THE EXPLANATION OF STOCK PRICE MOVEMENTS USING A SINGLE EQUATION NON-LINEAR MODEL

By

KELLY RICHARD GUIOU

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1968

Submitted to the Faculty
of the Graduate College
of the Oklahoma State University
in partial fulfillment of the
requirements for the Degree of
MASTER OF BUSINESS ADMINISTRATION
July, 1970

THE EXPLANATION OF STOCK PRICE

MOVEMENTS USING A SINGLE

EQUATION NON-LINEAR

MODEL

Thesis Approved:

George E. Pinches
Thesis Adviser

Winfield P. Betty

D. Durham
Dean of the Graduate College

PREFACE

This paper will attempt to show grounds for the application of a sophisticated, quantitative model to the prediction of securities prices. Time and resources have limited the scope of the study, and thus the model will not be tested. This paper is meant to accomplish the challenging of other individuals to further the present study. This study by no means exhausted the area of study, and I realize that it has only scratched the surface. It has, however, shown that there may be many applications for techniques such as the present.

The model proposed in this paper corrects interrelationships between variables which have been assumed to be absent in most financial research. In some cases this correction may have meant the difference between the success or failure of many of the linear models proposed by others. With this technique, the individual is equipped to handle more statistically complicated problems. Securities analysis has been dominated by the Random Walk Hypothesis for many years, as it is the only technique which has gained any statistical support.

My thesis has been that much of the reason for the lack of statistical backing for technical and/or fundamental analysis is the lack of a powerful enough statistical tool. The present study hopes to show that the distributive lag model coupled with the combining of the two disciplines, Technical and Fundamental Analysis, will show that predictive models may have validity. Technical and Fundamental Analysis are mutually reinforcing and together they form a strong coalition. (Technical Analysis takes the demand for a product into account, and Fundamental Analysis establishes the value of the security.) In economic research if the individual knows the value of and the demand for a product, he can anticipate the price of that product.

I believe that a coalition between these two conceptual areas will give us more ability to predict the value and demand of a security. The distributive lag model developed by James E. Martin gives us the ability to do this predicting. Thus, this study proposes that a fully developed model based on the distributive lag principles will provide good predictive capability.

I would like to thank two individuals on the faculty of Oklahoma State University for their help and understanding. Dr. John R. Franzman was of immense help in the statistical development and computerization of this study's model; while Dr. George F. Pinches helped with the financial basis for the study. An extra thanks goes to Dr. Pinches, my adviser, for his help and patience while I was writing this paper. I would also like to thank my wife for help in editing, typing, and support during moments when I thought all was lost. Thanks goes to my typist, Raine Bridgewater, for her patience and help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER I

INTRODUCTION

General

Two theoretical fields of investment analysis, fundamental and technical, have been emphasized as approaches to investment decision making. These theoretical tools have been formed into two different fields of thought, and too often the thought of crossing from one field into the other has been rejected. The individual investor takes the factors of both fields into account before making his decision, whether by actual study or by intuitive thought. This study will include factors from both of these conceptual fields of investment analysis, in the hope of identifying some of the important factors from both fields which enter into the investor's decision process.

Purpose

This paper will attempt the development of a quantitative, computerized model for explaining securities price changes using factors from both fundamental and technical analysis. A non-linear distributive lag program which corrects for the auto-regressive errors caused by lagging certain independent variables will be used in the development of this quantitative model. The factors which will be lagged in this study will be technical in nature and thus are related to the demand for the stock. The objective of this paper will be to prove that a non-linear distributive lag model using both fundamental and technical variables can lead to a partial explanation of stock price changes. This study intends to show that the model includes many of the important variables, and that these variables lead to a significantly high coefficient of determination.

Methodology

Attempting the development of such a model requires constant attention to the fine points of statistical analysis. In a paper of this nature, one must first identify the factors which are of importance in such a model. In doing this the individual, in a way, biases the study since he has the tendency to put priority on those variables he deems as important. After the variables have been selected one must test these factors for significance; throwing out those which are not significant to the predetermined level. New variables must be added to the test group of variables to supplement those already in the study, and to take the place of those non-significant variables in order to increase the coefficient of determination.

A test group must be chosen, at this point, which is representative of the population to be studied. The test group for this study consists of thirty companies (Appendix A) chosen at random from the Standard and Poor's Composite 500 Index. These thirty companies represent twenty-nine different industrial groups, and give the study a wide representation of the population. The random selection of these companies will eliminate the bias associated with the selection of companies for the qualities desired in the study output.

The development of a truly meaningful model requires that proper quantitative techniques be used. In this study we will use two techniques: a standard linear regression method and a non-linear regressive method which corrects for the auto-regressive errors. The first technique, a linear regressive method, will be used to test the significance and intercorrelation of the variables, and to get the needed estimates of the parameters for each variable for the start vector analysis of the non-linear technique. The non-linear technique will be used in the development of our model, since with this type of mathematical model the fluctuations and interrelations of all the variables can be taken into account, and thus it can increase the explanatory capability of the model. This adjustment process is particularly good for relationships developing between two different periods' observations of the same variable. The development of a technique which corrects for these relationships should be widely researched in economic and financial analysis since so much in these areas of study is related to time series studies. This technique is being used in more and more areas of economic study, and this study is undertaken with the hope of furthering the knowledge needed for its use in financial analysis.

Plan of Organization

In Chapter II the principles of random walk, fundamental analysis, and technical analysis are examined. This chapter will give the financial basis of the study, but it will not attempt to cover any of these conceptual fields in great depth. Emphasis will be given those studies which have used methods similar to those employed in this study. This paper will attempt to show both the strong and weak points of these studies.

The quantitative techniques and program used in this study will be reviewed in Chapter III. The backgrounds and applications of the two quantitative models will be discussed and details of the development of both the regression program and James E. Martin's (autocorrelation program) will be given in order to provide the needed information about the use of these models.[1] A full explanation and discussion will be given in Chapter IV of all the various tests and statistical measures which were used in evaluating the results of the study and the application of the model in the future.

Limitations

In studies of the securities' markets and other economic factors limitations are commonplace, in that they are caused by the very nature of these fields of study. This study was limited to the large listed companies which are included in the Standard and Poor's 500 Composite Index, from which we drew our random sample of thirty companies, and thus this model cannot be expected to apply to all firms without further testing. The model also has the weakness of being applicable to the test companies only for the period of the study (January 1, 1966 to March 1, 1969,) and cannot be applied to other periods without adjustments. Time and finances limited this study since they placed constraints on the number of companies studied, the time period considered, the number of variables included, and the amount of computer time available, to name but a few of the limitations imposed by these constraints. More thorough testing and development of this model and others like it should be in order, and this development and testing may lead to a model with more widespread application.

The present study has value, in that, it furthers the knowledge of investment theory by proposing a quantitative model containing both fundamental and technical variables. The combination of variables from both of these areas may show the way to the rejection of the random walk theory. This study will strive to give some statistical basis for the acceptance of fundamental and technical analysis, and challenge someone to further study the development of a truly valid and highly explanatory model.

FOOTNOTES

[1]James E. Martin, "Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear Single Equation Models," Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968, p. 2-3.

CHAPTER II

A REVIEW OF INVESTMENT THEORY

Members of an affluent society have a great deal of money to spend and invest. More and more people have entered the capital market with their excess funds, and thus there is a growing demand for knowledge about stocks and other securities. This demand for knowledge about security analysis has led to the development of three basic areas of thought and research: random walk hypothesis, fundamental or security analysis, and technical analysis. These conceptual areas give each individual a reference point from which to develop his own theory of security analysis. The formalized structuring of these areas is rather new to the market structure, as most of the materials upon which they are based date back only a couple of centuries. The development of the computer for application to securities analysis was helpful, since the data needed to statistically prove or disprove the development of any theory is beyond the individual's computational ability. This chapter will cover the three conceptual divisions of investment theory in the order they were listed above.

The Random Walk Hypothesis

This theory casts doubt on the whole field of security analysis, in that it leaves no area open for the predicting and describing of the behavior of stock prices.

> Random walk theorists usually start from the premise that the major security exchanges are good examples of efficient markets. An efficient market is defined as a market where there are large numbers of rational profit maximizers actively competing, with each trying to predict future market values of individual securities, and where important current information is almost freely available to all participants.[1]

If the securities' markets act in the way suggested by the Random Walk Hypothesis, we can assume that the market price reflects all the information available about the given stock value, and that a purely competitive market exists. In this type of market the price is based on the current information, thus past information and prices have no bearing whatsoever on the current period's price. If the random walk hypothesis is a true description of reality, nothing could be gained by sophisticated investment analysis. Instead the individual investor could make as high a return on his capital by a random buy and hold policy as could be earned by any other system.

The random walk hypothesis does not actually suggest a purely competitive market, but rather suggests that some inperfections do exist. It suggests that the actual price at any time is a random deviation from that stock's intrinsic value, whereas the efficient market suggests that the actual price is a good estimate of its intrinsic value. This random fluctuation of price will be caused by the competition between buyers which reduces any systematic tendencies that exist. Competition of this kind causes information to be instantaneously reflected in the current prices; but since no new information can be fully understood, the instantaneous adjustment of prices has two implications.

> First, actual prices will initially overadjust to changes in intrinsic value as often as they will underadjust. Second, the lag in the complete adjustment of actual prices to successive new intrinsic values will itself be an independent, random variable with the adjustment of actual prices sometimes preceding the occurence of the event which is the basis of the change in intrinsic values (i.e., when the event is anticipated by the market before it actually occurs) and sometime following.[2]

The random walk theory suggests that a series of stock prices has no memory, thus a price at time t cannot be used to predict the price at time t + 1.

The testing of this theory is a problem in itself, but two techniques have been developed to accomplish this: (a) by running a series of prices and coming up with the serial correlation coefficients and testing these for significance, and (b) testing the different mechanical trading rules for predictive capability. The first method is more widely characteristic of the available research. The body of current published research tends to support the random walk hypothesis. Cootner suggests that systematic changes do exist, but to test them for significance is beyond the ability of the currently used statistical tools of financial analysis.[3] Any time series is filled with autocorrelation and the standard linear regressive technique cannot take this correlation between independent variables into account. Most of the tools used in this type of research are not powerful enough to detect the significance levels of autocorrelation associated with time series studies, but it is not enough to say that present techniques are not significantly powerful enough unless a model can be offered which identifies this autocorrelation. The identification of this autocorrelation is the aim of the present study.

The second technique, testing the different mechanical trading rules for significant predictability, is less clear cut, and the question arises as to whether the disproving of these rules is justification for the acceptance of the random walk theory. The lack of prediction by any mechanical trading model does not prove the truth of the random walk hypothesis, but only proves the invalidity of the model itself. Too much freedom has been taken in this type of study to prove something that the model studied is actually designed to disprove. The error in prediction may have been an error in the rules of the model and not in the model itself, and this would not justify the acceptance of the hypothesis that the random walk theory was a good estimate of reality. Seelinfreund, Parker and VanHorne suggest in their study of a quadratic formula for the prediction of stock price that the model was applied to an extensive series of daily price data and the results are found to be consistent with the random walk theory of stock price behavior.[4] But they also stated that one study of a mechanical trading rule is not enough to prove or disprove the random walk hypothesis. The proving or disproving of this theory will take many studies; at the present time all we can hope to do is to further the knowledge available. The random walk theory may gain more support while we are increasing the knowledge of stock price behavior, but the real proof will lie in proving that no significant systematic tendencies exist, and not in whether or not a mechanical system is valid.

It is not enough to prove that systematic price series exist. In order for this to support the development of mechanical trading rules, there must also be significant profits which can be made by using such a rule. These trading techniques must make a level of profits which, after commissions are subtracted, will yield a higher net profit than can be expected from a random buy and hold policy. Thus any mechanical system which will not bring forth such a yield is not worth the effort put into its development or the investment of funds using the system. The explaining and developing of these mechanical trading rules is the basic objective of the other two fields of Financial Investment Analysis.

