. It is not at all clear that the quadratic loss function of OLS is superior in any given application to the linear one of MAD, but the solution to the former is certainly easier computationally.

Pattern recognition has a different objective than OLS. It seeks only to make a binary estimate (on/off, yes/no, up/down, malignant/benign), not a continuous one. The parameters of the estimating equation are computed in a manner to obtain the best separation between the two classes. Mathematically, the problem is to find the set of coefficients w that give the "best" solution to the set of inconsistent simultaneous inequalities

$$Yw > 0$$

Again, "best" can be variously defined. A frequently used definition is the minimum probability of error, i.e., best means to solve the greatest number of inequalities simultaneously. Unfortunately, there is no known analytic solution to the mathematical problem just posed, and in fact the literature of pattern classification is largely devoted to finding approximate solutions to the set of inequalities.

Some problems lend themselves to OLS estimates, some to PR estimates: the birth weight of a fetus is to be estimated by OLS regression, the sex of the fetus by PR

techniques. Some other problems lend themselves to both. The direction of change in interest rates can be estimated using either technique since the OLS estimate of the amount of change also provides an estimate of the direction of change. The question is: if one is interested only in the direction of change, which technique--OLS or PR--will give the better estimate.

Mathematically, it would be suprising if the parameters that best solved the set of inconsistent equations also gave the best solution to the set of inconsistent inequalities.

The following example illustrates the superiority of PR over OLS when the objective is binary estimation. Let the values of the dependent variable (Y) and the independent variable (X) be given by

Y = -10 -3 -2 -1 -1 15 17 13 10 20 when

X = 5 8 8 9 10 12 13 13 13 15

The OLS equation is

$$Y = -28.09 + 3.20X$$

with a coefficient of determination equal to .92. If the classification task is to estimate whether the variable Y is greater or less than zero, however, use of the regression estimates misclassifies 2 of the 10 points for an error rate of 20%.

When the minimum squared error classifier discussed in Chapter III is applied to the problem, the decision

function is

$$d(X) = .30X - 3.19$$

which correctly classifies all points.

Ashby's law of requisite variety, mentioned above, may provide some insight. A generalized version of one formulation of Ashby's law is

$$V(O) \geq V(D) - \log k - V(R)$$

where $V(O)$ is the variety of the outcome (which one wants to reduce to zero), $V(D)$ is the variety of the disturbances (the complexity of the economy), $V(R)$ is the variety of the regulator (the complexity of the forecasting system), and $\log k$ is related to the number of repetitions in the outcome that arise from any given action by the regulator. The effect of using pattern recognition is to increase the value of k ; when k is large the regulator has less variety to control and hence it might be expected that the variety of the outcome would be less. This is what Hare (supra) means when he speaks of "those [methods] that we use to simplify, restrict, partition or otherwise cut the system variety as the need arises."

A less sophisticated, but perhaps more telling argument is an analogy drawn from the experience of those who hunt rabbits. When shooting for "sport" the marksman will carry a .22 caliber rifle and aim for a point in the midsection; when dinner is the objective, he will carry a shotgun and fire along the rabbit's apparent trajectory.

Precision is sacrificed for effectiveness.

APPENDIX C

USE OF THE HYPERBOLIC TANGENT ALGORITHM

This appendix reports on the use of the hyperbolic tangent algorithm (HTA) derived in Chapter III to refine estimates of the parameters in the investment model used in Chapter IV.

Rather than start the iteration of the hyperbolic tangent algorithm with an arbitrary weight vector, it is much more convenient to use the weight vector from the corresponding iteration of the recursively updated weighted least squared error algorithm (WLSA). This starting vector is known to give reasonable results, so the job of the HTA is to refine the values of the vector so as to maximize the criterion function of equation (31).

In 1965, Powell proposed a search technique for finding the optimum of a multivariate function without using derivatives. Neave and Shaftel (1978) improved the method and gave computer programs for its implementation in their paper, "An Accelerated Technique for Ridge Following Using Conjugate Directions." The details are unimportant here, but the method is considered highly efficient (Beveridge and Schechter, 1970).

The HTA estimates of the parameters of the investment

model were superior to those of the WLSA as the following summary results demonstrate:

	WLSA	HTA
Net Profit	1195	1331
Number of Hits	101	107

Only one model has been refined this way because the computational burdens are several orders of magnitude greater with the hyperbolic tangent algorithm than with the weighted least square algorithm. For example, the computer used to make the calculations reported here performs an ordinary least square regression involving 168 observations on one independent variable in 1.8 seconds; the recursively updated weighted least square algorithm uses 38 seconds; but the hyperbolic tangent takes 3.2 hours, about 300 times as long as the WLSA. To use the HTA for all the models and conduct L-method tests would mean that the same computer would need to run 24 hours a day for more than 5 months.

APPENDIX D

USING THE MODELS TO HEDGE A FIXED-INCOME PORTFOLIO

The models developed in Chapter IV were designed to forecast the direction of change of interest rates on long-term government bonds. But forecasts are not ends in themselves, they are made to improve decision making. This appendix examines how a hypothetical investment manager might use the models to improve investment performance.

On January 1, 1969, a small group of private investors establishes a mutual fund with unconventional provisions. Together the investors put up \$1 million dollars. All money is invested in twenty-year, constant-maturity U.S. government bonds. What makes the fund unconventional is its monthly accounting and pay-out/pay-in procedure. At the end of each month the value of the fund is determined. This value is the then-current market value of the bond portfolio including interest accrued and received during the month. If the fund value exceeds \$1 million, the excess is paid out to the investors (bonds are sold if necessary to raise cash for this purpose). If the fund value is less than \$1 million (because interest rates rose and the value of the bonds fell), then the investors agree to pay in the amount

necessary to restore the fund to \$1 million. Consequently, the value of the fund at the beginning of each month is exactly \$1 million. Interest is assumed to be earned at the rate in effect at the beginning of the month, and capital gains or losses are computed based on the change in rates during the month. Operated this way, the fund would have made a net cumulative pay-out (all pay-outs less all pay-ins) of \$911,000 by December 31, 1982, and the fund would still be worth the original \$1 million.

Now let the fund manager adopt a fully hedged strategy, that is, he balances his long bond portfolio by selling an equivalent amount of bonds in the futures market. Assume that the gain (loss) on the futures contracts exactly offsets the loss (gain) on the actual bonds (the so-called perfect hedge). Therefore there can be no net capital gain or loss, just interest earnings. This strategy is both less risky (because capital loss is impossible) and more profitable because the cumulative pay-out by December 31, 1982, totaled \$1,210,000

Next let the fund manager adopt a different hedge strategy, one based on the investment model developed in Chapter IV. If the model forecasts that interest rates will increase during the next month, then the manager hedges the portfolio by selling futures contracts for \$1 million. Again assume that the gain (loss) on the futures contracts

exactly offsets the loss or (gain) on the actual bonds. Therefore, any month that the portfolio is hedged, there can be no net capital gain or loss, only interest earnings. If the model forecasts that interest rates will fall during the month, the portfolio is held unhedged, so that any gain or loss is fully reflected in its value. Operated this way, the fund would have made a net cumulative pay-out of \$1,566,000 by December 31, 1982, and the fund would still have its original \$1 million. The risk accepted is more than that of the fully hedged strategy, but less than that of the unhedged.

The cumulative pay-outs under these three strategies are compared graphically in Figure 55.

The manager might have based his hedge strategy on any of the other simple models developed in Chapter IV. The financial results would have been very similar as the following table shows:

Unhedged	\$ 911,000
Full Hedge	1,210,000
Investment Model Hedge	1,566,000
Price-Investment Model Hedge	1,594,000
Labor Model Hedge	1,547,000
Atheoretical Model Hedge	1,700,000

These results are fairly impressive, particularly since the models were not specifically designed to accommodate a hedging strategy. That is to say, rate

