1972

Evaluation and prediction of common stock prices: A statistical study using correlation and regression analysis

James Richard Keyes
The University of Montana

THE EVALUATION AND PREDICTION OF COMMON STOCK PRICES: A STATISTICAL STUDY USING CORRELATION AND REGRESSION ANALYSIS

By

James R. Keyes

B.S.B.A., University of Minnesota, 1963

Presented in partial fulfillment of the requirements
for the degree of

Master of Business Administration

UNIVERSITY OF MONTANA

1972

Approved by:

Chairman, Board of Examiners

Dean, Graduate School

June 5, 1972
Date

6-5-72

ACKNOWLEDGEMENTS

I would first like to thank Doctor Bernard J. Bowlen, Chairman of my Examining Committee and Resident Administrator of the Air Force Institute of Technology Minuteman School. His encouragement and assistance during the writing of this paper and his guidance throughout my enrollment in the Minuteman Education Program have been invaluable.

I also wish to acknowledge the critical evaluation and constructive suggestions provided by Dean Rudyard B. Goode, Dean of the School of Business Administration and Examining Committee Member and Professor Carl J. Schwendiman, Examining Committee Member.

Appreciation must be expressed to Mrs. Floranne Boyd, my typist; to Mrs. Virginia Gilmore, AFIT Librarian, for her help in the research and preparation of this paper; and to my wife, Carmen, for her unending assistance, patience and understanding.

Finally, I am indebted to the Strategic Air Command and the Air Force Institute of Technology for establishing the Minuteman Education Program at Malmstrom Air Force Base.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

During calendar year 1971, the percentage of appreciation of individual common stocks listed on the New York Stock Exchange ranged from +462.3% (Winnebago Industries) to -58.7% (Boise Cascade). Of the 729 major U.S. companies rated in Forbes magazine's 24th Annual Report on American Industry, 279 declined in value over the past five years.[1] In the Forbes Report of January 1971, over half (354) of the 659 companies ranked had decreased in market value over the previous five year period.[2] How does the prospective investor pick the winners and avoid the losers? The stock analyst traditionally looks at various fundamental indicators of corporate performance in assessing the future market value of a stock. These data, together with information on the quality of management, the nature of the industry, and knowledge concerning general economic, social, and governmental factors, provide evidence which will aid in determining the prospects for the stock.

---

[1] Forbes, (January 1, 1972), p. 64.

[2] "23rd Annual Report on American Industry," Forbes, (January 1, 1971), p. 51.

An attempt will be made to answer the following question in this paper: Is it possible to analyze the statistical data regarding a company's past and present performance, while disregarding all qualitative information, and still say something meaningful about the future market value of the company's stock? If the data could be formulated in some way so as to predict future performance of stock prices, the alternatives available to the investor would be narrowed considerably. The investor's knowledge of management, the industry, and the economy coupled with the statistical determinations would serve to maximize investment profits.

Methodology

The principal statistical tools used in this study are correlation and regression analysis. Various statistical measures of performance have been compared with stock price appreciation to determine the degree of correlation. This comparison is accomplished with two separate groups of companies for the calendar year 1971. The rationale for the use of one year's data in the study is based on the cross-sectional approach in analyzing statistical information. Cross section analysis is structured on knowledge and interpretation of current forces rather than projection of past trends. The time-series approach, on the other hand, employs past variations in the data as a basis for projection. In this study, rather than following the pattern of historical data for individual companies, information available at the

beginning of the year will be compared with the change in stock price for many companies. The assumption is made here that variable data used in the comparisons is not time-sensitive and that feasible projections can be made from one year's corporate data.[3]