Fundamental Securities Analysis

Fundamental Analysis is the taking of all available information, related to the economic evaluation of the security, and attempting to project an intrinsic value for that security. The fundamentalist believes that the price of a security always strives toward its intrinsic value, and that profit potential exists when the intrinsic value deviates significantly from the market price. This theory advocates the buying and selling of securities based on the deviations between the intrinsic value and their market price. This theory advocates the buying and selling of securities based on the deviations between the intrinsic value and their market price. Fundamental analysis is not too far removed from the random walk hypothesis; since the only real point of disagreement is that in fundamental analysis information tends to be restrictive and thus there is profit to be made by the prediction of price. Fama supports this in his article on random walks.

> The analyst will do better than the investor who follows a simple buy and hold policy as long as he can more quickly identify discrepancies between actual prices and intrinsic values than other analysts and investors, and if he is better able to predict the occurrence of important events and evaluate their effects on intrinsic values.[5]

However, he cautions that the real proof lies in the fundamentalist being able to produce profits greater than a random buy and hold policy, after all incremental costs are subtracted. Several articles on the random walk hypothesis state that a few individuals practicing good fundamental analysis are needed in order that the market will retain its random character. They state ~~that~~ these investors perform the function of transferring information and thus helping the stock prices to reflect all the current information. These investors, they say, must be able to make returns above the level of the random buy and hold policy to pay them for their services. This suggestion weakens the random walk proponents own thesis, in that they open the door for profits greater than their own policy would produce, and turns their own challenge to produce higher profits back upon themselves.

The fundamental analyst is concerned with every bit of information which might help in predicting the economic growth of the firm in question. Some of the common fundamental tools are: sales, earnings, dividends, management, plant expansion, return on assets, share of market, sales area expansion, research and development, etc. Any information which may affect any of these areas must be considered if the fundamentalist is to produce an accurate estimation of the value of the security. In order for the fundamental analyst to project the intrinsic value of the security profitably, he must have access to first-hand information and must have the resources to take advantage of the information gathered. The insights of analysts are of no real value unless they are borne out by the market, that is, unless the price moves in the predicted direction. In order for the fundamentalist to claim success he must be able to produce these insights consistently and with substantial profits.

Many studies have been done on the ability of the fundamental analyst to produce meaningful predictions. Berstein lists the following tests of a true growth company: a true growth must have a better-than-average uptrend in sales and earnings, but it must also meet the following criterion:

> (1) The uptrend in earnings should be relatively smooth. Earnings need not rise every single year, but they should show an increase in at least as many years as the "average" company's earnings increase. (2) What is true of earnings should also be true, in most cases, of dividends. (3) Return on net worth should be maintained; and if the dividend payout is abnormally low, then return on net worth should actually be rising. (4) Increases in earnings and/or net worth should of course reflect an

> increase in the physical volume of output or at least a beneficial shift in product mix - rather than merely larger dollar results reflecting a rising price level.[6]

A firm can have any one of these and give the outward appearance of being a growth company as long as it has rising sales and earnings. The real test is whether the combination of all four of these factors exists with rising sales and earnings. The fundamental analyst must take all these factors into consideration before projecting an intrinsic value, or his analysis will be faulty and inconsistent.

Many individuals equate the concept of growth company with that of a growth stock, but the two are not related. A growth stock is identified only by hindsight and is simply a stock which underwent a rapid increase in price. A growth company on the other hand is a company whose stock, because of good management, etc. increased in value at a reasonable ratio to increase in earning power over time and is likely to appreciate in the future.[7] The concept of a growth company is dynamic and reflects from an overall good corporate structure and perspective. The security of a growing company is more likely to withstand the downtrends of the market and grow more rapidly in an upward market trend.

This is not to suggest that the job of the fundamentalist is easy when actually it is very difficult. The analyst must project sales, earnings, investments, dividends, etc. far into the future, and the methods of accomplishing these projections are themselves weak. Accounting's experience with cost and profit budgeting gives us an example. Projections of costs and profits six months to a year into the future can be fairly accurate when we are given all the available information, but any projections of periods further into the future is not recommended as being accurate in cost accounting. The same limitations are true in fundamental prediction, and Holt states that the uncertainty of forecasting earnings in the distant future is a much more important consideration for growth stocks than for nongrowth stocks.[8] Holt does show that by assuming a constant growth rate that a trend diagram can be drawn which will offer a framework for evaluating growth companies. The technique is called a Nomograph and it gives the expected duration of future growth given the existence of a constant dividend policy.

It must be noted that the current discounting methods are not significantly powerful enough to give the full future effect of fundamental information. As more sophisticated techniques become available, fundamental analysis and its sister predictive conceptual division, technical analysis, will achieve greater acceptance in the world of financial investment theory. Both of these disciplines are not structured for the average investor, since the techniques and models needed for success in these fields is beyond the average investor's comprehension. After these fields are proven the average investor will be able to use some of the simpler models which will spin off the more complicated ones. Fundamental analysis will have less application here as it requires the most advanced information available for the prediction of a security's intrinsic value, and the small investor does not have the facilities to collect and use this information. This discipline cannot be discounted as it will take a major part in the overthrow of the random walk theory, if it is overthrown, but a combination of both fundamental analysis and technical analysis is a stronger approach to the prediction of securities prices.

Technical Financial Analysis

There has been a lot of criticism of technical analysis since its conception. The random walk theorists suggest that you cannot use past supply and demand figures to present a competent prediction of the future price of a security, and the fundamentalist states that you cannot predict the price of the stock without predicting its value. In spite of this criticism more and more investors are turning to technical models for the prediction of securities prices. This is true because anyone can develop their own model by using the principles of this discipline to set the rules and constraints the way they want. Not all of these models are good ones, for various reasons, and the ones which one hears about are not those which are likely to be predictive. If a model is good, then it is best to keep it to oneself, since if it becomes general knowledge the predictability and/or profitability may be lost.

Technical analysis is, in essence, the recording of the actual history of trading (including both price movement and the volume of transactions) for one stock or a group of equities, and deducing the future trend from this historical analysis.

> Technical theory can be summarized as follows: (1) Market value is determined solely by the interaction of supply and demand. (2) Supply and demand are governed by numerous factors, both rational and irrational. (3) Disregarding minor fluctuations in the market, stock prices tend to move in trends which persist for an appreciable length of time. (4) Changes in trends are caused by the shifts in supply and demand relationships.[9]

There are more individual technical schemes than can be related here, as each investor modifies a model to fit his own personal needs. The basis of all technical models is the old adage that history tends to repeat itself, thus the use of historical data is justified in the prediction of securities' prices. The technician suggests that, even if the fundamentalist did predict the intrinsic value of a security and if this belief were not shared by others, the chance of it coming true is little and none. Supply and demand must be considered in any predictive model for securities' prices.

There has been little proof set forth which backs technical analysis, and most of the current research tends to discredit and/or dismiss it completely. Levy lists four major criticisms of technical analysis.

> First, it is contended that the behavior of the stock market in the past may not be indicative of its behavior in the years to come. Second, that technical traders acting on the results of their studies tend to create the very patterns and trends which they claim have predictive significance. Third, that, if technical analysis is continually successful, an influx of technical traders will neutralize whatever profit potential exists. There are four reasons why this criticism is not fatal to technical analysis. (1) It is quite possible that extremely successful technical systems have been developed but, for this very reason, have not been publicized. (2) It is likely that those who are not engaged in technical analysis would be reluctant to believe the claims of successful technicians. (3) To the extent that technical analysis may depend in part upon the use of electronic computers and sophisticated mathematical techniques, both the expense and the requisite training and knowledge will prevent its exploitation by the majority. (4) Most people will not take the time, patience, or effort to do the work necessary to achieve results by technical analysis. Fourth, technical analysis is too subjective.[10]

The critical point the technician and the fundamentalist have in common is one of taking data from past and present periods to project future prices. The two areas use different types of information and this has been a major stumbling stone to any combinations of the two. The technician is faced with easily available information for any time period he wishes to use, but the fundamentalist is faced by information which is published at different times, and in differing quantities. Thus, in this study it is felt that the use of two different types of data need not be a barrier between the two conceptual areas and the study hopes to show that much can be gained by the combining of these two areas of study.

Suggested Model

Alone neither of the latter two theoretical divisions will be able to gain enough support to overthrow the random walk hypothesis, but together they form a good coalition with one area covering the other's weak points. The problem which confronts those who wish to combine these disciplines is mainly one of structure. A model containing two completely different types of independent variables is more complex and harder with which to work. The technical factors are readily available, but must be entered in the way of a time series and thus the model must correct for the correlation between these independent variables as members of a time series. The fundamental variables are not as readily available, and their gathering often causes problems. There are three ways of handling these fundamental variables when they are to be entered in the same model as technical factors: (1) report them only on the day of their availability while on all other days they are carried as zero; (2) report them on the day of their availability and everyday until the next report as the same value; or (3) develop a program which will figure the value of that item for every day between reports. Each of these has its advantages and disadvantages and the most accurate is the third, but in order to use this method one must have unlimited time and resources.

The model proposed in this study will use a three-period lag on technical factors in order to establish the desired time series. The model used for this study will correct for the correlation between these lag periods and lagged variables by using a predesigned program for estimating the error parameters associated with these lags. The fundamental factors of the study will be carried as the same value from one reporting date until the next, which is assuming a constant affect of these variables, because due to the lack of time and resources a program cannot be developed to figure each day's value. This is an additional limitation on the part of the fundamental factors. In the next chapter the quantitative background of the proposed model will be explained in depth.

FOOTNOTES

[1]Eugene F. Fama, "Random Walks in Stock Market Prices," Financial Analysts Journal, (September-October), 1965, p. 56.

[2]Ibid., p. 56.

[3]Paul H. Cootner, "Stock Prices: Random vs. Systematic Changes," Industrial Management Review, Vol. III, No. 2 (Spring, 1962), p. 44.

[4]Alan Seelinfreund, George C. Parker, and James C. Van Horne, "Stock Price Behavior and Trading," Journal of Financial and Quantitative Analysis, Vol. III (September, 1968), p. 265.

[5]Fama, p. 58.

[6]Peter L. Bernstein, "Growth Companies vs. Growth Stocks," Harvard Business Review, Vol. XXXIV, No. 5 (September-October, 1956), p. 93-94.

[7]Ibid., p. 98.

[8]Charles C. Holt, "The Influence of Growth Duration on Share Prices," The Journal of Finance, Vol. XVII, No. 3, (September, 1962), p. 475.

[9]Robert A. Levy, "Conceptual Foundations of Technical Analysis," Financial Analysts Journal, (July-August, 1966), p. 83.

[10]Ibid., p. 87-88.

CHAPTER III

DEVELOPMENT OF THE QUANTITATIVE MODEL

The gathering of data and its correct use within a quantitative model will determine the validity of that model. Thus precautions must be taken to guarantee the validity and applications of the information used. This will increase the chance of the study being free of errors and inconsistencies, which arise when inappropriate data units or types are used. Proper collection and use of a data base will save many hours of needless work and effort when it comes to correction and explanation of the study's findings.

Investment Data: Its Sources and Uses

Data on securities markets is readily available to all persons, but items must be chosen which have application to the particular study. Then care must be taken in adjusting and using this information. The availability of these statistics many times forces adjustments to be made in order to adapt them to the study. An example of this can be found within the present study, i.e., earnings could have been handled in any one of three ways: (1) Entered on the date of their publication as the full value but zero on all other days; (2) Entered on all days as the value at their last reported date; or (3) A different value could be computed for the effect of these earnings on each given day. Each of these would have a different effect upon the model's effectiveness. When choosing these data units the ones which will give the study the most validity must be chosen but yet will not make the data collecting more costly than the benefits to be gained.

In Appendix B a full list of the statistics and their sources is given. In this study more than one source was sought for each statistic in order to assure the availability of all the needed information. If a volume of the primary source were missing, the information was readily available from an alternate source.