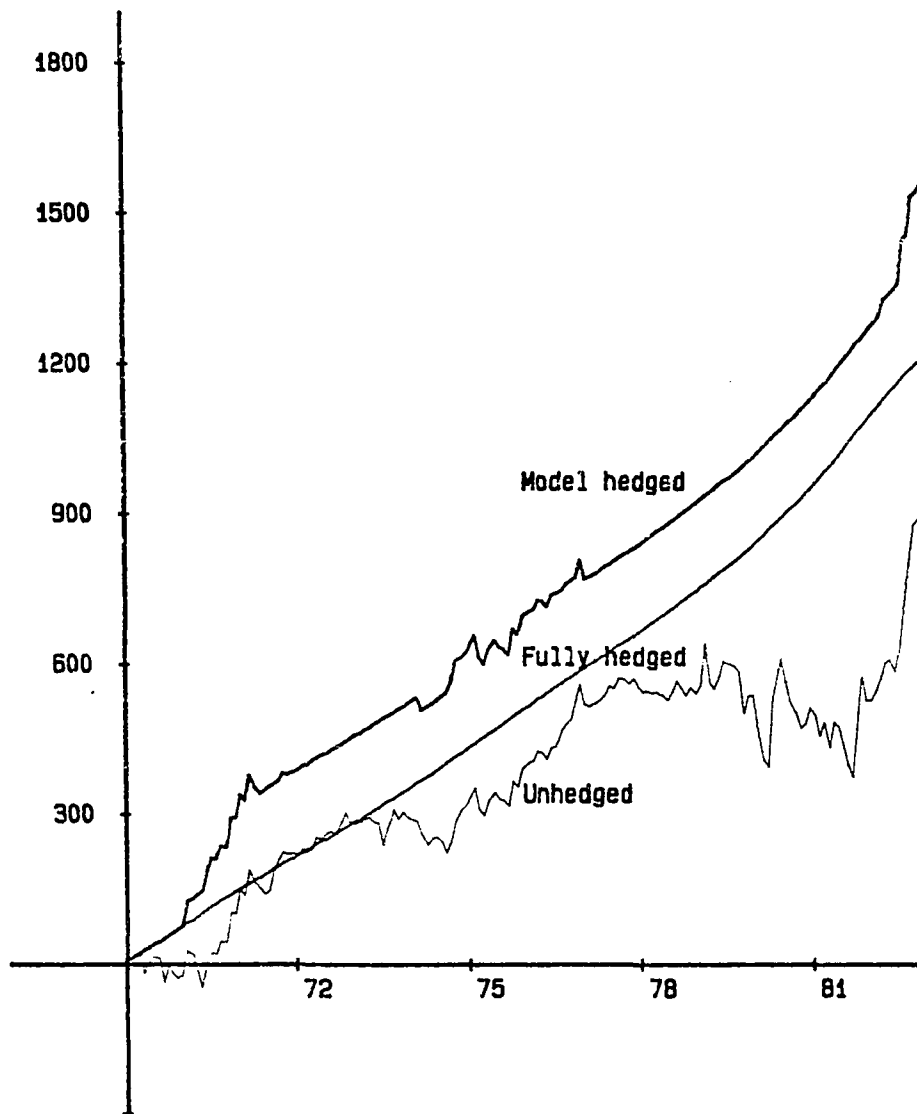


Figure 55. Cumulative pay-out under three fixed-income portfolio investment strategies (in \$1000s)

increases and decreases were given equal weight in the model design; had the models been designed with greater weight being given to identifying those cases where rates decreased, better hedging performance could have been expected.

APPENDIX E

DATA

This appendix contains the series used in the forecasting models (both as reported by the source and as detrended) as well as the coefficients of the models as they developed over time.

The following legend explains the meaning of the column headings.

YR	Year of the series
MO	Month of the series
CC	Construction contracts series
CO	Contracts and orders series
IN	Commodity price series
ER	Employment ratio series
RCL	Ratio of coincident to lagging indicators
INT	Interest rate series
CON	Constant term in the forecasting models

The construction series (CC) is reproduced by permission of the F.W. Dodge Division of McGraw-Hill Information Systems and the price series is used by permission of the Commodity Research Bureau.

TABLE XVIII
 SERIES AS REPORTED BY THE SOURCE

YR	MO	CC	CO	IN	ER	RCL	INT
1967	1	51.27	5.30	105.0	55.69	100.1	4.48
1967	2	57.84	5.69	104.1	55.57	99.2	4.70
1967	3	54.68	5.81	102.0	55.37	98.6	4.56
1967	4	60.36	5.70	100.1	55.64	98.9	4.81
1967	5	54.67	5.88	101.0	55.53	98.8	4.85
1967	6	61.55	6.11	100.8	55.78	98.8	5.08
1967	7	57.71	6.05	99.1	55.88	99.0	5.06
1967	8	61.17	6.26	98.7	55.94	100.5	5.14
1967	9	62.89	6.09	97.9	55.92	100.1	5.19
1967	10	58.87	6.19	96.9	55.99	100.2	5.52
1967	11	60.06	6.22	96.9	55.97	102.4	5.65
1967	12	60.85	6.40	98.1	56.15	103.7	5.56
1968	1	63.08	7.74	98.1	55.51	103.0	5.37
1968	2	59.78	7.81	98.4	55.83	103.1	5.40
1968	3	66.95	9.63	99.0	55.88	103.7	5.62
1968	4	54.03	7.97	98.0	55.95	103.3	5.50
1968	5	62.20	7.32	96.7	56.27	103.6	5.50
1968	6	64.66	7.24	96.1	56.25	104.1	5.36
1968	7	71.99	8.30	95.4	56.10	105.0	5.19
1968	8	67.83	8.39	95.6	55.98	104.5	5.22
1968	9	65.91	7.77	96.4	55.99	104.5	5.32
1968	10	75.31	9.29	97.1	55.96	105.0	5.46
1968	11	69.40	7.98	100.1	56.07	105.1	5.66
1968	12	71.06	8.75	100.8	56.20	104.4	5.98
1969	1	84.15	8.98	102.9	56.14	104.0	6.14
1969	2	70.98	9.33	105.1	56.47	104.1	6.17
1969	3	67.37	8.69	106.2	56.41	104.3	6.22
1969	4	71.45	9.93	107.8	56.45	103.6	5.98
1969	5	82.47	9.35	109.8	56.28	102.6	6.43
1969	6	81.66	8.85	111.3	56.54	102.0	6.28
1969	7	70.30	8.83	111.3	56.52	102.6	6.23
1969	8	72.35	8.75	113.2	56.65	102.4	6.32
1969	9	73.69	9.93	113.9	56.57	102.0	6.82
1969	10	78.65	8.84	113.6	56.62	101.6	6.58
1969	11	61.78	8.81	114.2	56.61	100.5	6.80
1969	12	84.18	8.95	114.6	56.68	100.2	6.94

TABLE XVIII (continued)
 SERIES AS REPORTED BY THE SOURCE

YR	MO	CC	CO	IN	ER	RCL	INT
1970	1	82.65	9.20	116.1	56.64	96.5	6.96
1970	2	81.53	8.86	116.6	56.50	96.5	6.55
1970	3	71.78	8.37	116.9	56.53	96.3	6.68
1970	4	66.91	8.00	116.3	56.50	97.2	7.07
1970	5	55.79	8.10	115.0	56.15	96.9	7.46
1970	6	58.91	7.80	113.8	55.97	96.2	7.11
1970	7	65.39	8.15	112.9	56.09	96.7	6.91
1970	8	64.81	7.72	112.9	55.92	95.8	7.00
1970	9	58.80	7.94	112.0	55.74	96.0	6.81
1970	10	52.68	7.04	110.9	55.78	94.0	6.90
1970	11	53.37	8.11	109.2	55.66	93.8	6.39
1970	12	53.53	8.90	107.2	55.53	96.9	6.45
1971	1	53.75	8.21	107.1	55.63	99.5	6.07
1971	2	51.66	8.94	109.9	55.43	99.7	6.24
1971	3	67.01	9.02	109.3	55.26	100.5	5.83
1971	4	57.86	8.89	109.7	55.45	101.8	6.10
1971	5	63.04	8.65	108.8	55.47	102.1	6.27
1971	6	62.66	9.73	108.1	55.11	104.4	6.44
1971	7	58.75	8.00	108.3	55.40	103.4	6.43
1971	8	57.09	8.88	108.3	55.48	102.1	6.09
1971	9	64.47	9.39	107.4	55.49	103.0	5.97
1971	10	61.98	8.49	106.7	55.55	103.6	5.87
1971	11	69.84	9.25	105.8	55.73	104.7	5.95
1971	12	65.84	9.54	106.7	55.76	105.9	5.99
1972	1	63.71	8.75	110.3	55.71	109.2	6.09
1972	2	65.86	9.23	112.4	55.72	109.8	6.03
1972	3	68.12	9.94	114.4	55.96	110.7	6.11
1972	4	65.59	9.81	115.6	55.95	111.4	6.12
1972	5	85.60	10.79	119.2	56.02	111.5	5.96
1972	6	69.61	9.39	119.1	56.06	110.4	6.05
1972	7	66.67	10.47	119.8	56.06	111.8	5.98
1972	8	70.71	9.69	121.0	56.21	113.6	5.99
1972	9	71.36	11.07	122.7	56.09	114.3	6.09
1972	10	77.00	10.65	124.5	56.05	115.9	5.93
1972	11	76.52	10.98	126.2	56.21	117.2	5.82
1972	12	80.83	11.21	130.8	56.39	118.8	6.03