Those variables showing the highest correlation following the initial calculations were selected for further analysis. The data from the two groups of firms were used in computing the linear regression formula $Y = a+bX$. This formula was used in determining the Y (estimated stock price appreciation) values for a third randomly selected group of companies. These values, in turn, were compared with the actual percentage change in the market value of the stock for the companies studied. The percentage of occurrences in which the regression equation accurately predicts the direction taken by the market price of the stock was determined in the analysis. Furthermore, the data obtained through the above calculations and analysis provide a clue to a more sophisticated use of the regression analysis which will increase the accuracy of correctly predicting stock price direction and aid in maximizing investor profits. Finally, a brief introduction to multiple correlation and regression analysis is offered. Although only two independent variables were used in the multiple regression computations, there is an indica-

---

[3]Louis Siegelman and Milton H. Spencer, Managerial Economics, Decision Making and Forward Planning, (Homewood, Ill.: Richard D. Irwin, Inc., 1964), p. 122.

tion that further study using more companies and a greater number of independent variables could prove rewarding in stock price prediction using multiple regression.

Statistical Studies in the Field of Investment Analysis

The use of correlation and regression in investment analysis, although relatively new, has been suggested and evaluated by several individuals in the field of security analysis and statistics. Mathematical techniques for the analysis of portfolios of securities have been developed by Markowitz. Portfolio models for the efficient diversification of investments are outlined in his monograph.[4] Sharpe has extended Markowitz' work by presenting a simplified model for portfolio analysis.[5] Using regression analysis, Mlynarczyk attempted to determine from empirical data the effects of one set of specific alternative accounting methods on security prices by means of multivariate statistical techniques.[6] Whitbeck and Kisor studied price/earnings multiples and growth and stability of earnings in determining

---

[4]Harry M. Markowitz, Portfolio Selection, Efficient Diversification of Investments, (New Haven, Conn.: Yale University Press, 1959)

[5]William F. Sharpe, "A Simplified Model for Portfolio Analysis," Frontiers of Investment Analysis, (Scranton, Pa.: International Textbook Company, 1971), pp. 90-107.

[6]Francis A. Mlynarczyk, "An Empirical Study of Accounting Methods and Stock Prices," Frontiers of Investment Analysis, pp. 498-517.

the price to be paid for stock.[7] In an analysis of 45 major companies, Cootner evaluated the theories of random versus systematic changes in the price of stocks using statistical techniques.[8] An analysis of formula planning techniques was conducted by Dince using regression equations involving the Dow-Jones Industrial Average and Gross National Product.[9] A paper by Moore concerns itself with successive movements in the price of common stocks. Moore's study is oriented toward estimation of characteristics of stock behavior through correlation analysis.[10]

The studies described above are essays which have appeared in various financial and statistical journals. There have been few books written in which a comprehensive discussion of the interrelationship of quantitative analysis and investments is provided. One of the most useful books in the field is by J. Peter Williamson.[11] Various statistical

---

[7]Manown Kisor, Jr. and Volkert S. Whitbeck, "A New Tool in Investment Decision Making," Frontiers of Investment Analysis, pp. 567-582.

[8]Paul H. Cootner, "Stock Prices: Random vs. Systematic Changes," Frontiers of Investment Analysis, pp. 690-711.

[9]Robert R. Dince, "Another View of Formula Planning," The Investment Process, (Scranton, Pa.: International Textbook Company, 1970), pp. 443-452.

[10]Arnold B. Moore, "Some Characteristics of Changes in Common Stock Prices," The Random Character of Stock Market Prices, (Cambridge, Mass.: The M.I.T. Press, 1964), pp. 139-161.

[11]J. Peter Williamson, Investments, New Analytic Techniques, (New York: Praeger Publishers, 1970)

techniques such as regression analysis and simulation are outlined in this text with specific reference to their application in analyzing common stocks. Williamson also included work on portfolio selection and stock-selection techniques. The models developed by Markowitz and Sharpe and the equations tested by Whitbeck and Kisor, described above, are stated and evaluated in these sources.

The Focus of This Paper

The various analytical studies in the field of security analysis have been, for the most part, of a macro nature. They have dealt with the broad and general aspects of the economy and investments. Furthermore, they have been highly theoretical, employing advanced mathematical and statistical methods. The focus of this paper, on the other hand, will be directed toward individual investment in a common stock. The variables employed are those basic to corporate performance and are frequently used by the investor in determining a stock purchase. The data references are also those used by the typical investor and are easily understood.