This study proposes the use of two different types of data in the construction of a quantitative model for predicting securities prices and values. These variables, technical and fundamental, were explained in the previous chapter, but it must be recognized that the technical variables tend to be readily available while the fundamental variables tend to be less available and systematized. Thus, in this study, the fundamental variables will be reported at the quarterly reported value until the next report, while the technical variables were reported at their weekly value. This way all of the variables have values for all periods, but it is not the optimal technique in that it assumes that the quarterly value will have a constant effect on securities prices. The alternative approach of reporting the value only on the day of the report is based on the premise that they only affect the security on the day of their issue, and the limited time and resources precluded the development of a program to figure the exact effect of an earnings report on securities' prices. Next comes the job of developing, adjusting, proving, and using the proposed model.

Linear Regression Model

Linear regression is a quantitative technique commonly used in financial model making. This technique is often used in models involving a time series, but this is beyond its statistical capabilities. It was designed for the specific purpose of giving a value estimate or equation model given that the following assumptions were not violated: (1) linear relationships exist between independent variables, (2) there is no interaction between independent variables, and (3) there is an independence of the error terms (μ) between observations of any and all variables.[1] When a time series is included involving the proposed variables this automatically eliminates the linear model, since one cannot expect two time periods' values of the same variable to be completely unrelated. This would violate all the statistical theory known to man. This interrelation of time series variables was shown in a prior study and paper.

This study related basically to the same type of a model as the one proposed in the present study except that it suggested the use of a linear rather than a distributive lag technique. It covered the same thirty companies (Appendix A) which are included in the present study, but only ten variables were considered: market index, industry index, price, earnings, and dividends. These variables were chosen for their accessibility and they were used to try to prove that some degree of explanation could be achieved using a linear regressive model. The first three variable groups were involved in time series with the first two being lagged one period while price was lagged two periods. This prior study used the same computer linear regression model which will be used in the present study (BMDO3R) to estimate parameters and check for significance. The general form of the equation was:

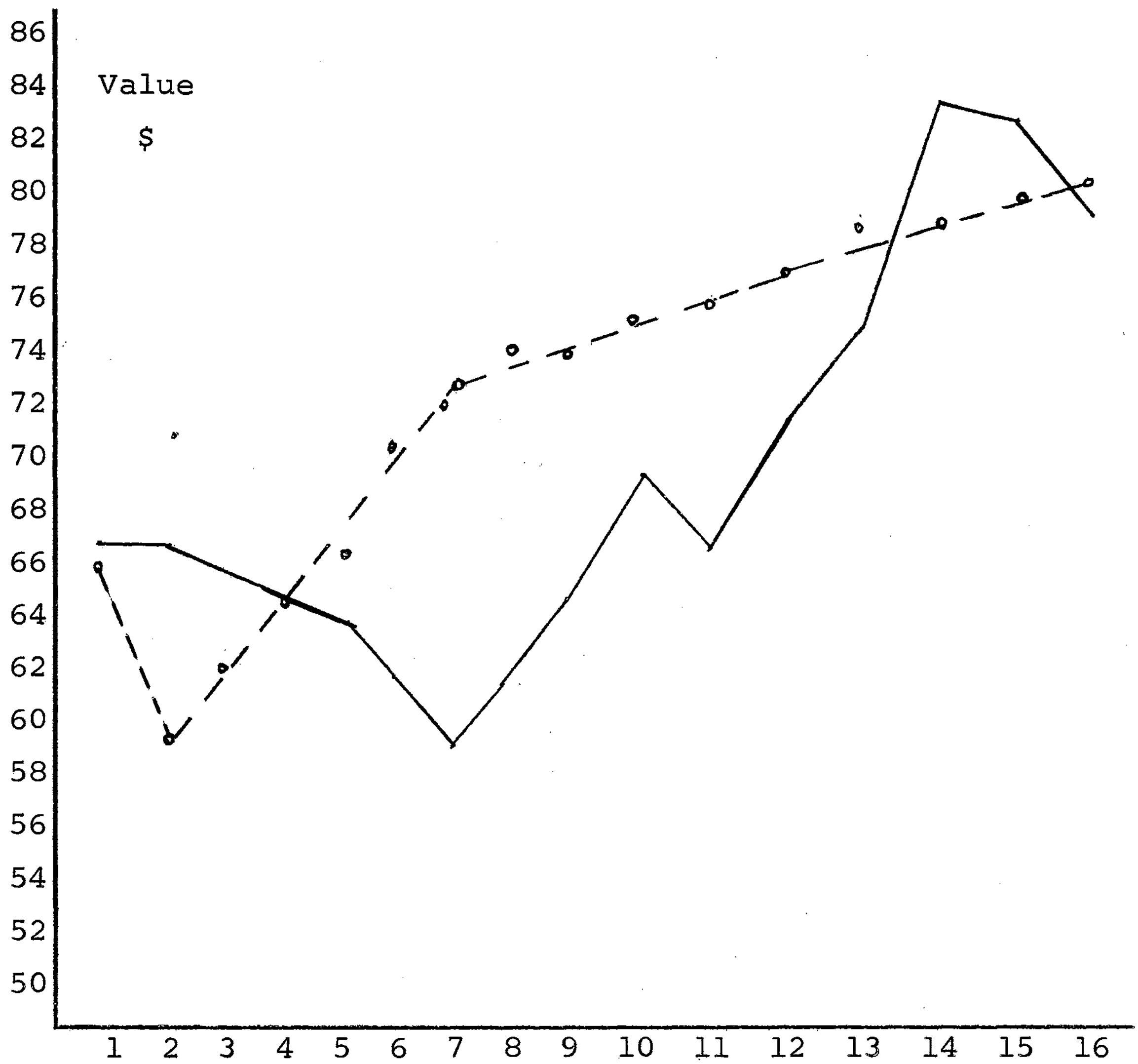
(1) $$P_t = A_o + b_1 P_{t-1} + b_2 P_{t-2} + b_3 E_t + b_4 F_t + \ldots$$

$$b_5 A_t + b_6 B_t + b_8 B_{t-1} + b_9 \text{ Dummy} + et$$

The assumptions listed earlier concerning linear regression were assumed to be true and the model above was considered to be a good estimate of price behavior for the purpose of the prior study.

This assumption had to be rejected when a graphical plot was made of several periods of a security's price behavior (Figure 1). It showed that the model explained only a small amount of the price fluctuation, and tended only to form a trend ignoring the fluctuations which it could not explain by its least squares technique. But this equation could not be completely abandoned, since all but two of the variables, market index at t and industry index at t, were significant at the .01 level (Table 1). The coefficient of determination was a low .3983, and showed that more variables needed to be added to increase the explanatory ability of the model. It was not hard at this point, to suspect that no true linear relationship exists, since there are undoubtedly relationships between lagged observations of the same variables but an alternate technique had to be found.

This called for the abandonment of any assumptions of a linear equation, and a new approach to the suggestion of a more sophisticated model. There were two possible approaches which could be taken: Spectral Analysis, or Distributive Lags. Spectral Analysis would establish a



----- Original Assumption of Linear Function of Price Movement Over Time

——— Price Movements Ignored by Linear Function of Price Movement

Figure 1. Relationship Between the Predicted and Actual Values Using a Linear Equation

TABLE I

THE IMPORTANT STATISTICS FROM THE LINEAR MODEL APPLIED IN A PRIOR STUDY

Variable	Regression Coefficient	Computed T Value	Prop. Var. Cum.
Pt-1	.03442*	12.40624	.12981
Pt-2	.03208*	11.45616	.01888
EPS	28.29495*	35.82999	.19153
ΔEPS	-10.42887*	-13.99233	.03149
II	.01532	1.26632	.01938
IIt-1	.04189*	3.39820	.00226
MI	.02822***	2.39893	.00162
MIt-1	.03914**	3.19543	.00158
ΔDiv	-2.68971*	-3.28808	.00176

Coefficient of Determination .3983

* Significant at the .001 level
** Significant at the .01 level
*** Significant at the .02 level

confidence interval around the Linear Regression Line, and would explain the fluctuation outside this range as random if no pattern existed and systematic if these fluctuations tended to pattern themselves. A Distributive Lag Model would tend to follow a path of estimation, which would closely approximate the plotted price path of the observed security price. In this study the Distributive Lag Model will be assumed to be the better approximation of stock price behavior. Equation 2 shows the model which was proposed by the previous study as an approximation of the security price movements. This model can be compared with the model developed in this study, and one may be able to evaluate the worthiness of the model suggested by the present study.

$$(2) \quad P_t = .12631 + .03442\, P_{t-1} + .03208\, P_{t-2} + 28.29495\, Et - 10.42887\, Ft + .01532\, A_t + .04189\, A_{t-1} + .02822\, A_t + .03914\, B_{t-1} - 2.68971\, \text{Dummy} - 6.1268$$

The present study is divorced from the simple linear regression model and is undertaking the use of the more statistically powerful Distributive Lag Model. However, due to time and a limited computer budget it was not practical to use the more sophisticated model for all the needed functions. Thus the simple linear technique will be used to estimate parameters needed in the Distributive Lag

Model, and to test the variables for significance. The linear model can perform both of these functions with some degree of accuracy. It is most accurate in the testing of the variables for significance, as this has to do with their linear relationships to the dependent variable.

To perform this function for the study, a canned program was chosen from the University of California at Los Angeles Biomedical Research Series. The specific program (BMDO3R) is a linear regression program which allows for up to 100 variables and 99,999 observations per variable. This program also gives the study the ability to test any size sub-sample from within the total data deck. Thus any portion of the data can be drawn out of the sample and tested against the outcomes of the sample tests. This gives a quick test of the conformity of the sub-samples with the over-all study. The output of the program gives the correlation matrix of all the variables (which can be used to check the correlation between variables), the coefficient of determination (showing the degree of explanation of the model), the correlation coefficient for each variable (which will be used to estimate the needed parameters), and the T-test value for each variable (used to test the variables for significance.) The two statistics which are of most importance for our further study is the correlation coefficients for each variable, and the T-test value for each variable, but we will use the other two in a supporting role. The correlation coefficient gives us the parameter estimate which we need for the start vector cards in the Distributive Lag Model. These are the values which start our model into operation; using the T-test value we can find out whether the variable is significant or insignificant by using a standard "t" table from a statistics textbook.

This study has used a standard linear regressive model as a tool in a more sophisticated model. The linear technique is by no means relegated to strictly a supporting role, for if the variables of a model fit the assumptions stated earlier it would be foolish to use a more powerful model than is needed to do the job. Thus the linear regression model is useful and cannot be discarded, but when the conditions are not those which call for linear regression, we must strive to recognize these conditions, and use other methods which are powerful enough to do the job. Based on this fact the present study chose to use the James E. Martin Distributive Two-Lag Model for the development of a single equation non-linear model.[2]

The Distributive Lag Model

The Distributive Two-Lag Model (Appendix C) is not without statistical backing, in fact it is a highly advanced form of distributive lag model, which has been set forth in the study of economics over the last few decades. The development of these non-linear models has become a meaningful part of economic research since Koyck and Nerlove suggested that most economic prediction is based on past time periods. Thus Nerlove suggests that some adjustment process must be developed to take into consideration the interrelationships of the independent variables and particularly, those which are members of a time series.[3] Nerlove recommends two equations to take account for this autoregressive behavior of independent variables, and they are developed, in equation form, in Appendix C. These equations give two basic ways of considering time series studies, and both of them are single lag type studies. The single lag study assumes that all the lagged independent variables belong to one group and are associated with one lag parameter even though the second parameter is computed. These studies and many others give a formal basis for the Two-Lag Model and several other authors suggest that a two-lag model may be more statistically significant, but its formal development has been undertaken mainly by James E. Martin.