TABLE XVIII (continued)
 SERIES AS REPORTED BY THE SOURCE

YR	MO	CC	CO	IN	ER	RCL	INT
1973	1	90.73	11.26	134.4	56.14	118.2	6.86
1973	2	88.09	11.95	143.0	56.57	118.4	6.90
1973	3	88.28	12.01	149.9	56.84	118.4	6.89
1973	4	84.53	12.16	152.9	56.82	116.3	6.91
1973	5	81.81	12.85	161.1	56.79	116.3	7.05
1973	6	84.05	12.73	171.2	57.06	115.4	7.13
1973	7	93.16	13.04	181.9	57.04	114.9	7.63
1973	8	89.80	13.11	207.8	56.89	114.1	7.35
1973	9	80.88	13.02	194.9	56.98	113.7	7.05
1973	10	91.60	14.41	192.0	57.22	114.5	7.30
1973	11	87.38	14.55	192.1	57.35	115.1	7.22
1973	12	73.02	13.90	204.3	57.31	113.4	7.38
1974	1	75.89	13.88	213.3	57.32	112.3	7.48
1974	2	84.49	14.27	232.0	57.36	111.9	7.57
1974	3	77.05	14.64	233.0	57.32	112.5	7.89
1974	4	85.92	13.93	230.8	57.12	110.4	8.19
1974	5	75.91	15.34	221.6	57.17	109.9	8.13
1974	6	72.49	14.17	224.4	57.11	109.6	8.17
1974	7	73.37	16.64	236.9	57.16	109.5	8.33
1974	8	85.19	15.12	240.8	56.99	108.6	8.64
1974	9	73.48	15.61	230.5	56.91	106.7	8.49
1974	10	58.95	14.94	231.5	56.78	105.5	8.12
1974	11	57.97	13.52	227.8	56.45	102.1	7.96
1974	12	56.95	14.71	213.1	56.04	97.8	7.93
1975	1	53.68	13.38	205.1	55.62	96.0	7.82
1975	2	49.74	12.45	201.9	55.29	96.3	7.72
1975	3	42.90	12.03	198.6	55.17	95.3	8.24
1975	4	54.06	13.77	201.2	55.09	98.4	8.44
1975	5	45.17	14.34	194.5	55.17	101.2	8.20
1975	6	50.17	14.15	187.2	55.01	105.8	8.09
1975	7	47.89	13.24	195.5	55.24	107.0	8.31
1975	8	43.86	14.63	205.6	55.36	109.7	8.42
1975	9	51.36	12.53	206.3	55.30	111.5	8.61
1975	10	50.61	12.52	201.4	55.28	111.8	8.15
1975	11	42.38	12.76	191.7	55.25	113.6	8.34
1975	12	52.37	12.04	189.7	55.29	114.2	8.04

TABLE XVIII (continued)
 SERIES AS REPORTED BY THE SOURCE

YR	MO	CC	CO	IN	ER	RCL	INT
1976	1	44.59	14.47	190.4	55.67	116.4	8.00
1976	2	50.10	14.15	193.2	55.77	118.4	8.02
1976	3	52.62	14.90	196.0	55.88	119.7	7.90
1976	4	51.77	14.81	202.3	56.08	120.9	7.99
1976	5	52.70	13.66	202.8	56.23	120.5	8.19
1976	6	52.53	15.97	207.4	55.99	121.3	8.02
1976	7	53.45	17.16	216.0	56.24	121.4	8.04
1976	8	52.50	15.32	206.9	56.23	121.6	7.81
1976	9	45.74	16.55	203.1	56.11	120.9	7.75
1976	10	51.30	16.98	196.5	56.10	120.2	7.72
1976	11	56.11	16.23	197.2	56.24	122.2	7.44
1976	12	53.03	16.49	200.6	56.27	123.2	7.17
1977	1	54.91	16.66	207.3	56.27	123.6	7.61
1977	2	53.92	16.61	213.0	56.45	123.9	7.73
1977	3	63.42	16.06	218.4	56.66	125.6	7.72
1977	4	55.88	17.46	220.8	56.87	125.7	7.72
1977	5	63.02	19.14	218.7	57.07	126.0	7.68
1977	6	58.53	18.83	208.5	57.10	125.7	7.57
1977	7	59.64	16.84	204.1	57.08	126.5	7.68
1977	8	73.25	18.92	200.8	57.22	125.6	7.53
1977	9	66.59	20.29	201.3	57.29	126.1	7.61
1977	10	66.05	18.34	203.3	57.42	126.1	7.76
1977	11	68.58	18.86	205.9	57.78	125.9	7.75
1977	12	72.79	20.96	212.7	57.91	126.2	7.98
1978	1	80.68	20.20	218.0	58.00	123.1	8.16
1978	2	69.58	22.43	220.3	58.00	123.9	8.21
1978	3	67.01	20.54	226.4	58.07	124.7	8.31
1978	4	76.04	20.56	228.0	58.39	128.1	8.37
1978	5	89.34	22.89	228.1	58.53	126.9	8.50
1978	6	84.54	21.22	229.7	58.78	126.8	8.67
1978	7	79.20	22.97	228.9	58.54	126.9	8.59
1978	8	80.35	24.17	236.2	58.66	127.1	8.43
1978	9	87.28	24.91	243.0	58.72	126.6	8.63
1978	10	84.95	28.26	251.0	58.90	127.8	8.90
1978	11	86.12	25.33	252.2	59.03	126.7	8.80
1978	12	82.87	22.78	250.8	59.00	126.9	8.99

TABLE XVIII (continued)
 SERIES AS REPORTED BY THE SOURCE

YR	MO	CC	CO	IN	ER	RCL	INT
1979	1	85.78	25.43	255.3	59.14	125.3	8.90
1979	2	104.38	27.22	268.1	59.28	124.8	8.12
1979	3	94.15	29.81	277.4	59.27	127.3	9.07
1979	4	96.06	27.48	276.3	59.01	122.5	9.26
1979	5	89.32	24.74	277.1	59.03	124.5	9.08
1979	6	86.61	26.29	278.1	59.16	123.1	8.82
1979	7	92.79	26.62	281.2	59.27	123.5	8.98
1979	8	84.75	25.36	279.5	59.06	122.4	9.10
1979	9	91.05	26.36	281.1	59.31	120.6	9.31
1979	10	95.23	26.29	283.8	59.21	119.5	10.40
1979	11	81.97	28.14	281.0	59.23	118.7	10.10
1979	12	84.18	27.63	286.2	59.36	119.0	10.16
1980	1	99.43	28.47	287.1	59.23	119.4	11.16
1980	2	82.08	25.74	294.1	59.20	117.7	12.12
1980	3	78.31	26.24	285.3	58.95	113.7	12.44
1980	4	72.76	26.23	272.5	58.60	109.7	10.92
1980	5	67.35	23.46	264.1	58.36	109.9	10.46
1980	6	71.59	25.34	260.3	58.19	112.4	10.08
1980	7	74.62	27.09	274.6	58.12	115.6	10.80
1980	8	71.41	26.52	288.7	58.06	117.2	11.42
1980	9	64.15	26.75	292.8	58.15	119.5	11.84
1980	10	73.46	26.74	296.6	58.21	121.2	12.38
1980	11	90.80	27.61	298.4	58.29	121.0	12.38
1980	12	87.75	28.70	287.7	58.27	118.8	12.09
1981	1	83.72	28.70	281.7	58.38	120.6	12.46
1981	2	83.86	25.75	273.4	58.43	122.0	13.18
1981	3	83.79	28.23	275.1	58.58	123.7	12.95
1981	4	79.64	30.24	276.0	58.80	123.6	13.87
1981	5	84.75	28.54	270.1	58.72	120.2	13.31
1981	6	81.01	28.38	267.9	58.31	120.5	13.59
1981	7	73.46	28.62	274.2	58.44	120.5	14.31
1981	8	78.67	28.27	272.6	58.36	119.5	15.15
1981	9	68.12	27.92	266.4	57.94	117.5	15.78
1981	10	74.26	26.96	260.5	58.02	115.6	14.75
1981	11	70.77	27.88	254.8	57.88	114.9	13.13
1981	12	70.65	26.66	249.4	57.51	113.3	14.04

TABLE XVIII (continued)
 SERIES AS REPORTED BY THE SOURCE

YR	MO	CC	CO	IN	ER	RCL	INT
1982	1	58.18	26.62	251.5	57.46	109.8	14.21
1982	2	63.29	28.51	252.6	57.41	111.7	14.09
1982	3	61.15	25.78	247.2	57.29	111.3	13.89
1982	4	58.93	25.48	245.8	57.17	109.7	13.57
1982	5	53.71	23.33	248.6	57.40	110.9	13.63
1982	6	64.87	23.31	240.0	57.17	109.9	14.18
1982	7	57.80	23.33	241.0	57.06	109.8	13.63
1982	8	59.78	23.03	238.4	57.06	110.6	12.59
1982	9	55.95	24.54	237.7	56.92	110.7	11.65
1982	10	54.65	23.51	232.8	56.65	109.6	10.79
1982	11	50.69	23.82	228.0	56.57	111.9	10.79
1982	12	49.55	24.84	226.8	56.50	113.4	10.62