It must be stressed from the start that this study is limited in nature due to the number of firms evaluated and the number of variables analyzed. The equations and models presented are formulated to test the methodology of these techniques. There are indications, however, that a more detailed, comprehensive study could be especially worthwhile in providing data which would assist the investor in knowing

which common stocks to purchase. Even here, the information obtained through correlation and regression analysis would not give all the answers. It would be useful, not to replace, but to supplement traditional fundamental and technical stock market guidance.

CHAPTER II

VARIABLES USED IN THE STUDY

Dependent Variable

The dependent variable, Y, used throughout this study is the percentage change in common stock market price. Since the analysis is based on the calendar year 1971, the percentage change is from the closing market price for 1970 to the closing market price for 1971. The sources of these data were the annual financial tabulations of all listed stocks reported by the Associated Press which appeared in major metropolitan newspapers on January 2, 1972.[12]

The rationale for the use of stock price change as the Y variable was that appreciation of market value is the primary goal of investors interested in growth of their investment in common stocks. Even those investors who are primarily interested in income and seek out rising dividends and high yield still take a considerable interest in whether a stock appreciates or depreciates in value. The goal of this study was to find a statistical method to aid in avoiding purchase of those stocks which will decrease in market

---

[12]Annual Financial Tabulations, Las Vegas Review-Journal, Jan. 2, 1972, p. 29.

value. All investors and potential investors share in this goal. The use of the calendar year period was a matter of convenience - data were readily available for analysis for this length of time.

Independent Variables

There are 16 independent variables used initially in this study. They are as follows:

1. 5-Year Return on Equity.
2. Latest 12-Month Return on Equity.
3. 5-Year Return on Total Capital.
4. 5-Year Annual Sales Growth.
5. 5-Year Annual Earnings Per Share Growth.
6. 5-Year Price Gain.
7. Dividend Yield.
8. Price-Earnings Ratio.
9. The difference between the current price-earnings ratio and the median price-earnings ratio for the previous 15 year period.
10. The Value Line Quality (Dependability) rating. This rating is now called "safety".
11. The Value Line 12-Month Performance Rating.
12. The Value Line estimated price appreciation for the next 3 to 5 year period.
13. The percentage change from previous year's sales to current year's sales.
14. The percentage change from current year's sales to

the following year's sales, as estimated by Value Line.

15. The percentage change from previous year's earnings per share to current year's earnings per share.

16. The percentage change from current year's earnings per share to the coming year's expected earnings per share, as estimated by Value Line.

The source of data for the first six variables listed above was Forbes magazine's 23rd Annual Report on American Industry.[13] Each year Forbes ranks the top U.S. companies on three basic scales: Profitability, Growth, and Stock Market Performance. Profitability measures how efficently management is using the assets entrusted to it. The profitability measures include the 5-Year Return on Equity, the Latest 12-Month Return on Equity, and the 5-Year Return on Total Capital. The Five-Year Return on Equity is the average profit which management has made on stockholder investment over five years. The Latest 12-Month Return on Equity is the same figure for the most recent 12-month reporting period. Five-Year Return on Total Capital is, like the first yardstick, an average of the last five years. But, while return on equity is the profit management made on stockholders' investment, this figure is the profit made on total capital-stockholders' investment and debt. The use of five-year averages for return on equity and capital tends to average out the high and low values over a period of time resulting in a more

---

[13]Forbes, (January 1, 1971), p. 70.

reliable variable for forecasting purposes. The second primary test of management effectiveness, growth, is represented by variables numbered 4 and 5. Both the sales and earnings per share growth figures are five-year annual compounded rates. The final Forbes yardstick of stock market performance is shown by the 5-Year Price Gain. This is based on a comparison of a company's stock price at the end of 1970 to its price five years earlier, the last trading day of 1965.