Martin, in Computer Algorithms for Estimated the Parameters of Selected Classes of Non-Linear, Single Equation Models, suggests the use of not only his own two lag model, but also shows five other models which can be used from the same program.[4] It is important, he stresses, that in any study one uses the model which will give him the most efficient application to his research. The use of a tool which is more powerful than needed is as much a mistake as the use of one that will not perform the needed function. This is true because the study may get bogged down in the statistical output, and overlook the statistics which will give it the answers to the research in question.

The programs developed by Martin are aimed at giving economic studies better, more powerful tools with which to work. It has been used in countless tasks in the field of Agricultural Economics, but the major emphasis has been in the prediction and explanation of demand curves for agricultural products. It has performed this function admirably, but the program's application to financial research has not yet been implemented. Distributive Lag Models have been used in the financial field in the last decade, but these have been mainly single lag models of less statistical complexity than the models developed by either Nerlove or Martin. Koyck suggests the use of a single-lag model which computes only one error term in his study on the use of distributive lags in a study of capacity adjustments and internal investment possibilities related to the steel industry. He sets forth reasons for the lag in the adjustment process which causes distributive lags, Table II. Though his study was inconclusive, he suggested that it furthered man's knowledge and that it showed that a time series adjustment can lead to a more significantly powerful quantitative model.

TABLE II

CAUSES OF DISTRIBUTIVE LAGS

(A) Objective Reasons

(1) Technological Factors

(a) It takes time to adjust the production of investment and consumers' goods to the demand.

(b) The lifetime of durable goods, investment and consumers' goods causes resistence to complete and immediate adaptation to changed circumstances.

(c) The restriction in the behavior of entrepreneurs with respect to durable investment goods are of great importance in dynamic theory.

(2) Institutional factors -- lagged reaction caused by the rules observed, through law and custom.

(B) Subjective Reasons

(1) Imperfect knowledge of the market

(2) Psychological inertia of economic subjects

SOURCE: L. M. Koyck, Distributed Lags and Investment Analysis, Contributions to Economic Analysis, North Holland Publishing: Amsterdam, Holland, 1954.

Most of the other studies of distributive lags in the finance area are equally inconclusive, but they all consist of weaker statistical techniques than are available with the Distributive Lag Models developed by Nerlove and Martin. Thus these more sophisticated models, particularly the two-lag model, were studied as possible techniques for the present study. Nerlove listed the assumption of all lagged independent variables were in one group as a drawback in the application of his model.[6] Martin suggested the use of the Nerlove model if all the lagged independent variables were correlated between .50 and .75 with one another.[7] Since all of this study's variables are not correlated to that level, the two lag model will be used in this study in an attempt to develop a truly meaningful explanatory model.

The application of this model to the present study required very little alteration of Martin's proposed model. The assumptions required were nominal, that industry index and market index are related and should be included under the same lag parameter, and that company volume and market volume are related and should be included under the same lag parameter, and thus were thrown together to form the dummy section of the equation. The major implication and assumption was that contrary to the random walk hypothesis, past time periods can be used in the explanation and prediction of future price movements. This is qualified by the assumption that alone these technical factors cannot accomplish an accurate prediction of price, thus fundamental facts were added in order to support these technical factors. This model is set forth in Table III, and all variables used in the model are defined within this table. It will not answer all the problems, and its development in this paper may not be proven or tested, but it is a step in the direction of increasing man's knowledge, and by doing this it may challenge someone else to further the study and prove or disprove its hypothesis. If it accomplishes this task it will be worth more than the time and money invested in its development.

TABLE III

THE PROPOSED NON-LINEAR MODEL

$$
\begin{aligned}
P_t = {} & a_0(1-\mu)(1-\lambda)(1-\beta) + \sum_{i=1}^{A} a_i V_{it-1} - (\mu+\beta)\sum_{i=1}^{A} a_i V_{it-2} \\
& + \mu\beta \sum_{i=1}^{A} a_i V_{it-3} + \sum_{j=1}^{B} b_j U_{jt-1} - (\mu+\beta)\sum_{j=1}^{B} b_j U_{jt-2} + \\
& \mu\beta \sum_{j=1}^{B} b_j U_{jt-3} + \sum_{k=1}^{C} c_k A_{kt-1} - (\lambda+\beta)\sum_{k=1}^{C} c_k A_{kt-2} + \\
& \lambda\beta \sum_{k=1}^{C} c_k A_{kt-3} + \sum_{l=1}^{D} d_l B_{lt-1} - (\lambda+\beta)\sum_{l=1}^{D} d_l B_{lt-2} + \\
& \lambda\beta \sum_{l=1}^{D} d_l B_{lt-3} + \sum_{m=1}^{O} o_m E_{mt-1} + \sum_{m=1}^{P} p_m F_{mt-1} + \sum_{m=1}^{Q} q_m G_{mt-1} \\
& + \sum_{m=1}^{T} t_m R_{mt-1} + \sum_{m=1}^{U} u_m S_{mt-1} + \sum_{m=1}^{V} v_m X_{mt-1} + \sum_{m=1}^{W} w_m Y_{mt-1} \\
& + \sum_{m=1}^{Z} z_m D_{mt-1} + \sum_{m=1}^{X} x_m P_{mt-4} + (\lambda+\mu+\beta) P_{t-1} - \\
& [(\lambda+\mu)\beta + \lambda\mu] P_{t-2} + \lambda\mu\beta P_{t-3} + e_t
\end{aligned}
$$

TABLE III (Continued)

WHERE:

P_{t-m} = the current and lagged values of the dependent variable (m = 0, 1, 2, 3)

Vi_{t-m} = the lagged values of the company volume associated with lag parameter μ (m = 1, 2, 3)

Vj_{t-m} = the lagged values of the market volume associated with lag parameter μ (m = 1, 2, 3)

Ak_{t-m} = the lagged values of the Industry Index associated with lag parameter λ (m = 1, 2, 3)

Bl_{t-m} = the lagged values of the Market Index associated with lag parameter λ (m = 1, 2, 3)

Em_t = Quarterly earnings per share

Fm_t = Change in quarterly earnings per share

Gm_t = Percent change in quarterly earnings per share

Rm_t = Quarterly dividend per share

Sm_t = Percent change in quarterly dividend per share

Xm_t = Percent change in company volume

Ym_t = Percent change in market volume

Dm_t = Dummy (profitability measure)

μ = Error associated with variables V and U

λ = Error associated with variables A and B

β = Autoregressive error

A_o = The pure constant term

et = Error not associated with lags

FOOTNOTES

[1]N. R. Draper and H. Smith, Applied Regression Analysis, John Wiley & Sons: New York, 1966, p. 4-6.

[2]James E. Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station: Oklahoma State University, May, 1968, p. 2.

[3]William Addison and Marc Nerlove, "Statistical Estimation of Long-Run Elasticities of Supply and Demand," Journal of Farm Economy, Vol. 50 (November, 1958), p. 869.

[4]Martin, p. 2.

[5]L. M. Koyck, "Distributed Lags and Investment Analysis," Contributions to Economic Analysis, North Holland Publishing: Amsterdam, Holland, 1954.

[6]Nerlove, p. 375.

[7]Martin, p. 16.

CHAPTER IV

RESULTS OF THE STUDY

In this paper an attempt has been made to cover fully all the informational and statistical bases upon which this study finds support. The study is based on the hypothesis that movements in securities prices are systematic, because of this a model can be constructed which will explain these movements. A qualification arises in that this study is of the opinion that the only way to build such a model is to include variables relating to both the value of the security and the demand for that security. Thus both fundamental and technical analysis factors were included in this study's model. The inclusion of both of these theoretical areas in the study gives the model a wide conceptual base in investment theory. If this base is supported by a significantly powerful statistical tool, then a meaningful explanatory model should be the result. The statistical results from both the BMD Linear Regression program and the Non-Linear Study Model will be covered in this chapter.

Results From the BMD Linear Regression Program

Though this program has been used mainly in a supporting role in this study, its proposed model will be established. This will give us a comparison between this simple linear regression model and the distributive lag model. The principles behind the linear model and its output have been explained before and need no reiterating in this chapter. It is of importance that it be remembered that any linear regression problem is making the three assumptions listed earlier in this paper. When a linear model is used, it assumes independence of all the variables, with the exception of the dependent variable, from one another. This will undoubtedly cause some error, especially when the time series variables are considered by the model.

In undertaking the job of analyzing the output of the model, the first thing undertaken is a check of the significance of the variables. Using a Standard T table, from a statistics textbook, you compare the computed T value with the table and establish the significance level for the variable (Table IV.) It can easily be seen that a high level of significance exists for most of the variables. Thirteen of the variables are significant at the .001 level while four of them are significant at the .01 level; these variables are highly correlated with the future price of the security. The other three variables (company volume at t-2, industry index at t-1, and percent change in company volume) should be replaced by variables which are more

TABLE IV

THE IMPORTANT STATISTICS FROM THE LINEAR MODEL

Variable	Regression Coefficient	Computed T Value	Proportion of Cumult. Coeff. of Det.
V_{t-2}	0.01281	1.76981****	0.17562
V_{t-1}	0.04726	3.37315**	0.03348
U_{t-2}	-0.00016	-5.20233*	0.00612
U_{t-1}	0.00005	10.12024*	0.00189
A_{t-2}	-0.01264	-17.02196*	0.02450
A_{t-1}	0.01105	1.73693****	0.00082
B_{t-2}	-0.02169	-2.20114***	0.00002
B_{t-1}	-0.00190	-7.72891*	0.00223
E_t	0.34665	6.39314*	0.20761
F_t	0.00226	-3.29493**	0.00298
G_t	0.00467	5.29822*	0.01203
R_t	-0.34209	-6.96533*	0.00081
S_t	-0.50850	4.68442*	0.02407
X_t	0.03859	-0.73133	0.00010
Y_t	0.78241	3.33435**	0.00892
D_t	-0.07549	-9.30785*	0.00008
P_{t-4}	0.00017	-3.21906**	0.46585
P_{t-3}	0.15532	16.57763*	0.00954
P_{t-2}	-0.11803	-14.67481*	0.00627
P_{t-1}	0.98143	17.04218*	0.01276

Coefficient of Determination	0.9859
Multiple Correlation Coefficient	0.9929
Variance of the Estimate	6.06445
Standard Error of Estimate	2.46261
Intercept (A Value)	1.93159
Competence Check of Final Coefficient	0.98141
Range of the Residuals	17.469
Range/Std. Error of Estimate	7.094

* Significant at .001 level
** Significant at .01 level
*** Significant at .02 level
**** Significant at .05 level

significant and explanatory. The variables which were significant to the level .001 give an indication that some autocorrelation might exist.

In order to check for this autocorrelation, we must go to the correlation matrix and see if any of these variables are highly correlated to the others. Upon completion, we find that the five price variables are all highly correlated, and thus their explanatory ability is questionable. There is nothing which can be done with the normal linear model to compensate for this autocorrelation except to eliminate those variables from the test group. This would affect the coefficient of determination in an adverse way.

The coefficient of determination for the original group was .9859 which was unusually high for a linear model, but it showed the possibility of a highly explanatory model. With the price variables removed from the test the coefficient is reduced by .49442, more than half, to .49148. This is caused by the inability of the linear model to cope with the autocorrelation of the price variables, and having to drop the highly autocorrelated variables because they were correlated with one another above the .75 level. The inability of the linear model to take these interrelationships into account forces one to assume away this autocorrelation or forces the rejection of the model as being non-explanatory. This rejection has been the case in investment theory, and many models which may have predictive value have been cast aside, or used in violation of the recommended cut off point (.75) for correlation between two independent variables.