TABLE XIX
SERIES AS DETRENDED

YR	MO	CC	CO	IN	ER	RCL
1967	1	0.0000	0.0000	0.0000	0.0000	0.0000
1967	2	0.0000	0.0000	0.0000	0.0000	0.0000
1967	3	-9.7300	-0.2700	-1.2000	-0.0800	0.3000
1967	4	2.3760	-0.4094	-0.5972	0.4169	1.0993
1967	5	-7.3920	-0.0733	2.4035	0.0486	0.6458
1967	6	3.0228	0.0844	1.9520	0.2940	0.5448
1967	7	-3.7675	-0.1488	0.1389	0.2213	0.6673
1967	8	0.5073	0.0183	0.5747	0.1360	1.9924
1967	9	1.0224	-0.2735	0.3851	0.0056	0.7231
1967	10	-4.4489	-0.1646	0.0641	0.0199	0.4668
1967	11	-2.5151	-0.1634	0.8643	-0.0618	2.3975
1967	12	-1.6095	-0.0095	2.5676	0.0863	2.7060
1968	1	0.4697	1.2549	2.4792	-0.6348	0.8710
1968	2	-3.6110	0.8528	2.6831	-0.1709	0.3741
1968	3	3.9632	2.3031	3.0979	-0.1048	0.5118
1968	4	-10.5809	-0.1310	1.7717	-0.0349	-0.3920
1968	5	-0.3459	-0.9528	0.4568	0.2680	-0.3690
1968	6	1.7935	-1.0050	0.1450	0.1580	-0.1483
1968	7	8.3166	0.1016	-0.1989	-0.0588	0.4246
1968	8	1.9236	0.0049	0.4308	-0.1996	-0.5257
1968	9	-0.9596	-0.7825	1.5318	-0.1810	-0.7824
1968	10	8.0342	0.7251	2.3137	-0.2050	-0.4858
1968	11	0.0300	-0.8772	5.2409	-0.0836	-0.6411
1968	12	1.0028	-0.1138	5.3297	0.0369	-1.5662
1969	1	13.2371	-0.0189	6.7811	-0.0531	-2.0285
1969	2	-2.8294	0.1806	8.0701	0.2617	-1.9077
1969	3	-6.8073	-0.6418	8.0280	0.1362	-1.6990
1969	4	-2.4703	0.5425	8.4653	0.1292	-2.4161
1969	5	8.1800	-0.2683	9.2120	-0.0871	-3.3208
1969	6	5.4686	-0.8834	9.3259	0.1574	-3.6876
1969	7	-7.4361	-0.9317	7.8984	0.0880	-2.7939
1969	8	-5.1945	-1.0312	8.5403	0.1775	-2.8118
1969	9	-3.9397	0.1447	7.8783	0.0451	-3.0206
1969	10	0.7867	-1.0974	6.2811	0.0592	-3.1961
1969	11	-16.8958	-1.1253	5.7652	0.0114	-4.0431
1969	12	6.8290	-0.9776	5.0986	0.0493	-3.9811

TABLE XIX (continued)

SERIES AS DETRENDED

YR	MO	CC	CO	IN	ER	RCL
1970	1	3.8624	-0.7351	5.5991	-0.0273	-7.3183
1970	2	1.6335	-1.1086	5.0325	-0.1952	-6.5649
1970	3	-8.9827	-1.5889	4.3194	-0.1742	-6.0827
1970	4	-13.5614	-1.8949	2.7755	-0.2152	-4.5419
1970	5	-23.8866	-1.6955	0.6893	-0.5716	-4.3557
1970	6	-18.8734	-1.9140	-1.0835	-0.7204	-4.5817
1970	7	-10.9758	-1.4577	-2.3766	-0.5532	-3.5782
1970	8	-10.8991	-1.8241	-2.6390	-0.6916	-4.0682
1970	9	-16.2448	-1.5028	-3.7726	-0.8258	-3.4053
1970	10	-21.1744	-2.3297	-4.9941	-0.7262	-5.0005
1970	11	-18.8110	-1.1072	-6.6963	-0.7948	-4.6418
1970	12	-17.1688	-0.2742	-8.5358	-0.8663	-1.0099
1971	1	-15.5923	-0.9950	-8.3016	-0.7004	1.8001
1971	2	-16.4456	-0.2317	-5.1783	-0.8481	1.9626
1971	3	0.2325	-0.1836	-5.7170	-0.9523	2.7091
1971	4	-8.9943	-0.3493	-5.2043	-0.6867	3.8897
1971	5	-3.1182	-0.6110	-6.0286	-0.6123	3.9687
1971	6	-3.2770	0.4692	-6.5798	-0.9234	6.0372
1971	7	-6.9499	-1.3474	-6.1805	-0.5587	4.6344
1971	8	-8.0772	-0.4099	-6.0069	-0.4321	3.0377
1971	9	-0.0700	0.0855	-6.7412	-0.3850	3.7618
1971	10	-2.5428	-0.8670	-7.2136	-0.2911	4.1274
1971	11	5.5221	-0.0872	-7.8444	-0.0840	4.9620
1971	12	1.1265	0.1650	-6.6220	-0.0422	5.8308
1972	1	-1.0801	-0.6813	-2.7832	-0.0833	8.7317
1972	2	1.1525	-0.1964	-0.7176	-0.0615	8.7176
1972	3	3.3363	0.4844	1.1026	0.1889	8.9978
1972	4	0.5754	0.2774	1.9951	0.1717	9.0518
1972	5	20.5439	1.1944	5.2249	0.2355	8.4955
1972	6	3.1442	-0.3315	4.5319	0.2648	6.7708
1972	7	-0.0440	0.7248	4.6822	0.2520	7.6569
1972	8	3.9582	-0.1491	5.3194	0.3899	8.8792
1972	9	4.3071	1.1940	6.4116	0.2485	8.9161
1972	10	9.6208	0.6500	7.5295	0.1961	9.8449
1972	11	8.4710	0.8901	8.4716	0.3469	10.4085
1972	12	12.1781	1.0146	12.2494	0.5080	11.2304

TABLE XIX (continued)

SERIES AS DETRENDED

YR	MO	CC	CO	IN	ER	RCL
1973	1	21.2382	0.9508	14.7858	0.2288	9.7946
1973	2	17.1912	1.5303	22.1582	0.6466	9.2410
1973	3	16.2092	1.4439	27.3720	0.8787	8.5155
1973	4	11.3368	1.4519	28.3586	0.8064	5.7287
1973	5	7.7762	1.9985	34.4713	0.7280	5.2034
1973	6	9.3794	1.7022	42.1082	0.9538	3.8060
1973	7	17.7551	1.8521	49.8815	0.8760	2.8874
1973	8	13.1728	1.7525	72.3848	0.6722	1.7198
1973	9	3.2830	1.4976	54.7790	0.7196	1.0173
1973	10	13.5847	2.7362	48.1308	0.9140	1.5536
1973	11	8.3697	2.6549	44.8272	0.9872	1.8595
1973	12	-6.7038	1.7868	53.7832	0.8859	-0.1523
1974	1	-3.7267	1.5950	59.0283	0.8397	-1.4549
1974	2	4.8222	1.8227	73.6647	0.8255	-1.9869
1974	3	-3.1253	2.0175	69.7901	0.7318	-1.4898
1974	4	5.6567	1.1211	62.8883	0.4827	-3.7181
1974	5	-4.9013	2.3910	49.3188	0.4963	-4.2287
1974	6	-8.3223	1.0143	48.4285	0.3990	-4.5112
1974	7	-7.2650	3.3475	57.2633	0.4165	-4.5773
1974	8	4.6819	1.5712	57.0269	0.2130	-5.4382
1974	9	-7.5027	1.8934	42.5792	0.1096	-7.2536
1974	10	-21.8984	1.0388	40.1328	-0.0387	-8.2759
1974	11	-22.0261	-0.5241	33.0912	-0.3797	-11.4445
1974	12	-22.1790	0.5996	15.3808	-0.7839	-15.3540
1975	1	-24.5660	-0.8514	5.2219	-1.1782	-16.5683
1975	2	-27.4992	-1.8324	0.3486	-1.4632	-15.6181
1975	3	-33.1827	-2.2561	-4.3929	-1.5241	-16.0073
1975	4	-20.5860	-0.4989	-3.0091	-1.5415	-12.2724
1975	5	-28.6212	0.0062	-10.9889	-1.3975	-9.0078
1975	6	-22.3835	-0.2721	-19.1946	-1.4998	-4.0915
1975	7	-23.7051	-1.2576	-11.4159	-1.2068	-2.8001
1975	8	-26.7082	0.1024	-2.1884	-1.0368	-0.0667
1975	9	-18.0363	-2.0892	-2.7781	-1.0541	1.6430
1975	10	-17.9981	-2.0917	-8.9404	-1.0303	1.7753
1975	11	-25.4356	-1.8433	-19.6250	-1.0172	3.4011
1975	12	-14.3159	-2.5654	-22.1306	-0.9344	3.7538