The source of data for the remaining independent variables used in this study is the Value Line Survey.[14] The Dividend Yield is the dividend declarations estimated for the next 12 months divided by the price of the stock. The Price Earnings Ratio (P/E) is obtained by dividing the earnings estimated for the current 12 months into the price. Value Line also reports the median 15-year price earnings ratio based on the average of the 7th, 8th and 9th highest price-to-earnings ratios. The period used for the data in this paper is 1955-1969. Variable #9 represents the difference between this median and the latest P/E. The Quality rating is Value Line's rating for dependability and safety for the long term. The higher the rating, the better the chance (in the opinion of Value Line's analysts) of risk avoidance. The 12-Month Performance rating is another rating unique to Value Line which represents their analysts' price expectations over

---

[14]Value Line Survey, (New York: Arnold Bernhard & Co., Inc., 1970, 1971)

the next 12 months. Variable number 12 is the estimated percentage of appreciation for the period 3 to 5 years from now. The percentage figures for Variables 13 through 16 were derived by comparing sales and earnings for the past year and the current year and between the current year and an estimate for the next year.

The 16 independent variables chosen for use in this study are frequently used and readily understood by most investors. Not all statistical measures of corporate performance are represented. The balance sheets and income statements of most companies would yield more data should one care to avail himself of it. However, it was felt the items of information which were of principal interest to the investor were contained on this list. The reference sources utilized were chosen for much the same reason as were the variables; both Forbes and Value Line are accessable to most investors and the data contained therein are easily understood.

CHAPTER III

CORRELATION ANALYSIS

The Study

The objective of this study was to determine if one could use data concerning a company's past and present performance to statistically predict the future movement of its stock price. The first step of this procedure was to find those factors which had a relatively high correlation with the price of the stock.

The initial group of companies which are worked with in determining this degree of correlation are the 30 corporations comprising the Dow-Jones Industrial Average.[15] This group has been chosen for the first series of calculations because: (1) it is a convenient number of observations with which to work, (2) the Dow-Jones Industrial Average is the most popular stock market average and is frequently used for comparisons, and (3) data for these companies are readily available. The complete breakout of data for all 30 companies for each of the 16 independent variables is shown in Appendix B. The 16 independent variables are defined symbolically as X_1

---

[15]Maurice L. Farrell, The Dow Jones Investor's Handbook, (Princeton, N.J.: Dow Jones & Company, Inc., 1972), P. 11.

through X_{16}. The dependent variable, stock market price change, is defined as Y. The sums EX, $(EX)^2$, EX^2, EY, $(EY)^2$, EY^2, and EXY, which are required for calculations later in this study, are shown in Appendix C.

The Coefficient of Correlation

The following statistic, the coefficient of correlation, is the most widely used measure of the strength of linear relationships between two variables and will be used throughout this study.

$$r = \frac{n \cdot EXY - (EX)(EY)}{\sqrt{n \cdot EX^2 - (EX)^2}\ \sqrt{n \cdot EY^2 - (EY)^2}}$$

The coefficient of correlation is a measure of whether or not it is reasonable to say that a linear relationship (correlation) exists between X and Y. If r is close to 0, the relationship is weak or nonexistent; the closer r is to +1 or -1, the stronger the linear relationship positively or negatively.[16]

Correlation Coefficients: the Dow Jones 30 Industrials

Coefficients of correlation (r) were computed for each of the 16 independent variables based on the data shown in

---

[16]John E. Freund and Frank J. Williams, Modern Business Statistics, (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1958), p. 310.

Appendix B. The r values for each variable are shown below.