Keeping these inconsistencies in mind we can construct a model using the linear regression program's recommended correlation coefficients. It will be assumed, in order to facilitate the construction of this model, that the interrelationship between all variables is below the .75 cutoff point. Thus the model can be constructed as if no problems existed, and as if it were our true estimation of the proposed explanatory model. The proposed linear model is shown in Equation 3, which is constructed on the pattern suggested by Equation 1(a). We can use this model as a point of reference when we finish the non-linear model in the next section.

$$(3)\quad P_t = 1.93159 + (0.01281)\ CV_{t-2} + (0.04726)\ CV_{t-1} - (0.00016)\ MV_{t-2} + (0.00005)\ MV_{t-1} - (0.01264)\ II_{t-2} + (0.01105)\ II_{t-1} - (0.02169)\ MI_{t-3} - (0.00190)\ MI_{t-1} + (0.54665)\ EPS + (0.00226)\ \Delta EPS + (0.00467)\ \%\Delta EPS - (0.34209)\ Div. - (0.50850)\ \%\Delta Div. + (0.03859)\ \%\Delta CV + (0.78241)\ \%\Delta MV - (0.07589)\ Dummy + 0.00017\ P_{t-4} + (0.15532)\ P_{t-3} + (0.11803)\ P_{t-2} + (0.98143)\ P_{t-1} + 7.094$$

This study does not suggest that the above model will prove to be highly explanatory, but instead this model was set forth here to show the point in research where most of the investment studies tend to stop. This study has shown where and why this model would not be a good explanatory model, and when a financial or economic study contains a times series autocorrelation between two or more of its variables should be expected. Care should be taken to make sure the analysis included searching the correlation matrix for variables which are correlated above the .75 cutoff point. When some variables exceed this level of correlation, some other technique should be used for the study; and in the case of a time series a Distributive Lag Model is recommended for accuracy.

Results Using The Non-Linear Study Model

Appendix C gives a good explanation of the Martin two-lag model. This model gives one a tool for taking up where the linear model leaves off, and to accomplish things which cannot be realized by the use of the normal linear regression. It does have its faults and limitations, but within its area of operating ability it provides a powerful and accurate tool. The important statistics in the output of the two-lag model for this study will be the pure constant term, the theta estimate of the parameters, the "T" value test of significance, the predicted, observed, and residual values of each observation, and the Durbin-Watson "d" statistic.[1] These will give us the basic elements we need for the development of the model.

In using a model such as this the interpretation of the different statistical output must be consistent with the intentions for which they were meant. This study was interpreted in accordance with the articles written by James E. Martin.[2] The important statistics are shown in Table V, except for the observed, predicted, and residual values of each observation which will be shown later in our analysis. The Durbin-Watson "d" statistic in this case shows us very little as it is in the range of inconclusiveness, and thus does not truly indicate whether positive or negative autocorrelation has been present in the proposed model. The level of this statistic tends toward the upper range of inconclusiveness, so one might intuitively suggest slight positive autocorrelation in this study. This would not be a statistical interpretation, but would be a response one might make to a statistic which is autocorrelated more positive than negative. In this study it will not be interpreted as either positive or negative without further information.

There is quite a lot of discrepancy between the T-values computed by the non-linear program and those of the

TABLE V

THE IMPORTANT STATISTICS FOR THE NON-LINEAR MODEL

Variable	Theta	T-Value
V_t	0.03067	- 0.84222
U_t	-0.08691	0.61435
A_t	0.00057	1.72959****
B_t	0.00061	2.16545***
λ	0.70421	31.59242*
μ	0.77121	30.84149*
β	0.67817	33.65592*
E_t	-0.06453	5.68488*
F_t	-0.00084	0.11593
G_t	0.02011	0.74031
R_t	0.06258	1.23887****
S_t	18.32309	6.13492*
X_t	16.14764	3.45573**
Y_t	-13.51444	3.69817*
D_t	-4.21401	2.03873***
R Square (Coefficient of Determination)		.92131
Pure Constant Term		-0.61241436
Durbin-Watson "d" Statistic		1.73608780

* Significant at the .001 level

** Significant at the .01 level

*** Significant at the .02 level

**** Significant at the .05 level

linear program. It is readily noticeable that the non-linear model produces more insignificant variables and a lot fewer variables which were significant at the .001 level. If all those variables which were significant below the .01 level were removed from the study based on their proportion of the cumulative coefficient of determination (Table IV) the "R" Square would be reduced by .25625 to .66506. This shows that the model, even though it is more restrictive, will yield a higher coefficient of determination than the linear model. Even if those variables yielding a significance level of .001 were checked for correlation above the cutoff point, the results will be the same. The three correlation adjustment parameters have absorbed this autocorrelation of the time series, and thus have reduced the chance of high correlation between variables. Thus the significance level is established for this model and is significantly higher than the .49148 of the linear regression technique. This study is now ready to construct the non-linear model.

The Proposed Model

The final model, based on the distributive lag programs of James E. Martin, has undergone a complex development process.[3] It has been suggested that some of the variables are insignificant and should be replaced, and the inconclusiveness of the Durbin-Watson statistic for one job has been explained. Five jobs were run using different sets of the four lagged variables associated with the two lag parameters, but in all of these jobs the "d" statistic remained in the inconclusive range. Thus the study suggests the possibility of a lack of any high degree of autocorrelation, or the lack of statistical power on the part of the "d" statistic in testing for positive or negative autocorrelation.

The model suggested by the computer output is shown in Equation Four, which is set up using the format shown in Table III. This is a complex model whose handling would be

(4) $$P_t = -0.61241\ (1 - 0.70421)\ (1 - 0.77121)\ (1 - 0.67817) + 0.03067\ V_{t-1} - (0.77121 + 0.67817)\ 0.03067\ V_{t-2} + (0.77121)\ (0.67817)\ 0.03067\ V_{t-3} - 0.08691\ U_{t-1} + (0.77121 + 0.67817)\ 0.08691\ U_{t-2} - (0.77121)\ (0.67817)\ (0.08691)\ U_{t-3} + 0.00057\ A_{t-1} - (0.70921 + 0.67817)\ 0.00057\ A_{t-2} + (0.70421)\ (0.67817)\ 0.00057\ A_{t-3} + 0.00061\ B_{t-1} - (0.70921 + 0.67817)\ 0.00061\ B_{t-2} + (0.70421)\ (0.67817)\ 0.00061\ B_{t-3} + 0.06453\ EPS - 0.00084 \Delta EPS + 0.02011\ \%\Delta EPS + 0.06258\ Div. + 18.32309\ \%\Delta Div. + 16.14764\ \%\Delta CV - 13.51444\ \%\Delta MV - 4.21401\ Dummy + (0.70421 + 0.77121 + 0.67817)\ P_{t-1} - [(0.70421 + 0.77121)\ 0.67817 + (0.70421)\ (0.77121)]\ P_{t-2} + (0.70421)(0.77121)(0.67817)\ P_{t-3} - 0.61241436$$

extremely difficult. It must be realized that this is not a tool for the individual investor, because of the need for a computer to do the computation of the predicted value. We will now make predictions using both the linear and non-linear techniques and use these regressions to compare the two.

In looking at the printed observed, predicted, and residual values for the non-linear model, it is noted that it tends to be above the market when it is in a decline and below the market in an increase. In movements from one stage to the other the predicted value tends to move up one period in advance of the stock price but tends to move down at the same time. The predicted value for "t" will tend to show the price at "t" to begin a decrease so prior knowledge would be present. This cannot be taken to mean that this model will accomplish this sort of a lag and prediction in all cases. In examining this study one must remember the qualifications that without further study and testing, this model can only be applied to the companies tested for the time period in which they were tested. Thus, it would be a statistical testing error to suggest that from no more evidence than is present in this study that this model's predicted values will always lead the observed value by any specific period.

TABLE VI

PREDICTED, OBSERVED, AND RESIDUAL VALUES FOR SECURITY Z

Observation	Actual Price	Non-Linear Predicted	Non-Linear Residual	Linear Predicted	Linear Residual
1	66.5	68.693	-2.219	65.329	.171
2	66.5	68.80	-2.30	59.135	6.365
3	65.25	68.791	-3.54	62.0	3.25
4	64.25	67.861	-3.61	64.25	0.0
5	63.625	67.261	-3.63	66.003	-2.378
6	61.875	67.024	-5.149	70.581	-8.706
7	58.75	64.725	-5.975	71.521	-12.771
8	61.25	62.80	-1.50	74.039	-12.789
9	64.5	65.417	- .917	73.821	-9.321
10	68.75	67.244	1.50	75.125	-6.375
11	66.5	70.561	-4.061	76.751	-10.251
12	71.0	68.92	2.07	76.511	-5.511
13	74.25	72.139	2.11	78.115	-3.865
14	83.25	74.056	9.19	78.415	4.835
15	82.5	81.338	1.16	79.125	3.375
16	79.0	80.23	-1.23	80.051	-1.051

Shown in Figure 2.

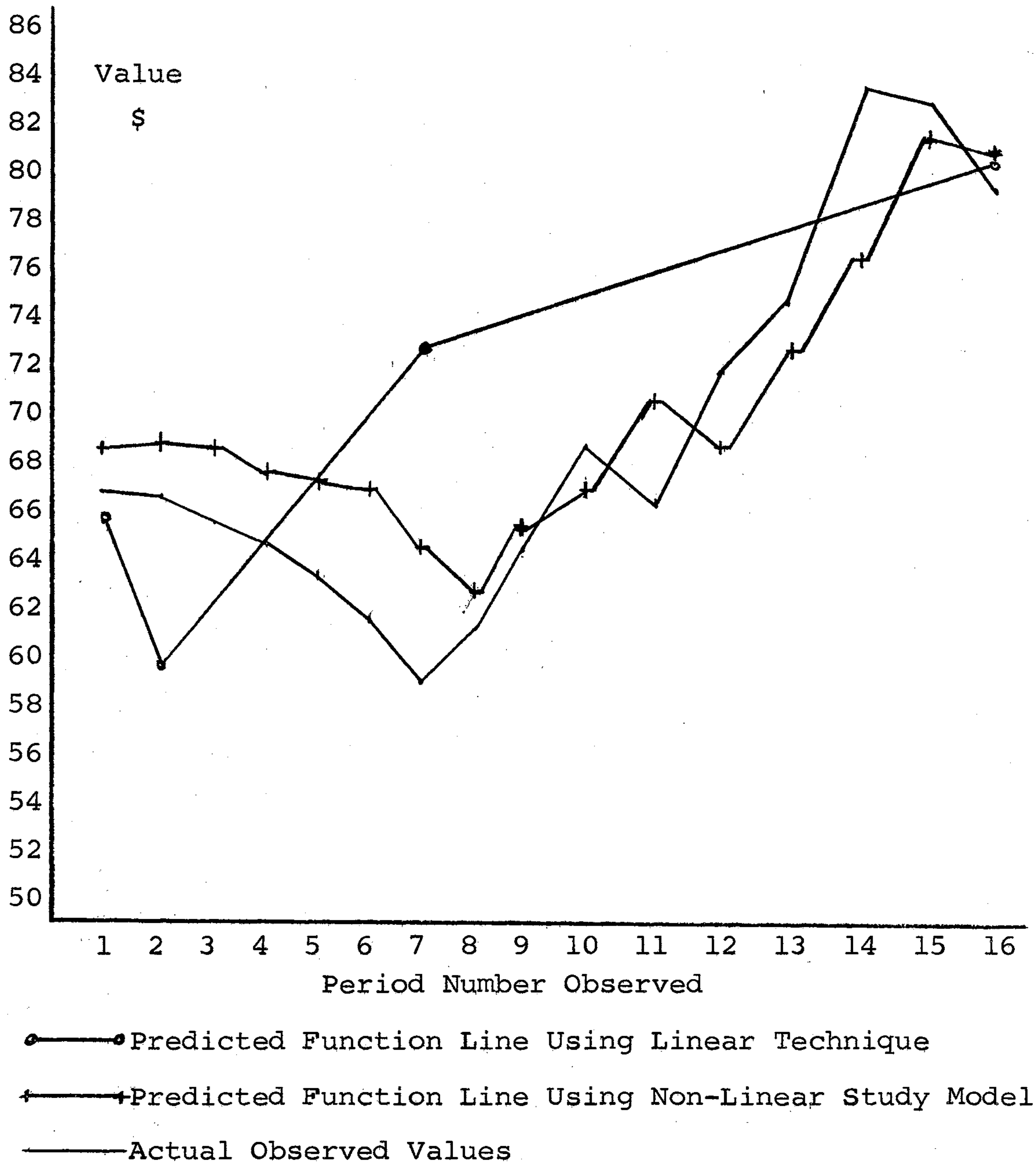


Figure 2. Relationship Between the Actual Values, the Predicted Values of the Linear Technique, and the Predicted Values of the Non-Linear Technique

FOOTNOTES

[1]James E. Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968.