TABLE XIX (continued)

SERIES AS DETRENDED

YR	MO	CC	CO	IN	ER	RCL
1976	1	-21.4497	-0.1049	-21.8191	-0.5149	5.6899
1976	2	-14.9760	-0.5016	-19.4143	-0.3935	7.3399
1976	3	-11.7676	0.1890	-17.1074	-0.2673	8.2162
1976	4	-12.0633	0.0099	-11.3943	-0.0564	8.9523
1976	5	-10.5627	-1.2215	-11.7213	0.0956	8.0544
1976	6	-10.2225	1.0595	-7.9309	-0.1489	8.3920
1976	7	-8.8035	2.1243	-0.2973	0.1069	8.0130
1976	8	-9.3112	0.1142	-10.6821	0.0920	7.7475
1976	9	-15.6052	1.2572	-15.3337	-0.0323	6.5907
1976	10	-9.3158	1.5529	-22.5896	-0.0415	5.4797
1976	11	-4.0308	0.6561	-22.2428	0.0997	7.1125
1976	12	-6.8490	0.8055	-19.2039	0.1251	7.6770
1977	1	-4.5918	0.8586	-12.9822	0.1195	7.6167
1977	2	-5.2937	0.6894	-8.0061	0.2941	7.4568
1977	3	4.5238	0.0267	-3.5256	0.4916	8.7012
1977	4	-3.0879	1.3402	-2.2212	0.6812	8.2942
1977	5	4.2801	2.8818	-5.4675	0.8532	8.1010
1977	6	-0.2704	2.3725	-16.6857	0.8483	7.3132
1977	7	0.9555	0.2024	-21.6649	0.7934	7.6540
1977	8	14.6340	2.1858	-25.3470	0.9003	6.2799
1977	9	7.5153	3.3827	-25.0817	0.9330	6.3567
1977	10	6.7855	1.2132	-23.3203	1.0241	5.9290
1977	11	9.1517	1.5955	-21.0199	1.3416	5.3161
1977	12	13.1067	3.5430	-14.6007	1.4168	5.2249
1978	1	20.5909	2.5570	-9.9179	1.4489	1.7358
1978	2	8.8023	4.5969	-8.4065	1.3894	2.2759
1978	3	5.9768	2.4405	-3.1490	1.4018	2.7955
1978	4	14.8535	2.2726	-2.5837	1.6635	5.8954
1978	5	27.6730	4.4203	-3.5384	1.7352	4.2809
1978	6	21.9212	2.4892	-2.9576	1.9139	3.8239
1978	7	15.8319	4.0471	-4.7972	1.5957	3.5826
1978	8	16.4475	4.9982	1.5303	1.6486	3.4491
1978	9	22.8175	5.4543	7.1318	1.6392	2.6196
1978	10	19.6965	8.5031	13.7330	1.7498	3.5190
1978	11	20.1814	5.1624	13.2984	1.8061	2.0859
1978	12	16.2249	2.3178	10.2761	1.7001	2.0026

TABLE XIX (continued)

SERIES AS DETRENDED

YR	MO	CC	CO	IN	ER	RCL
1979	1	18.5631	4.7721	13.2572	1.7674	0.1218
1979	2	36.5062	6.2801	24.4319	1.8320	-0.5937
1979	3	24.9920	8.5347	31.7157	1.7440	1.7156
1979	4	26.0083	5.7902	28.3427	1.4086	-3.3548
1979	5	18.3345	2.7282	26.9792	1.3644	-1.4512
1979	6	14.9478	4.0598	25.8566	1.4315	-3.0121
1979	7	20.5627	4.1255	26.8667	1.4759	-2.7195
1979	8	11.7645	2.5982	23.0375	1.1986	-3.9363
1979	9	17.5987	3.3815	22.6320	1.3904	-5.8115
1979	10	21.1149	3.0681	23.3355	1.2255	-6.9231
1979	11	7.0705	4.6844	18.5112	1.1858	-7.6965
1979	12	8.9581	3.8865	21.8425	1.2572	-7.3430
1980	1	23.8220	4.4641	20.7602	1.0660	-6.8996
1980	2	5.5958	1.4520	25.8094	0.9808	-8.5694
1980	3	1.5412	1.7675	14.8895	0.6783	-12.4833
1980	4	-4.1625	1.5625	0.3203	0.2854	-16.2682
1980	5	-9.5417	-1.3962	-9.3797	0.0151	-15.7282
1980	6	-5.0967	0.3902	-14.1669	-0.1766	-12.9026
1980	7	-2.0026	1.9896	-0.6988	-0.2620	-9.4648
1980	8	-5.2465	1.2178	12.1416	-0.3348	-7.7346
1980	9	-12.4368	1.2701	14.5737	-0.2551	-5.3577
1980	10	-2.8286	1.0806	16.6267	-0.2079	-3.6546
1980	11	14.5091	1.7769	16.6124	-0.1422	-3.9043
1980	12	10.9132	2.6711	4.0957	-0.1785	-6.1455
1981	1	6.4471	2.4470	-3.3330	-0.0836	-4.3161
1981	2	6.2881	-0.7205	-12.8313	-0.0517	-2.9424
1981	3	5.9229	1.6397	-12.0349	0.0793	-1.3106
1981	4	1.4878	3.4572	-12.0609	0.2762	-1.5286
1981	5	6.4480	1.5086	-18.8841	0.1671	-5.0396
1981	6	2.4062	1.1590	-21.7967	-0.2687	-4.7431
1981	7	-5.3236	1.2198	-16.1175	-0.1513	-4.7548
1981	8	-0.0594	0.6886	-18.5069	-0.2474	-5.7653
1981	9	-10.7134	0.1732	-25.4215	-0.6806	-7.7446
1981	10	-4.3570	-0.9368	-31.8253	-0.6007	-9.5634
1981	11	-7.8192	-0.1336	-37.8333	-0.7431	-10.1266
1981	12	-7.8075	-1.4943	-43.3570	-1.1111	-11.5714

TABLE XIX (continued)

SERIES AS DETRENDED

YR	MO	CC	CO	IN	ER	RCL
1982	1	-20.1449	-1.6344	-41.2107	-1.1481	-14.8717
1982	2	-14.5362	0.1598	-40.1211	-1.1839	-12.6725
1982	3	-16.3395	-2.7186	-45.5570	-1.2884	-12.8359
1982	4	-18.1677	-3.0826	-46.8273	-1.3896	-14.1924
1982	5	-22.9399	-5.2855	-43.8534	-1.1377	-12.7073
1982	6	-11.1904	-5.2938	-52.3588	-1.3530	-13.4633
1982	7	-18.0081	-5.2611	-51.0109	-1.4418	-13.2954
1982	8	-15.5772	-5.5485	-53.2941	-1.4178	-12.2304
1982	9	-19.0235	-4.0168	-53.6038	-1.5343	-11.8939
1982	10	-19.8387	-5.0683	-58.0967	-1.7772	-12.7653
1982	11	-23.2878	-4.7491	-62.3533	-1.8230	-10.2101
1982	12	-23.8158	-3.7282	-62.8802	-1.8572	-8.5256

TABLE XX
COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Investment Model		Price-Investment Model		
		CON	CC	CON	CO	IN
1968	1	-0.0481	0.2140	-0.1312	0.0216	-0.3370
1968	2	3.6012	1.0342	56.1761	-7.0217	-22.7183
1968	3	2.2033	0.7135	60.9305	29.4428	-22.2324
1968	4	0.1583	0.1249	-35.3640	-5.4954	13.3507
1968	5	0.1587	0.1228	-0.8108	-0.6968	0.3023
1968	6	-0.3489	-0.0605	0.2797	-0.5397	-0.1119
1968	7	-0.3398	0.0296	-0.8379	-0.2275	0.2741
1968	8	-0.2813	0.0348	-0.5817	-0.3314	0.1859
1968	9	-0.1057	0.0588	-0.2670	-0.4821	0.0905
1968	10	-0.0071	0.0820	-0.0979	-0.5078	0.0772
1968	11	0.1252	0.0938	-0.0680	-0.5455	0.1415
1968	12	0.3077	0.0913	-0.0447	-0.5379	0.1239
1969	1	0.3043	0.0902	-0.1578	-0.4486	0.1884
1969	2	0.3154	0.0901	-0.1421	-0.4428	0.1795
1969	3	0.3308	0.0907	-0.1159	-0.4476	0.1672
1969	4	0.1818	0.0065	0.2872	-0.5400	-0.0202
1969	5	0.3385	-0.0085	0.0967	-0.4677	0.0619
1969	6	0.2564	0.0066	0.2023	-0.3675	0.0145
1969	7	0.2337	0.0084	0.2332	-0.3944	0.0038
1969	8	0.2527	0.0129	0.2246	-0.4016	0.0098
1969	9	0.3364	0.0251	0.2078	-0.4573	0.0213
1969	10	0.2480	0.0421	0.2002	-0.3135	0.0050
1969	11	0.3068	0.0336	0.2262	-0.3537	0.0034
1969	12	0.3365	0.0302	0.2436	-0.3324	0.0060
1970	1	0.3397	0.0302	0.2469	-0.3360	0.0056
1970	2	0.2911	0.0512	0.0430	-0.0781	0.0278
1970	3	0.3029	0.0526	0.0772	-0.1086	0.0238
1970	4	0.3442	0.0559	0.1895	-0.1683	0.0102
1970	5	0.3864	0.0577	0.3028	-0.2420	-0.0087
1970	6	0.3416	0.0659	0.0144	-0.0344	0.0481
1970	7	0.3314	0.0692	-0.0847	0.0353	0.0691
1970	8	0.3435	0.0567	-0.0183	0.0063	0.0558
1970	9	0.3399	0.0587	-0.0855	0.0434	0.0701
1970	10	0.3529	0.0564	-0.0155	0.0310	0.0568
1970	11	0.3170	0.0627	-0.1213	0.0631	0.0788
1970	12	0.3229	0.0599	-0.0735	0.0646	0.0698