INDEPENDENT VARIABLE	COEFFICIENT OF CORRELATION (r) DOW-JONES 30 INDUSTRIALS
X_1	.127
X_2	.011
X_3	.182
X_4	-.006
X_5	-.096
X_6	.237
X_7	-.151
X_8	.622
X_9	.458
X_{10}	.146
X_{11}	-.181
X_{12}	-.287
X_{13}	-.028
X_{14}	.097
X_{15}	-.086
X_{16}	.083

Based on the results shown above, it was decided that at least four X variables were worthy of further analysis. These four, all of which had a r value larger than ± .200, are X_6, X_8, X_9, and X_{12}. These X values symbolize respectively, 5-Year Price Gain, Price Earnings Ratio, the difference between the current price earnings ratio and the median P/E for the previous 15 year period, and the estimated price

appreciation for the next 3 to 5 year period. It should be pointed out here that few of these variables had high coefficients of correlation. On a relative basis, however, they had considerably higher r values then the other variables used in the test.

Correlation Coefficients: 30 Randomly Selected Firms

A second group of 30 firms was chosen randomly for a comparison analysis with the initial group of 30 Dow-Jones Industrials. These 30 companies were picked from a table of random numbers using accepted procedures as specified in the Standard Mathematical Tables.[17] The random numbers were limited to numbers 1 through 647, the number of companies ranked in Forbes 23rd Annual Report on American Industry. The data for these 30 firms for each of the 16 independent variables used earlier are shown in Appendix D. The summary totals required in the computation of the coefficients of correlation for this second group of companies for each of the independent variables are depicted on the following page with the original 16 r values shown for comparison.

The r values which were the highest for the first group of stocks continued to be quite high for the second group. To further establish the accuracy of these calculations, the

---

[17]Samuel M. Selby, Standard Mathematical Tables, Sixteenth Edition, (Cleveland, Ohio: The Chemical Rubber Co., 1968), pp. 594-98.

r values were computed for the entire combined group of 60 companies with the following results: X_6, .360; X_8, .336; X_9, .298; and X_{12}, -.270.

INDEPENDENT VARIABLE	COEFFICIENTS OF CORRELATION (r) 30 Firms Randomly Selected	Dow-Jones 30 Industrials
X_1	.018	.127
X_2	.140	.011
X_3	.121	.182
X_4	.136	-.006
X_5	.261	-.096
X_6	.333	.237
X_7	-.060	-.151
X_8	.209	.622
X_9	.141	.458
X_{10}	.075	.146
X_{11}	.169	-.181
X_{12}	-.215	-.287
X_{13}	-.305	-.028
X_{14}	-.077	.097
X_{15}	-.001	-.086
X_{16}	-.299	.083

In the test involving the 30 randomly selected firms, X_5, X_{13}, and X_{16} all had significantly higher r values than

they had earlier. Because of this fact, the r values were recomputed for the combined 60 stocks for these X variables also. The variables X_{13} and X_{16} had very low r values of -.134 and -.087 respectively for this test and were eliminated from further consideration. The variable X_5, however, had a coefficient of correlation of .197 overall, and it was decided to include it along with the previously selected four variables for further analysis. Summary computations required for variables X_5, X_6, X_8, X_9, X_{12}, X_{13}, and X_{16} are shown in Appendix F.

CHAPTER IV

REGRESSION ANALYSIS

The Regression Equation

The objective of this statistical investigation was to predict or forecast changes in common stock prices. Whenever possible, statisticians in business and economics strive to express relationships between variables, namely, relationships between quantities that are known and quantities that are to be predicted, in terms of mathematical equations. One of the simplest and most widely used equations for expressing relationships in various fields is the linear equation which is of the form

$$Y = a + bX$$

In an effort to test for a linear relationship between the X and Y variables, data for the five independent variables selected in the previous chapter were plotted as points on graph paper for each of the 60 companies in this study. These graphs are depicted on the following pages as Figures 1 through 5. The change in common stock market price is plotted on the Y axis and the values for the independent variable on the X axis. This arrangement follows a convention by which the variable which is to be predicted (stock