[2]James E. Martin, "The Use of Distributed Lag Models Containing Two Lag Parameters in the Estimation of Elasticities of Demand," Journal of Farm Economy, Vol. XLV (December, 1963).

[3]Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models.

CHAPTER V

CONCLUSIONS AND PROPOSALS

A complex mathematical model has been applied in this paper to the explanation of stock price movements. The model has been shown to have some predictive as well as explanatory capability, but further testing needs to be done on similar models to prove its value. The value of such a study may not be the model itself, but the bringing together of previously divided areas. This paper has suggested that its application to investment theory is only an extension of its widespread economic applicability and capability.

This model suggests the combining of two previously separate concepts of investment theory. The fundamentalist has always used the information available to predict an intrinsic value, and would base his buying and selling decisions on this value without taking into consideration the demand for the security. The technical analyst believes that history tends to repeat itself, and thus an individual can use past data and information to project future price movements for a security. In this study an attempt has been made to include factors from both of these fields of analysis. The results have been gratifying, since the explanatory ability of the model, even without the insignificant variables, is beyond expectations. It is suggested that more widespread use of combination models be undertaken in research of securities price movements.

The statistical technique employed in this study has shown its possible applications in research concerning securities analysis. The ability to properly take into account the autocorrelation of the time series studies will allow for the development of even better and more sophisticated models than the one proposed in this study; and if the random walk hypothesis is to be shown to be invalid, models such as this non-linear distributive lag technique must be developed. This area of study has not yet exhausted the possibilities of research within its narrow scope, but has merely opened the door to new and more complete analysis of the variables, companies, and the time period covered by the study.

This study should be expanded on any one or all of the three following courses: new and more variables; more companies; and longer time period. Some of the variables in the present study should be eliminated because of their insignificance or lack of predictability. New variables need to be added to replace these and others should be added to increase the model's predictability, but this can not be undertaken in wholesale fashion. Each of these new variables must be chosen for their possible predictive capacity, and must be tested for significance and a level of predictability above that of the original variable. The choosing of a better variable set will have a great effect on the explanatory capability of the model given that the other two constraints are fixed.

The study could also have its predictability increased by adding more companies, and a second group of companies could be used to form a test group to prove the model's validity. Research should be undertaken to establish the predictability of non-linear models for income stocks, growth company stocks, and growth stocks. This could be accomplished by selecting companies to complete both a study group and test group from each of these areas and then develop and test a model oriented toward each of these kinds of securities. These models could then be compared to a model consisting of randomly selected companies to find if any significant differences were present. An alternative to this would be the expansion of the present random selection to between 50 and 100 companies. There is little doubt that any one or all of these expansions of the present study would significantly increase the application and predictability of the presently proposed model.

Another way of increasing explanation of the proposed model is to increase the time period studied. There are two possible alternatives here: (1) extend the present period back X years or (2) select X number of years at random from the years security markets have been in existence. The first of these alternatives is the more simple, since the entire study period can be placed in the model at once. The second will give a better explanation of the overall life time of the securities markets, but each year must be tested separately and thus would not lead to the development of a single equation model. For this reason, the first alternative is recommended for possible expansion of the time period in a future study. The explanatory capability of the technique should be tested by choosing a second time interval of the same length and testing the developed model for predictability during this period.

In order to expand this study on all three possible courses, the individual would have to have almost limitless time and resources. It would take between twenty and thirty hours of computer time, along with many hours of keypunching, coding, and gathering of the data. For this reason, the expansion of this study should be taken in segments or by a group working in coalition with one another. This is not meant to discourage anyone from pursuing the expansion of any part of the study, because this can be accomplished without the long periods of dedicated study and work. This study, in fact, was aimed at challenging others to expand the present model and its principles.

A SELECTED BIBLIOGRAPHY

Addison, William, and Marc Nerlove. "Statistical Estimation of Long-Run Elasticities of Supply and Demand," _Journal of Farm Economy_, Vol. L (November, 1958), pp. 868-881.

Alt, F. L. "Distributed Lags," _Econometrica_, Vol. X, No. 2 (1942).

Baumol, William J. _The Stock Market and Economic Efficiency_. New York: Fordham University Press, 1965.

Bernstien, Peter L. "Growth Companies vs. Growth Stocks," _Harvard Business Review_, Vol. XXXIV, No. 5 (September-October, 1956), pp. 87-98.

Cagan, P. "The Monetary Dynamics of Hyper-Inflations," _Studies in the Quantity Theory of Money_, Illinois: University of Chicago Press, 1956.

Cohen, Jerome B., and Edward D. Zinbarg. _Investment Analysis and Portfolio Management_, Illinois: R. D. Irwin, 1967.

Cootner, Paul H. "Stock Prices: Random vs. Systematic Changes," _Industrial Management Review_, Vol. III, No. 2 (Spring, 1962), pp. 24-45.

Cootner, Paul H. _The Random Character of Stock Prices_. Massachusetts: MIT Press, 1967.

Durand, David. "Growth Stocks and The Petersburg Paradox," _The Journal of Finance_, Vol. XII, No. 3 (September, 1957), pp. 348-363.

Durand, David and Alan M. May. "The Ex-Dividend Behavior of American Telephone and Telegraph Stock," _The Journal of Finance_, Vol. XV, No. 1 (March, 1960), pp. 19-31.

Fama, Eugene F. "Random Walks in Stock Market Prices," _Financial Analysts Journal_, Vol. XXI (September-October, 1965), pp. 55-59.

Fuller, Wayne A., and James E. Martin. "A Note on the Effects of Autocorrelated Errors on the Statistical Estimation of Distributed Lag Models," Journal of Farm Economy, Vol. ILIV, (May, 1962), pp. 407-410.

Fuller, Wayne A., and James E. Martin. "The Effects of Autocorrelated Errors on the Statistical Estimation of Distributed Lag Models, Journal of Farm Economy, Vol. ILIII (February, 1961), pp. 71-82.

Griliches, Zvi. "A Note on Serial Correlation Bias in Estimates of Distributed Lags," Econometrica, Vol. XXIX, (1961), pp. 65-73.

Griliches, Zvi. "Distributed Lags: A Survey," Econometrica, Vol. XXXV, No. 1 (January, 1967), pp. 16-47.

Griliches, Zvi and N. Wallace. "The Determinents of Investment Revisited," International Economics Review, Vol. VI, No. 3 (1965).

Hartley, H. O. "The Modified Gauss-Newton Method for Fitting of Non-Linear Regression Functions by Least Squares, Technometrics, Vol. III (May, 1961), pp. 269-280.

Holt, Charles C. "The Influence of Growth Duration on Share Prices," The Journal of Finance, Vol. XVII, No. 3, (September, 1962), pp. 465-475.

Koyck, L. M. "Distributed Lags and Investment Analysis," Contributions to Economic Analysis. Amsterdam, Holland: North Holland Publishing, 1954.

Krainer, Robert E. "Liquidity and Stock Market Speculation," Journal of Financial and Quantitative Analysis,

Levy, Robert A. "Conceptual Foundations of Technical Analysis," Financial Analysts Journal, Vol. XXII (July-August, 1966), pp. 83-89.

McWilliams, James D. "Prices, Earnings, and P-E Ratios," Financial Analysts Journal, Vol. XXII (May-June, 1966), pp. 137-142.

Martin, James E. "Isolation of Lagged Economic Responses," Journal of Farm Economy, Vol. ILIX (February, 1967), pp. 160-168.

Martin, James E. "The Use of Distributed Lag Models Containing Two Lag Parameters in the Estimation of Elasticities of Demand," Journal of Farm Economy, Vol. ILV (December, 1963), pp. 1474-1481.

Martin, James E. Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, Oklahoma, 1968.

Mayor, Thomas H. "Short Trading Activities and the Price of Equities: Some Simulation and Regression Results," Journal of Financial and Quantitative Analysis, Vol. III (September, 1968), pp. 283-298.

Seelinfruend, Alan, George C. Parker, and James C. VanHorne. "Stock Price Behavior and Trading," Journal of Financial and Quantitative Analysis, Vol. III (September, 1968), pp. 263-281.

Smidt, Seymour. "A New Look at the Random Walk Hypothesis," Journal of Financial and Quantitative Analysis, Vol. III (September, 1968), pp. 235-261.

Soldofsky, Robert M. "Growth Yields," Financial Analysts Journal, Vol. XVII, No. 5 (September-October, 1961), pp. 43-47.

Walter, James E. "Dividend Policy: Its Influences on the Value of the Enterprise," The Journal of Finance, Vol. XVII, No. 2 (May, 1963), pp. 280-291.

Zakon, Alan J., and James C. Pennypacker. "An Analysis of the Advance-Decline Line as A Stock Market Indicator," Journal of Financial and Quantitative Analysis, Vol. III (September, 1968), pp. 299-314.

APPENDIX A

STUDY GROUP OF COMPANIES

	Company	Industry
1.	American Crystal	Beet Sugar
2.	American Enka	Synthetic Fibers
3.	American Metal Climax	Metals-Misc.
4.	Babcock Wilcox	Machinery-Steam
5.	Copper Range	Copper
6.	Cummings Engine	Auto-Trucks
7.	Dayton Power and Light	Electric Utilities
8.	Dome Mines	Gold Mining
9.	Fibreboard	Containers-Paper
10.	General Acceptance Finance	Finance
11.	General Signal	Railroad Equipment
12.	Halliburton	Machinery and Service-Oil Well
13.	International Harvester	Machinery-Agricultural
14.	Kroger	Food Chain
15.	Lockheed	Aerospace
16.	May Department Stores	Department Stores
17.	National Biscuit	Biscuit Bakers
18.	North American Coal	Coal
19.	Owens-Corning	Miscellaneous
20.	Pan American World Airlines	Air Transport
21.	Pittsburg Plate Glass	Miscellaneous
22.	Publicker Industries	Distillers
23.	Reynolds Tobacco	Tobacco-Cigarettes
24.	Royal Crown Cola	Soft Drinks

25.	Servomation	Vending Machines
26.	Southern Natural Gas	Natural Gas-Pipe Line
27.	Texas Gulf Sulphur	Sulphur
28.	United States Gypsum	Wallboard
29.	Whirlpool	Appliances
30.	Youngstown Sheet and Tubing	Steel

APPENDIX B

VARIABLES AND THEIR SOURCE REFERENCES

Technical Variables

Variable	Primary Source	Secondary Source
Weekly Price Data	Wall Street Journal	Barron's Commercial and Financial Chronicle
Weekly Volume Data	Barron's	Commercial and Financial Chronicle
Weekly Indices*	Standard and Poor's Trade & Statistics	Standard and Poor's Outlook
Weekly Indices**	Standard and Poor's Outlook	-----

Fundamental Variables

Variable	Primary Source	Secondary Source
Earnings Data	Wall Street Journal	Barron's Commercial and Financial Chronicle
Dividend Data	Barron's	Wall Street Journal Commercial and Financial Chronicle
Profitability	Wall Street Journal and Moody's Handbook	

* Prior to January 1, 1969

** After January 1, 1969

Wall Street Journal is published daily (except Saturday and Sunday) by Dow Jones and Company.

Barron's (National Business and Financial Weekly) is published weekly by Dow Jones and Company.

Commercial and Financial Chronicle is published weekly by William B. Dana Company.

Standard and Poor's Outlook is published weekly by Standard and Poor's Corporation.

Standard and Poor's Trade and Statistics is published weekly by Standard and Poor's Corporation.