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Investment Model		Price-Investment Model		
		CON	CC	CON	CO	IN
1971	1	0.3225	0.0605	-0.0891	0.0711	0.0737
1971	2	0.3325	0.0535	0.0460	0.1193	0.0521
1971	3	0.3260	0.0564	-0.0761	0.0542	0.0674
1971	4	0.3454	0.0505	0.0253	0.0813	0.0520
1971	5	0.3550	0.0468	0.0991	0.1233	0.0429
1971	6	0.3726	0.0477	0.1661	0.1626	0.0344
1971	7	0.3718	0.0477	0.1633	0.1611	0.0347
1971	8	0.3167	0.0464	0.0962	0.1320	0.0443
1971	9	0.2987	0.0460	0.0580	0.1028	0.0486
1971	10	0.2888	0.0462	0.0513	0.1050	0.0501
1971	11	0.2967	0.0459	0.0771	0.1187	0.0464
1971	12	0.3010	0.0461	0.0926	0.1295	0.0443
1972	1	0.3116	0.0464	0.1137	0.1344	0.0408
1972	2	0.2941	0.0450	0.1005	0.1263	0.0419
1972	3	0.3021	0.0455	0.1174	0.1379	0.0409
1972	4	0.3031	0.0455	0.1184	0.1379	0.0409
1972	5	0.2707	0.0437	0.0887	0.1237	0.0409
1972	6	0.2792	0.0443	0.1006	0.1329	0.0413
1972	7	0.2663	0.0436	0.0834	0.1209	0.0408
1972	8	0.2659	0.0436	0.0851	0.1225	0.0408
1972	9	0.2751	0.0442	0.0927	0.1241	0.0418
1972	10	0.2483	0.0429	0.0451	0.0838	0.0399
1972	11	0.2245	0.0412	0.0314	0.0800	0.0372
1972	12	0.2432	0.0426	0.0559	0.1040	0.0383
1973	1	0.2480	0.0430	0.0595	0.1068	0.0392
1973	2	0.2506	0.0433	0.0606	0.1078	0.0395
1973	3	0.2473	0.0429	0.0598	0.1069	0.0384
1973	4	0.2465	0.0428	0.0599	0.1069	0.0381
1973	5	0.2470	0.0429	0.0591	0.1048	0.0358
1973	6	0.2478	0.0430	0.0607	0.1054	0.0336
1973	7	0.2657	0.0448	0.0763	0.1165	0.0238
1973	8	0.2126	0.0404	0.1057	0.1258	0.0066
1973	9	0.1582	0.0353	0.1795	0.1946	-0.0075
1973	10	0.1647	0.0362	0.1767	0.1950	-0.0048
1973	11	0.1503	0.0346	0.1748	0.1921	-0.0058
1973	12	0.1610	0.0351	0.1767	0.1938	-0.0047

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Investment Model		Price-Investment Model		
		CON	CC	CON	CO	IN
1974	1	0.1648	0.0355	0.1803	0.2002	-0.0044
1974	2	0.1694	0.0359	0.1815	0.2036	-0.0041
1974	3	0.1940	0.0347	0.1563	0.1788	-0.0004
1974	4	0.2148	0.0343	0.1422	0.1638	0.0017
1974	5	0.2076	0.0339	0.1446	0.1656	0.0010
1974	6	0.2103	0.0339	0.1457	0.1670	0.0010
1974	7	0.2186	0.0344	0.1448	0.1639	0.0016
1974	8	0.2390	0.0338	0.1497	0.1722	0.0019
1974	9	0.2299	0.0344	0.1629	0.1920	-0.0001
1974	10	0.2072	0.0358	0.0619	0.0514	0.0041
1974	11	0.1894	0.0347	0.0516	0.0431	0.0038
1974	12	0.1877	0.0348	0.0478	0.0392	0.0039
1975	1	0.1859	0.0355	0.0316	0.0271	0.0045
1975	2	0.1845	0.0360	0.0234	0.0287	0.0045
1975	3	0.2098	0.0262	0.0984	0.0817	0.0014
1975	4	0.2166	0.0226	0.1165	0.0798	0.0010
1975	5	0.2143	0.0245	0.1086	0.1035	0.0004
1975	6	0.2141	0.0251	0.1045	0.1136	0.0002
1975	7	0.2224	0.0228	0.1380	0.1306	-0.0013
1975	8	0.2242	0.0210	0.1522	0.1392	-0.0019
1975	9	0.2299	0.0190	0.1694	0.1447	-0.0024
1975	10	0.2231	0.0219	0.1444	0.1668	-0.0026
1975	11	0.2261	0.0196	0.1658	0.1805	-0.0036
1975	12	0.2182	0.0209	0.1485	0.1944	-0.0034
1976	1	0.2172	0.0211	0.1460	0.1959	-0.0033
1976	2	0.2175	0.0209	0.1480	0.1953	-0.0034
1976	3	0.2133	0.0213	0.1444	0.2031	-0.0034
1976	4	0.2157	0.0205	0.1554	0.2105	-0.0040
1976	5	0.2235	0.0198	0.1752	0.2189	-0.0048
1976	6	0.2163	0.0201	0.1457	0.1974	-0.0033
1976	7	0.2172	0.0201	0.1476	0.1986	-0.0034
1976	8	0.2068	0.0205	0.1394	0.2092	-0.0037
1976	9	0.2041	0.0206	0.1254	0.1957	-0.0029
1976	10	0.2026	0.0206	0.1149	0.1838	-0.0023
1976	11	0.1893	0.0210	0.0674	0.1468	0.0003
1976	12	0.1821	0.0219	0.0152	0.0922	0.0035

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Investment Model		Price-Investment Model		
		CON	CC	CON	CO	IN
1977	1	0.2024	0.0213	0.0648	0.1469	0.0005
1977	2	0.2086	0.0214	0.0745	0.1546	0.0000
1977	3	0.2080	0.0214	0.0734	0.1537	0.0000
1977	4	0.2080	0.0214	0.0732	0.1535	0.0001
1977	5	0.2056	0.0214	0.0694	0.1509	0.0002
1977	6	0.1944	0.0208	0.0609	0.1475	0.0005
1977	7	0.2001	0.0209	0.0696	0.1567	0.0000
1977	8	0.1854	0.0201	0.0421	0.1170	0.0019
1977	9	0.1896	0.0203	0.0487	0.1257	0.0015
1977	10	0.1976	0.0206	0.0611	0.1333	0.0008
1977	11	0.1961	0.0205	0.0597	0.1316	0.0009
1977	12	0.2080	0.0213	0.0708	0.1493	0.0001
1978	1	0.2170	0.0218	0.0805	0.1578	-0.0004
1978	2	0.2193	0.0220	0.0829	0.1602	-0.0005
1978	3	0.2237	0.0223	0.0854	0.1647	-0.0007
1978	4	0.2256	0.0225	0.0873	0.1673	-0.0008
1978	5	0.2314	0.0229	0.0885	0.1699	-0.0009
1978	6	0.2389	0.0233	0.0934	0.1762	-0.0012
1978	7	0.2280	0.0224	0.0866	0.1680	-0.0008
1978	8	0.1976	0.0192	0.0707	0.1333	0.0004
1978	9	0.2044	0.0198	0.0765	0.1406	0.0001
1978	10	0.2141	0.0206	0.0803	0.1489	-0.0001
1978	11	0.2025	0.0197	0.0743	0.1308	0.0003
1978	12	0.2078	0.0202	0.0755	0.1347	0.0002
1979	1	0.1971	0.0192	0.0716	0.1060	0.0010
1979	2	0.1184	0.0118	0.0446	0.0275	0.0027
1979	3	0.1533	0.0147	0.0651	0.0403	0.0033
1979	4	0.1591	0.0152	0.0663	0.0457	0.0033
1979	5	0.1410	0.0126	0.0652	0.0297	0.0035
1979	6	0.1217	0.0105	0.0672	0.0019	0.0040
1979	7	0.1265	0.0111	0.0683	0.0078	0.0039
1979	8	0.1303	0.0114	0.0704	0.0090	0.0040
1979	9	0.1369	0.0119	0.0733	0.0136	0.0040
1979	10	0.1640	0.0144	0.0860	0.0345	0.0040
1979	11	0.1485	0.0136	0.0774	0.0308	0.0038
1979	12	0.1499	0.0137	0.0784	0.0316	0.0038