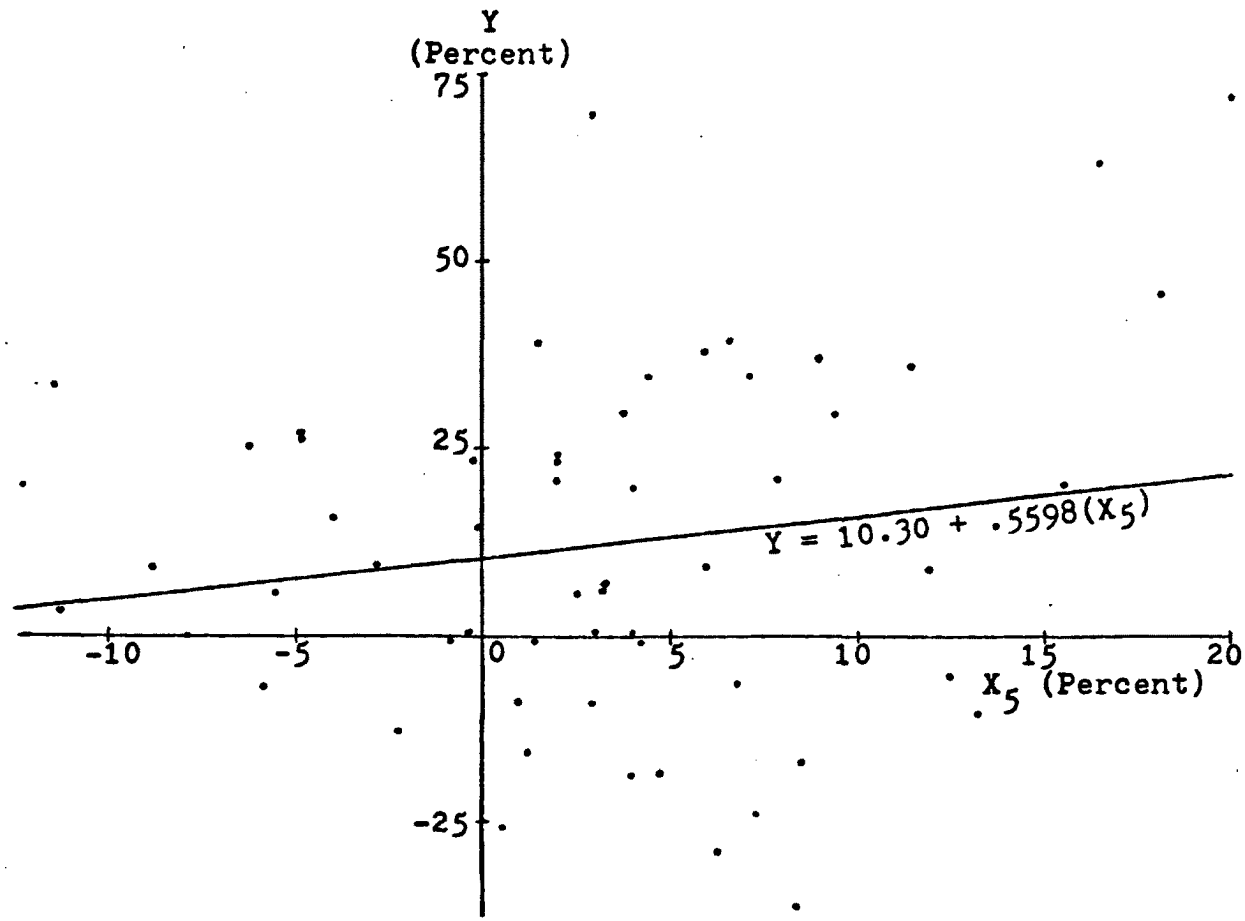

Fig. 1.--Scatter Diagram, Change in Common Stock Market Price (Y) versus 5-Year Annual Earnings Per Share Growth (X_5), 1971 Data for 60 Companies.