Standard and Poor's Trade and Statistics is published yearly with quarterly updates by Standard and Poor's Corporation.

Moody's Handbook is published quarterly by Moody's Investor's Service.

APPENDIX C

THE COMPUTERIZED DISTRIBUTIVE LAG MODEL

Introduction

The use of more sophisticated models requires more study and preparation to insure its complete applicability to the problem being studied. Even when this requirement is met the problem has not yet been solved. The study must orient all of its data gathering, structuring and testing towards the capabilities and needs of the model. Highly quantitative research models are designed for specific research functions, and require a certain form in their operational inputs and outputs. The application of a quantitative technique such as this study's model to a research project may force not only revisions in the data, but also might call for slight revisions in the technique itself. Care must be taken to insure that these adjustments do not interfere with the operation of the program. Only when all of this preparation is properly accomplished will the output be accurate and usable.

This Appendix is aimed at acquainting the individual with the sophisticated distributive lag model. An attempt will be made to bring this statistically oriented model down to the level of understanding of the normal business student, and make it possible for them to use the model without the hours of background study which was undertaken in this study. The detail with which this will be covered will be great, but the statistical detail will be less foreboding. This Appendix will cover the formats for the different cards needed, the limitations of the program, the designs of the different programs and job structures, the special operations available, and some of the problems and errors common with the use of this program.

The Program Structure

The capability of this one program ranges from a simple static multiple regression with independent error terms to a non-linear regression allowing for the lagging of two or more variables with three different autoregressive error terms (λ, μ, β). This is a technique which James E. Martin has developed into a widely applied tool of research in Agricultural Economics. The non-linear technique has universal application to all economic phenomenom involving time series studies. Its highest developed applications have been in the field of Econometrics and Agricultural Economics, but its application has been suggested, in the field of investment theory, by L. M. Koyck

(Distributed Lags and Investment Analysis)[1] and Zvi Griliches and N. Wallace (The Determinants of Investment Revisited).[2] These two applications of distributive lags to investment analysis are in the realm of company investment in plant and equipment, and their application to the securities investment is not supported. This study feels that this is overlooking a prominent and profitable field for these applications, thus the application of this technique to securities' analysis was undertaken.

The aim of the distributive lag models of James E. Martin is to set forth a computerized tool with widespread applications which can be applied to data with a minimum of adjustments in both data and programs.[3] There are two program decks associated with this model (Table VII) but each allows only a certain number of operations to be performed. Their basic operation is the same, but MRHS-1 is an "in-core" program which is highly efficient in the use of computer time; while MRHS-2 is for use on larger problems, but it requires an additional tape drive and is not efficient in the use of computer time. This allows for the use of a small efficient model on small jobs and a slower model for larger, more complex jobs. The problem cards, control cards, and start vector cards are set up the very same way for both programs and are interchangeable as long as the dimensions and limitations of the specific program are not exceeded.

TABLE VII

MAXIMUM DIMENSIONS AND LIMITATIONS
OF MRHS-1 AND MRHS-2

MRHS-1

(a) Data Deck:

(1) The total number of variables which may be included in the data deck must not exceed 20.

(2) The total number of observations per variable must not exceed 150.

(b) Job: The parameters of control card 1 which control a specific job must meet the following restrictions:

(1) $3(A+B) + C + 4 < 47$

(2) $A + B + C + 4 < 35$

Where:

(3) A = the total number of variables associated with lag parameter λ

(4) B = the total number of variables associated with the lag parameter μ

(5) C = the total number of variables which are not associated with lag parameters

MRHS-2

(a) Data Deck:

(1) The total number of variables which may be included in the data deck must not exceed 80.

(2) The total number of observations per variable is limited to 99,999. For large problems it is recommended that all variables be coded within the range, $\pm$ 10, if possible.

TABLE VII (Continued)

(b) Job: The parameters of control card 1 must meet the following restrictions:

(1) $3(A + B) + C + 4 < 83$

(2) $A + B + C + 4 < 75$

Where A, B, and C are defined by equations 3, 4, and 5 of MRHS-1 dimension restrictions.

(c) The user must not call for the correlation matrix if $A + B + C + 3 > 75$.

SOURCE: James E. Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968.

These programs are aimed at the estimation of seven distinct types of regressive equation techniques using the following input formats for the various control cards. The problem card controls the parameters which specify the number of variables, observations per variable, and the number of jobs which are associated with the data deck. The format is shown by Table VIII.

All the values punched in columns 1-20 of the problem card must be right justified integers; while the values punched in columns 21-80 may contain alphanumeric information. The data deck must be punched so that it is in row order, the first row contains all the first observations for each variable (X_{11}, X_{21}, ..., X_{A1}, Z_{11}, Z_{21}, ..., Z_{B1}, D_{11}, D_{21}, ..., D_{C1}, Y_{11}, Y_{21}, ..., Y_{y1}). The variables

TABLE VIII

THE FORMAT FOR SETTING UP OF THE PROBLEM CARD

Cols.	1- 5 . .	the problem number
Cols.	6-10 . .	the number of variables in the data deck
Cols.	11-15 . .	the number of observations per variable in the data deck
Cols.	16-20 . .	the number of jobs to be run using this data deck
Cols.	21-80 . .	blank or any identifying information the user wishes to use

SOURCE: James E. Martin, Computer Algorithms Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station: Oklahoma State University, May, 1968.

in the data deck must also conform as column vectors given the row order above as is shown in Table IX. The user must realize that only the most current observation of each variable is punched on any data card, thus a data card or group of data cards contain only one observation for each variable. The problem card and data cards give the basic structure of the problem and the information base, but the control over the functions to be performed, the variable to be used as the dependent variable, the tests to be run, etc. rests in the control cards.

There are two control cards associated with this program. The first control card defines the major parameters which control the types of estimations, the functions to be performed, and the tests parameters for the study,

TABLE IX

SPECIFICATIONS FOR DATA CARDS

$$X_{11}, X_{21}, \ldots, X_{A1}, Z_{11}, Z_{21}, \ldots, Z_{B1}, D_{11}, D_{21}, \ldots, D_{C1}, Y_{11}, Y_{21}, \ldots, Y_{Y1}$$

$$X_{12}, X_{22}, \ldots, X_{A2}, Z_{12}, Z_{22}, \ldots, Z_{B2}, D_{12}, D_{22}, \ldots, D_{C2}, Y_{12}, Y_{22}, \ldots, Y_{Y2}$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot$$

$$X_{1n}, X_{2n}, \ldots, X_{An}, Z_{1n}, Z_{2n}, \ldots, Z_{Bn}, D_{1n}, D_{2n}, \ldots, D_{Cn}, Y_{1n}, Y_{2n}, \ldots, Y_{Yn}$$

the number of variables to be associated with each autocorrelated parameter, and the variables not to be considered by that job. This card is set up on the format shown in Table X. Care must be taken to insure the proper format for this card, as it controls the overall operation of the program. If this format is not followed, the computer will terminate the operation and print out an error message associated with a control card error. The two control cards are directly related in that the second control card is used to set forth specific variable addresses related to columns 14 and 15 of control card one, in that it instructs the computer what parameters to skip. The programs are written so that the given variables are in row parameters, in the answer space, in relation to column location on the data cards. The exception to this is that the autocorre-

TABLE X

FORMAT AND SPECIFICATIONS FOR CONTROL CARD NUMBER ONE

Cols. 1- 5 . .	job number
Cols. 6- 7 . .	the total number of exogenous variables associated with lag parameter λ ($A \geq 0$)
Cols. 8- 9 . .	the total number of exogenous variables associated with lag parameter μ ($B \geq 0$)
Cols. 10-11 . .	the total number of exogenous variables and/or dummy variables which are not associated with a lag parameter ($C \geq 0$)
Cols. 12-13 . .	the column number of the variables in the data deck which is to be used as the dependent variable for this job
Cols. 14-15 . .	the total number of parameters or rows to be controlled or skipped, NSKIP(NSKIP $\geq$ 0)
Cols. 16-17 . .	"01" for least squares estimation, otherwise the maximum number of iterations to be performed in a non-linear estimation job ("25" is suggested for non-linear jobs).
Col. 18	1 if the calculation of the pure constant term is desired, otherwise, "0" when regressions through the origin are desired
Col. 19	1 if the calculation of the predicted values, residuals and Durbin-Watson "d" statistic is desired, otherwise, "0"
Col. 20	1 if a listing of the correlation matrix is desired, otherwise "0". Note: This option is available only when Col. 31 of control card 1 contains "0"
Cols. 21-30 . .	the test criterion to be used to halt the iterative process in non-linear jobs. The value of the test criterion is punched using FORTRAN Format (F10.4)... ("0.0001" is recommended for most non-linear jobs.)
Col. 31	"0" if the job to be run is: (1) the first job in a problem, (2) involves a different dependent variable from the preceeding job, or (3) involves a different computation on the pure constant term from the preceeding job; otherwise, "1".
Cols. 32-80 . .	Blank or any identifying information the user wishes to punch into this card.

SOURCE: James E. Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968.

lation parameters (λ, μ, and β) are injected in the three columns following the variables associated with the first two lag parameters. It is very important to identify the exact location of those variables and parameters to be skipped in the operation of the program. The format for control card 2 is [15(I2,2x)]. This is shown by Table XI.

TABLE XI

THE PARAMETER FORMAT FOR THE SKIP CONTROL CARD NUMBER TWO

Cols.	1- 2 . .	the actual row number of the first parameter to be controlled or skipped
Cols.	3- 4 . .	Blank
Cols.	5- 6 . .	the actual row number of the second parameter to be controlled or skipped
Cols.	7- 8 . .	Blank
.		
.		
.		
etc.		
.		
.		
.		
Cols.	61-80 . .	Blank

Note: The number of parameters which are punched into control card 2 and which are to be controlled or skipped must agree with the number punched into columns 14-15 of control card 1.

SOURCE: James E. Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968.

These control cards set the pattern and structure for the programs. But the model also requires the providing of estimates of the parameters which will result from the operation of the model. These estimates can be acquired by the use of the linear regression technique, but these linear estimates, though highly accurate, cannot be taken as the parameters in a non-linear problem. The final parameters of the non-linear program are not only the result of the linear relationship of the variable with the dependent variable, but also includes its relationships with the other independent variables.

The estimates which are needed for the programs are punched into the start vector cards using the format (8F10.5). In some cases the appropriate number of blank cards will suffice for the start vector cards. The number of start vector cards, NSVC, can be figured using the formula $NSVC = \frac{A+B+C+3}{8}$ where A is the number of variables associated with lag parameter λ, B is the number of variables associated with lag parameter μ, and C is the number of variables unassociated with any lag parameter. The nearer the user's estimates approximate the final parameter estimates, the shorter will be the computing time for the job. The actual punches in the start vector cards is determined by the type of a job estimation required in the operation of the program.

Types of Estimations

There are basically seven structures of equations for which this program will estimate the parameters. Each of these program structures call for some difference in the punches which occur on one or more of the job control cards. In this section the differences in these cards will be explained, along with an explanation of the uses of the particular models. Only those parameters and punches required to obtain each of the permissable combinations of mathematical models and types of estimations are presented here, and the format given before covers all punches other than those specified by each model.

The first model is a static model with independent errors. A static model is one whose equation contains only current exogenous and/or dummy variables as independent variables. Thus a least squares estimate, used in all linear regression techniques, can be used to estimate the parameters for this type of job. The program sets up three lagged periods of the dependent variable and assumes , , and to exist in every model; but in this model if they reduce to "0" it takes on the form of the normal static regression model, thus their incorporation into the model is merely to satisfy the program format. The general form of the model is shown by equation five. The identifying punches will instruct the computer which job is to be run using the data. A Type One estimation is non-iterative and is similar to the normal linear regression model.