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Investment Model		Price-Investment Model		
		CON	CC	CON	CO	IN
1980	1	0.1695	0.0156	0.0926	0.0418	0.0040
1980	2	0.1935	0.0163	0.0981	0.0572	0.0037
1980	3	0.2007	0.0166	0.1002	0.0603	0.0038
1980	4	0.1355	0.0089	0.0759	0.0133	0.0054
1980	5	0.1182	0.0086	0.0570	0.0112	0.0059
1980	6	0.1041	0.0087	0.0418	0.0075	0.0066
1980	7	0.1266	0.0081	0.0718	0.0143	0.0050
1980	8	0.1467	0.0071	0.1032	0.0046	0.0050
1980	9	0.1588	0.0067	0.1158	0.0020	0.0052
1980	10	0.1731	0.0065	0.1263	0.0033	0.0053
1980	11	0.1731	0.0065	0.1262	0.0033	0.0053
1980	12	0.1626	0.0071	0.1164	0.0043	0.0051
1981	1	0.1723	0.0069	0.1271	0.0043	0.0050
1981	2	0.1859	0.0079	0.1468	0.0089	0.0042
1981	3	0.1782	0.0075	0.1396	0.0047	0.0048
1981	4	0.1973	0.0079	0.1621	0.0158	0.0031
1981	5	0.1788	0.0075	0.1325	0.0210	0.0038
1981	6	0.1845	0.0076	0.1407	0.0234	0.0032
1981	7	0.2002	0.0076	0.1519	0.0360	0.0016
1981	8	0.2160	0.0080	0.1723	0.0402	0.0005
1981	9	0.2285	0.0080	0.1878	0.0423	-0.0003
1981	10	0.1954	0.0092	0.1506	0.0350	0.0022
1981	11	0.1475	0.0097	0.0995	0.0294	0.0056
1981	12	0.1704	0.0083	0.1290	0.0304	0.0036
1982	1	0.1741	0.0082	0.1357	0.0295	0.0033
1982	2	0.1707	0.0083	0.1320	0.0295	0.0035
1982	3	0.1651	0.0086	0.1244	0.0312	0.0038
1982	4	0.1566	0.0095	0.1132	0.0336	0.0043
1982	5	0.1581	0.0094	0.1148	0.0338	0.0042
1982	6	0.1719	0.0082	0.1426	0.0246	0.0033
1982	7	0.1581	0.0095	0.1226	0.0313	0.0041
1982	8	0.1357	0.0121	0.0877	0.0482	0.0046
1982	9	0.1153	0.0133	0.0672	0.0580	0.0050
1982	10	0.0997	0.0148	0.0536	0.0645	0.0053
1982	11	0.0998	0.0148	0.0538	0.0644	0.0053
1982	12	0.0969	0.0151	0.0519	0.0649	0.0054

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Labor Model		Atheoretical Model		
		CON	ER	CON	RCL	CC
1968	1	-0.9996	-0.0198	-0.0408	0.1814	-0.0787
1968	2	-0.5136	-24.4979	0.9241	-0.2125	-1.4862
1968	3	-0.4617	12.5865	4.8201	1.6019	0.6767
1968	4	0.0776	1.8699	-1.1403	-0.4148	-0.8312
1968	5	0.0814	1.8660	-1.1371	-0.4117	-0.8289
1968	6	-0.0922	1.8768	-0.7598	-0.2666	-0.7187
1968	7	-0.2697	1.7089	-0.5946	-0.0239	-0.4578
1968	8	-0.2299	1.8798	-0.5076	-0.0145	-0.4317
1968	9	-0.1181	2.2447	-0.2073	0.0319	-0.3045
1968	10	0.0332	2.2289	-0.1018	0.0567	-0.3008
1968	11	0.1965	1.6210	0.0470	0.0712	-0.2794
1968	12	0.3593	1.1262	0.2734	0.0757	-0.1889
1969	1	0.4120	0.9028	0.2814	0.0784	-0.1847
1969	2	0.4234	0.9096	0.2891	0.0779	-0.1926
1969	3	0.4437	0.9689	0.3026	0.0781	-0.1968
1969	4	0.2668	0.7452	0.1516	-0.0034	-0.3712
1969	5	0.3818	1.3885	0.3250	-0.0188	-0.3349
1969	6	0.2880	1.0601	0.2354	0.0018	-0.2315
1969	7	0.2605	0.9677	0.2125	0.0044	-0.1999
1969	8	0.2852	0.9276	0.2307	0.0087	-0.2027
1969	9	0.3682	1.2634	0.3385	0.0250	-0.0409
1969	10	0.2722	1.0504	0.2497	0.0421	-0.0360
1969	11	0.3006	1.1875	0.3080	0.0337	-0.0542
1969	12	0.3230	1.2045	0.3374	0.0303	-0.0621
1970	1	0.3259	1.2080	0.3403	0.0304	-0.0662
1970	2	0.2011	1.2361	0.2920	0.0512	-0.0738
1970	3	0.2218	1.2568	0.2944	0.0515	-0.0858
1970	4	0.2875	1.1591	0.3435	0.0556	-0.0419
1970	5	0.3563	0.6993	0.3871	0.0576	-0.0159
1970	6	0.2847	1.1265	0.3501	0.0640	-0.0935
1970	7	0.2534	1.3638	0.3393	0.0675	-0.0907
1970	8	0.2668	0.9908	0.3532	0.0550	-0.1174
1970	9	0.2613	1.2349	0.3509	0.0564	-0.1216
1970	10	0.2706	1.0307	0.3643	0.0540	-0.1381
1970	11	0.2626	1.3632	0.3331	0.0594	-0.1609
1970	12	0.2627	1.2134	0.3407	0.0566	-0.1913

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Labor Model		Atheoretical Model		
		CON	ER	CON	RCL	CC
1971	1	0.2604	1.3395	0.3401	0.0574	-0.1909
1971	2	0.2608	1.0761	0.3346	0.0530	-0.0148
1971	3	0.2622	1.1725	0.3334	0.0542	-0.0489
1971	4	0.2712	0.9677	0.3564	0.0472	-0.0784
1971	5	0.2690	0.8145	0.3662	0.0435	-0.0806
1971	6	0.2619	0.6572	0.3840	0.0450	-0.0667
1971	7	0.2617	0.6603	0.3830	0.0451	-0.0662
1971	8	0.2503	0.7467	0.3487	0.0414	-0.1301
1971	9	0.2519	0.7819	0.3210	0.0420	-0.1019
1971	10	0.2478	0.8007	0.3046	0.0431	-0.0792
1971	11	0.2545	0.7893	0.3123	0.0429	-0.0775
1971	12	0.2581	0.7855	0.3165	0.0431	-0.0771
1972	1	0.2682	0.7838	0.3269	0.0436	-0.0729
1972	2	0.2576	0.7734	0.3115	0.0420	-0.0800
1972	3	0.2672	0.7842	0.3182	0.0432	-0.0637
1972	4	0.2683	0.7854	0.3193	0.0432	-0.0638
1972	5	0.2398	0.7550	0.2873	0.0414	-0.0652
1972	6	0.2505	0.7740	0.2958	0.0419	-0.0655
1972	7	0.2325	0.7427	0.2809	0.0415	-0.0592
1972	8	0.2337	0.7449	0.2802	0.0414	-0.0589
1972	9	0.2449	0.7666	0.2890	0.0422	-0.0548
1972	10	0.2020	0.6848	0.2676	0.0404	-0.0705
1972	11	0.1694	0.6143	0.2433	0.0388	-0.0693
1972	12	0.1935	0.6598	0.2611	0.0406	-0.0603
1973	1	0.2021	0.6749	0.2660	0.0411	-0.0589
1973	2	0.2061	0.6834	0.2686	0.0414	-0.0577
1973	3	0.2030	0.6759	0.2645	0.0411	-0.0556
1973	4	0.2051	0.6797	0.2637	0.0410	-0.0556
1973	5	0.2152	0.7059	0.2634	0.0410	-0.0555
1973	6	0.2181	0.7142	0.2619	0.0409	-0.0542
1973	7	0.2342	0.7593	0.2792	0.0425	-0.0571
1973	8	0.1669	0.5805	0.2152	0.0398	-0.0119
1973	9	0.1024	0.3818	0.1566	0.0359	0.0117
1973	10	0.1179	0.4271	0.1635	0.0367	0.0098
1973	11	0.1064	0.3993	0.1488	0.0352	0.0119
1973	12	0.1160	0.4237	0.1587	0.0359	0.0167