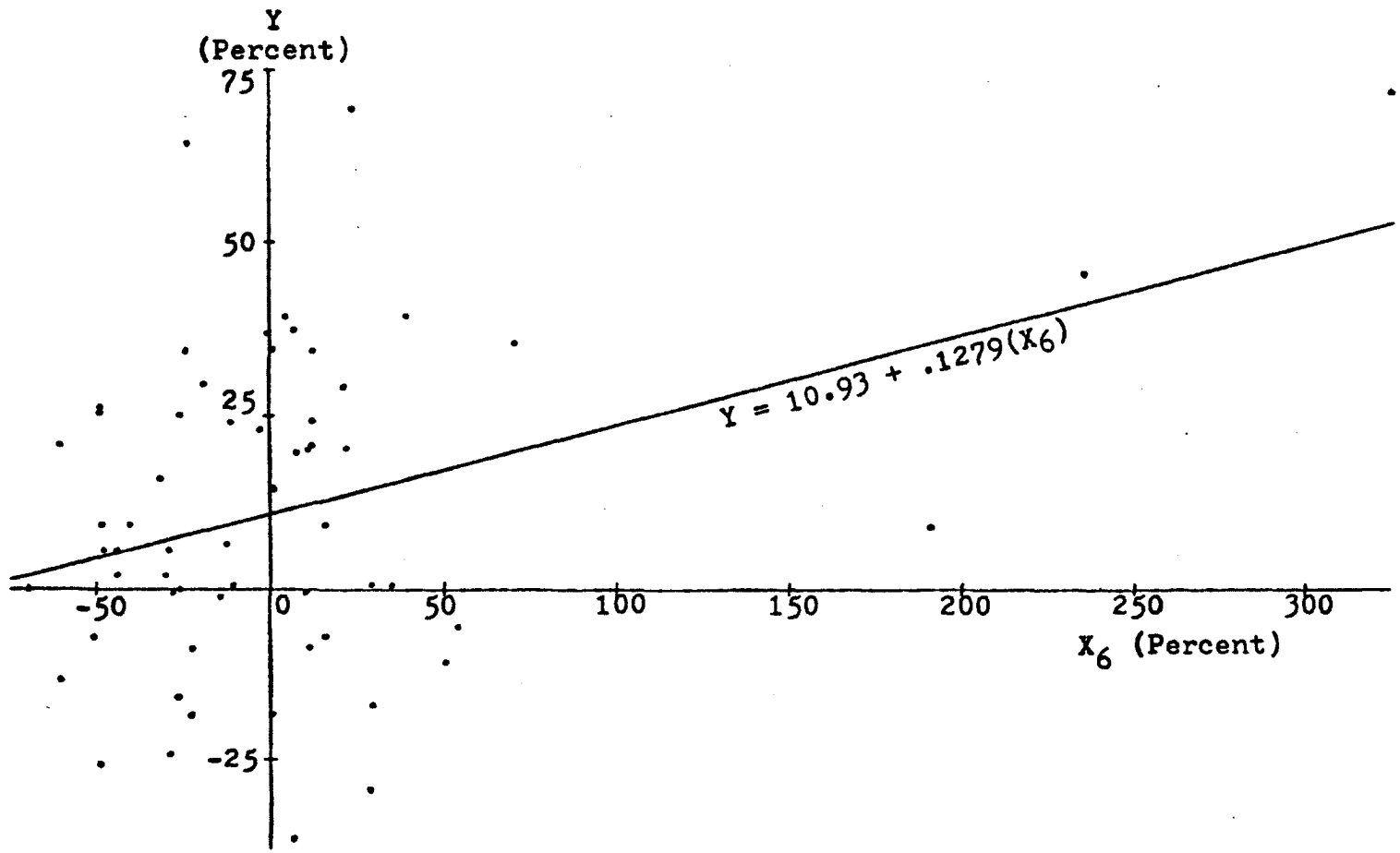

Fig. 2.--Scatter Diagram, Change in Common Stock Market Price (Y) versus 5-Year Price Gain (X_6), 1971 Data for 60 Companies.