(5) $Y_t = A_o (1-\lambda)(1-\mu)(1-\beta) + (\lambda+\mu+\beta) Y_{t-1} -$
$[(\lambda+\mu)\beta + \mu\lambda] Y_{t-2} + \lambda\mu\beta Y_{t-3} +$
$\sum_{k=1}^{c} dk\, Dkt + et$

A Type Two estimation is a little more complex, but it incorporates the assumption of independent error terms for each of the independent variables. All of the independent variables are assumed to belong to either Xi_t or Zj_t, whichever is chosen for the study. The general form of this equation and its reduced state are shown in equation six and seven. It must be realized that this model is only applicable when all the independent variables are closely related.

(6) $Y_t = A_o (1-\lambda)(1-\mu)(1-\beta) + \sum_{i=1}^{A} Ai\ Xi_t - (\mu + \beta)$
$\sum_{i=1}^{A} AiXi_{t-1} + \mu\beta \sum_{i=1}^{A} AiXi_{t-2} + (\lambda+\mu+\beta)Y_{t-1}$
$- [(\lambda+\mu)\beta + \lambda\mu]\ Y_{t-2} + \lambda\mu\beta\ Y_{t-3} + et$

If $\lambda \equiv \mu \equiv 0$ and $\beta = 1$, this equation reduces to:

(7) $Y_t - Y_{t-1} = \sum_{j=1}^{B} (Xj_t - Xj_{t-1}) + et$

The next model is the least squares estimate of the Nerlove Model with independent error, and this is the most sophisticated of the single lag models with independent error terms.[4] This model assumes independence of the errors terms since the equation using a single lagged independent error group can be reduced to an equation which can be solved by the use of a simple linear technique. Thus all lagged variables must be assumed to belong to the group Xi_t or Zj_t; but unlike the preceeding model, this model allows for the inclusion of a separate group of variables unassociated with any error. Its format is set forth in equation eight and nine. This model allows for more freedom in the selection of variables than do the first two models.

$$(8) \quad Y_t = Ao\,(1-\lambda)(1-\beta) + \sum_{i=1}^{A} AiXi_t - (\mu+\beta)\sum_{i=1}^{A} AiXi_{t-1} + \mu\beta\sum_{i=1}^{A} AiXi_{t-2} + (\lambda+\mu+\beta)\,Y_{t-1} - [(\lambda+\mu)\beta + \lambda\mu]\;Y_{t-2} + \lambda\mu\beta\,Y_{t-3} + \sum_{k=1}^{D} dkDk_t + et$$

reduces to:

$$(9) \quad Y_t = Ao\,(1-\lambda) + \sum_{i=1}^{A} AiXi_t + \lambda Y_{t-1} + \sum_{k=1}^{D} dkDk_t + et$$

In discussing the fourth type of estimation we come to our first model which corrects for the autoregressive errors. We now come into the field of correcting for the interrelationships between variables and time series. This type of estimate is computed upon a static model which has a one period lag associated with the independent variables and the dependent variable. All of the independent variables are assumed to belong to the same general group and are lagged accordingly. The autoregressive error (β) only takes into account that error existing between different time periods of the same variable, and not the interrelations between the variables themselves. This is represented in equation ten and is the first non-linear model set forth by Martin.[5]

$$(10) \quad Y_t = Ao\ (1-\beta) + \sum_{i=1}^{A} AiXi_t - \beta \sum_{i=1}^{A} AiXi_{t-1} + \beta Y_{t-1} + et$$

Type Five estimations are also related to a model set forth previously, Nerlove's Model, but this specific type is corrected for the autoregressive errors.[6] This technique allows for much freedom in the data collected, even though it restricts the lagging to one type of variable which may cause problems if the user wishes accuracy. This is caused by the assumption that all independent variables to be lagged are members of one group of variables. Though Martin has designed a two lag model, it is questionable how much better this is than the conventional one lag Nerlove Model. The equation for the Nerlove Model is set forth in equation five. This is a significantly powerful tool, and if all the lagged factors are rather highly correlated it would probably do a better job of estimating than a two lag model.

The study now has progressed to Martin's two lag model.[7] A two lag model is one in which two different independent variable groups can be lagged at the same time. This allows the individual to lag two completely different types of variables at the same time without hurting the accuracy of the study, but instead increasing the accuracy beyond the original. There are two types of estimates related to the two lag models and the first is with independent errors. The form of the equation is non-linear, but the estimates coming from this program are based on the linear assumptions. The operation is iterative, but because of the tendency for to approach "0" it takes on the appearance of a linear equation technique. The general form of its equation is shown by equation eleven.

$$(11)\quad Y_t = Ao(1-\lambda)(1-\beta) + \sum_{i=1}^{A} AiXi_t - \mu\sum_{i=1}^{A} AiXi_{t-1} + \sum_{j=1}^{B} bjZj_t - \lambda\sum_{j=1}^{B} bjZj_{t-1} + (\lambda+\mu)\, Y_{t-1} - \lambda\mu\, Y_{t-2} + \sum_{k=1}^{C} dkDk_t + et$$

The last type of estimate available using Martin's program is a two lag model with autoregressive errors.[8] This is the type of model for which the program was developed and designed, and it is the one which has proven to be the most error free of his models. The present study is designed on the basis of this model, and proposes its widespread use, where applicable, in the field of financial research. Using this model, the user can lag two separate categories of independent variables and have a separate category for exogenous variables not associated with either lag parameter, and lag the dependent variable three time periods. This model corrects for the errors associated with the time series, the interrelation of variables, and the autocorrelation of the equation as a whole. The general form of this equation is shown by equation twelve. Since this is the model to be used in this study, an explanation will be made of the adjustment process in order for the computer to know which job to run. Since the program was designed for this type of program, it takes less adjustment than the others, and this adjustment is shown in Table XII.

$$
\begin{aligned}
(12)\quad Y_t = {} & Ao\,(1-\lambda)(1-\mu)(1-\beta) + \sum_{i=1}^{A} AiXi_t - (\mu+\beta) \\
& \sum_{i=1}^{A} AiXi_{t-1} + \mu\beta \sum_{i=1}^{A} AiXi_{t-2} + \sum_{j=1}^{B} bjZj_t - \\
& (\lambda+\beta) \sum_{j=1}^{B} bjZj_{t-1} + \lambda\beta \sum_{j=1}^{B} bjZj_{t-2} + \\
& (\lambda+\mu+\beta) Y_{t-1} - [(\lambda+\mu)\beta + \lambda\mu] Y_{t-2} \\
& + \lambda\mu\beta Y_{t-3} + \sum_{k=1}^{C} dkDk_t + et
\end{aligned}
$$

TABLE XII

SPECIFIC PUNCHES FOR TWO LAG MODEL
WITH AUTOREGRESSIVE ERRORS

(a) Control Card 1: Punch 25 in column 16-17 of control card 1.

(b) Control Card 2: Using a blank card for control card 2.

(c) Start Vector Cards: Placing estimates of all parameter values which are expected as final parameter estimates in start vector cards.

SOURCE: James E. Martin, *Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models*, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968.

These types of model estimates, which can be developed from this one program, give one a wide range of capabilities. They range from a simple static model to the highly sophisticated two lag model with autoregressive errors, but each of these models has its place. In any given circumstance a different model application might be called for, and this model gives you the ability to make changes without having to change the entire text of the program. The last of these techniques, though not highly used at the present, gives the field of finance a tool which might be used in the future to show the relationship that exists by the nature of financial analysis. There are thorns that go along with the use of any study of this type, and this appendix will cover the frequent errors.

Common Errors

Most of the errors associated with these programs are caused by failure to follow one or more of the formats for the various cards. These errors are covered in Martin's article, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models.[9] The most common errors listed by Martin are concerned with having a negative punch somewhere on control card one, and the user must recognize that at no time is a negative punch called for on any of the cards but the data deck. Care must be taken to insure the avoidance of these negative punch errors, as any one of them will terminate the operation of the program.

The second type of error common to this program stems from the inability to invert the correlation matrix. This stems from correlation being too high between any of the variables in the data cards. It does not include the lagged values of the same variables since this lagging is an internal process and is corrected for by the program. When any two or more of the exogenous variables are correlated above the .75 range, then there is a chance you will get a singular correlation matrix. Thus at least one of the variables which are correlated at this level must be dropped from the study. A regular regression program can be used to compute these correlation coefficients, and action can be taken on the basis of the correlation matrix produced by the linear regression technique.

Another type of problem has arisen for the first time during this study, and this alignment problem is that it is caused by two statements of the same kind within the program being in different locations on their respective cards. The alignment problem stemmed from a mispunch in the binary deck. To offset this, the program source deck was used to yield a new binary deck and to show that the problem stemmed from the binary deck, since the source deck and other binary decks had been used many times before. This must be avoided, since when this happens the operation is terminated before any printout is begun, and it makes it twice as hard to locate the exact source of the error.

This Appendix was aimed at acquainting the individual with the specifics and operation of the distributive two-lag program. In this section an attempt has been made to bring the highly technical language used by Martin and other statistically oriented individuals to the level of the business student. Bridging this gap cannot be accomplished by one study which attempts to break down the statistical language to something more understandable. More statistical orientation needs to be developed in the business curriculum if meaningful statistical research is expected.

TABLE XIII

DEFINITION OF VARIABLES USED IN APPENDIX C

Y_{t-m} . . the current and lagged values of the dependent variable (m=0,1,2,3)

Xi_{t-m} . . the current and lagged values of the exogenous variables associated with the lag parameter (m=0,1,2)

Zj_{t-m} . . the current and lagged values of the exogenous variables associated with the lag parameter (m=0,1,2)

Dk_t . . the current exogenous and/or dummy variables which are not associated with a lag

et . . the errors in the equation

$Ao(1-\lambda)(1-\mu)(1-\beta)$ . . the pure constant term

Ai . . the parameters of the set of exogenous variables, Xi_t, i=(1,...,A)

bj . . the parameters of the set of exogenous variables, Zj_t, J=(1,...,B)

λ . . the lag parameters associated with the set of exogenous variables, Xi_t

μ . . the lag parameters associated with the set of exogenous variables, Zj_t

β . . the first order autocorrelation coefficient

dk . . the parameters associated with the set of exogenous variables, Dkt

SOURCE: James E. Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968.

FOOTNOTES

[1]L. M. Koyck, "Distributed Lags and Investment Analysis," Contributions to Economic Analysis, North Holland Publishing: Amsterdam, Holland, 1954.

[2]Zvi Griliches and N. Wallace, "The Determinants of Investment Revisited, International Economics Review, Vol. VI, No. 3, (1965).

[3]James E. Martin, Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear, Single Equation Models, Processed Series P-585, Agricultural Experiment Station, Oklahoma State University, May, 1968.

[4]William Addison and Marc Nerlove, "Statistical Estimation of Long-Run Elasticities of Supply and Demand," Journal of Farm Economy, Vol. L (November, 1958), p. 869.

[5]Martin, p. 10.

[6]Addison and Nerlove, p. 871.

[7]Martin, p. 16.

[8]Ibid., p. 18.

[9]Ibid., p. 19.

VITA

Kelly Richard Guiou

Candidate for the Degree of

Master of Business Administration

Thesis: THE EXPLANATION OF STOCK PRICE MOVEMENTS USING A SINGLE EQUATION NON-LINEAR MODEL

Major Field: Business Administration

Biographical:

Personal Data: Born in Webb City, Missouri, February 15, 1942, the son of Kelly R. and Wilma L. Guiou.

Education: Attended grade school and junior high school in Webb City, Missouri; graduated from Central High School of Tulsa, Oklahoma in 1960; received the Bachelor of Science degree from Oklahoma State University with a major in Personnel Management in January, 1968; completed requirements for the Master of Business Administration on July 31, 1970.

Professional Experience: Employed by American Oil Company as a Transportation and Distribution Analyst from January, 1968 to September, 1968; employed by F. W. Woolworth as an Assistant to the Regional Store Opener during the summers of 1965, 1966, and 1967; employed by F. W. Woolworth as an Assistant Manager from September, 1963 to January, 1965.

Organizations: Society for the Advancement of Management and Mu Kappa Tau.