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Labor Model		Atheoretical Model		
		CON	ER	CON	RCL	CC
1974	1	0.1209	0.4387	0.1620	0.0364	0.0187
1974	2	0.1248	0.4512	0.1679	0.0365	0.0121
1974	3	0.1379	0.4899	0.1973	0.0333	-0.0293
1974	4	0.1485	0.5200	0.2190	0.0323	-0.0440
1974	5	0.1412	0.4996	0.2122	0.0317	-0.0474
1974	6	0.1427	0.5033	0.2150	0.0314	-0.0525
1974	7	0.1501	0.5153	0.2232	0.0319	-0.0546
1974	8	0.1633	0.5375	0.2436	0.0308	-0.0654
1974	9	0.1486	0.5180	0.2339	0.0317	-0.0604
1974	10	0.1142	0.4700	0.2084	0.0345	-0.0303
1974	11	0.1009	0.4614	0.1897	0.0344	-0.0075
1974	12	0.0985	0.4608	0.1879	0.0346	-0.0051
1975	1	0.0901	0.4628	0.1862	0.0357	0.0058
1975	2	0.0835	0.4720	0.1851	0.0365	0.0143
1975	3	0.1283	0.3531	0.2074	0.0252	-0.0487
1975	4	0.1443	0.2892	0.2146	0.0217	-0.0425
1975	5	0.1348	0.3368	0.2125	0.0238	-0.0366
1975	6	0.1313	0.3554	0.2121	0.0242	-0.0407
1975	7	0.1458	0.2778	0.2224	0.0226	-0.0067
1975	8	0.1521	0.2482	0.2250	0.0216	0.0172
1975	9	0.1618	0.1980	0.2308	0.0198	0.0219
1975	10	0.1443	0.2632	0.2233	0.0218	-0.0015
1975	11	0.1535	0.2364	0.2264	0.0197	0.0037
1975	12	0.1425	0.2692	0.2186	0.0213	0.0094
1976	1	0.1411	0.2733	0.2176	0.0214	0.0083
1976	2	0.1420	0.2707	0.2179	0.0212	0.0077
1976	3	0.1376	0.2816	0.2136	0.0214	0.0025
1976	4	0.1421	0.2774	0.2160	0.0208	0.0053
1976	5	0.1522	0.2716	0.2238	0.0201	0.0070
1976	6	0.1423	0.2740	0.2166	0.0204	0.0057
1976	7	0.1433	0.2743	0.2174	0.0203	0.0047
1976	8	0.1268	0.2651	0.2071	0.0208	0.0056
1976	9	0.1231	0.2650	0.2044	0.0209	0.0073
1976	10	0.1210	0.2638	0.2029	0.0210	0.0080
1976	11	0.1018	0.2533	0.1902	0.0220	0.0225
1976	12	0.0852	0.2485	0.1839	0.0233	0.0350

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Labor Model		Atheoretical Model		
		CON	ER	CON	RCL	CC
1977	1	0.1075	0.2545	0.2036	0.0235	0.0525
1977	2	0.1135	0.2579	0.2096	0.0237	0.0539
1977	3	0.1128	0.2575	0.2091	0.0237	0.0540
1977	4	0.1127	0.2574	0.2091	0.0237	0.0540
1977	5	0.1098	0.2547	0.2065	0.0236	0.0525
1977	6	0.1011	0.2434	0.1955	0.0231	0.0534
1977	7	0.1065	0.2519	0.2012	0.0231	0.0524
1977	8	0.0930	0.2272	0.1870	0.0225	0.0560
1977	9	0.0969	0.2342	0.1911	0.0227	0.0567
1977	10	0.1038	0.2463	0.1996	0.0228	0.0513
1977	11	0.1029	0.2447	0.1981	0.0227	0.0510
1977	12	0.1127	0.2638	0.2101	0.0234	0.0506
1978	1	0.1197	0.2783	0.2192	0.0239	0.0496
1978	2	0.1214	0.2825	0.2215	0.0241	0.0499
1978	3	0.1244	0.2906	0.2267	0.0243	0.0443
1978	4	0.1262	0.2952	0.2283	0.0245	0.0450
1978	5	0.1298	0.3046	0.2337	0.0249	0.0470
1978	6	0.1342	0.3159	0.2394	0.0256	0.0548
1978	7	0.1265	0.2930	0.2289	0.0248	0.0583
1978	8	0.1121	0.2487	0.1978	0.0213	0.0525
1978	9	0.1163	0.2629	0.2039	0.0220	0.0537
1978	10	0.1222	0.2797	0.2131	0.0228	0.0553
1978	11	0.1149	0.2583	0.2016	0.0218	0.0563
1978	12	0.1188	0.2695	0.2053	0.0223	0.0585
1979	1	0.1127	0.2505	0.1948	0.0214	0.0616
1979	2	0.0704	0.1140	0.1155	0.0133	0.0429
1979	3	0.0909	0.1750	0.1535	0.0155	0.0192
1979	4	0.0940	0.1850	0.1592	0.0160	0.0192
1979	5	0.0873	0.1622	0.1415	0.0123	-0.0092
1979	6	0.0780	0.1337	0.1223	0.0116	0.0265
1979	7	0.0814	0.1409	0.1265	0.0123	0.0319
1979	8	0.0839	0.1460	0.1304	0.0125	0.0302
1979	9	0.0880	0.1550	0.1367	0.0131	0.0328
1979	10	0.1057	0.1963	0.1644	0.0155	0.0268
1979	11	0.0950	0.1795	0.1496	0.0151	0.0383
1979	12	0.0959	0.1815	0.1511	0.0152	0.0380

TABLE XX (continued)

COEFFICIENTS FOR THE FORECASTING MODELS

YR	MO	Labor Model		Atheoretical Model		
		CON	ER	CON	RCL	CC
1980	1	0.1125	0.2084	0.1705	0.0171	0.0391
1980	2	0.1268	0.2302	0.1938	0.0182	0.0481
1980	3	0.1309	0.2373	0.2005	0.0186	0.0514
1980	4	0.0832	0.1799	0.1377	0.0124	0.0829
1980	5	0.0695	0.1664	0.1269	0.0134	0.1100
1980	6	0.0569	0.1615	0.1199	0.0141	0.1269
1980	7	0.0787	0.1584	0.1414	0.0138	0.1319
1980	8	0.1000	0.1496	0.1553	0.0138	0.1468
1980	9	0.1154	0.1408	0.1610	0.0141	0.1551
1980	10	0.1352	0.1282	0.1710	0.0143	0.1627
1980	11	0.1352	0.1282	0.1708	0.0143	0.1625
1980	12	0.1226	0.1361	0.1588	0.0147	0.1568
1981	1	0.1353	0.1286	0.1693	0.0144	0.1548
1981	2	0.1582	0.1162	0.1873	0.0149	0.1400
1981	3	0.1488	0.1216	0.1775	0.0138	0.1279
1981	4	0.1755	0.1086	0.1908	0.0147	0.1391
1981	5	0.1541	0.1184	0.1700	0.0132	0.1186
1981	6	0.1613	0.1161	0.1757	0.0133	0.1186
1981	7	0.1773	0.1148	0.1997	0.0116	0.0831
1981	8	0.1965	0.1106	0.2144	0.0121	0.0871
1981	9	0.2138	0.0989	0.2267	0.0122	0.0876
1981	10	0.1761	0.1202	0.1970	0.0140	0.1033
1981	11	0.1206	0.1569	0.1604	0.0159	0.1358
1981	12	0.1500	0.1264	0.1867	0.0129	0.1109
1982	1	0.1550	0.1216	0.1905	0.0128	0.1100
1982	2	0.1508	0.1261	0.1880	0.0130	0.1118
1982	3	0.1436	0.1358	0.1852	0.0134	0.1167
1982	4	0.1326	0.1509	0.1742	0.0142	0.1112
1982	5	0.1348	0.1479	0.1758	0.0140	0.1108
1982	6	0.1544	0.1191	0.1903	0.0121	0.0990
1982	7	0.1367	0.1462	0.1743	0.0133	0.0936
1982	8	0.1077	0.1854	0.1577	0.0160	0.1046
1982	9	0.0857	0.2190	0.1374	0.0172	0.1049
1982	10	0.0686	0.2461	0.1198	0.0186	0.1004
1982	11	0.0687	0.2460	0.1198	0.0186	0.1004
1982	12	0.0657	0.2510	0.1178	0.0189	0.1019