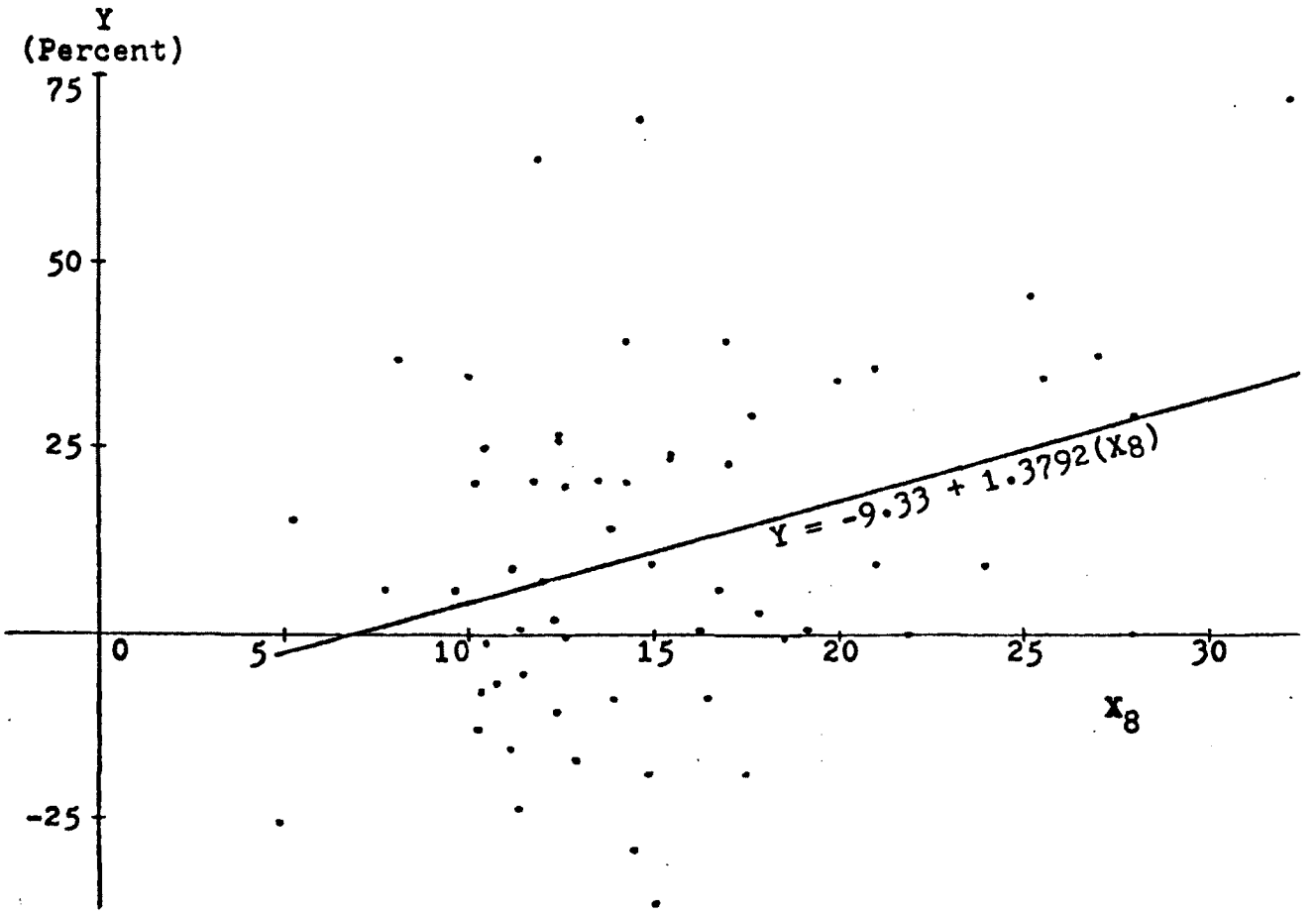

Fig. 3.--Scatter Diagram, Change in Common Stock Market Price (Y) versus Price Earnings Ratio (X_8), 1971 Data for 60 Companies.

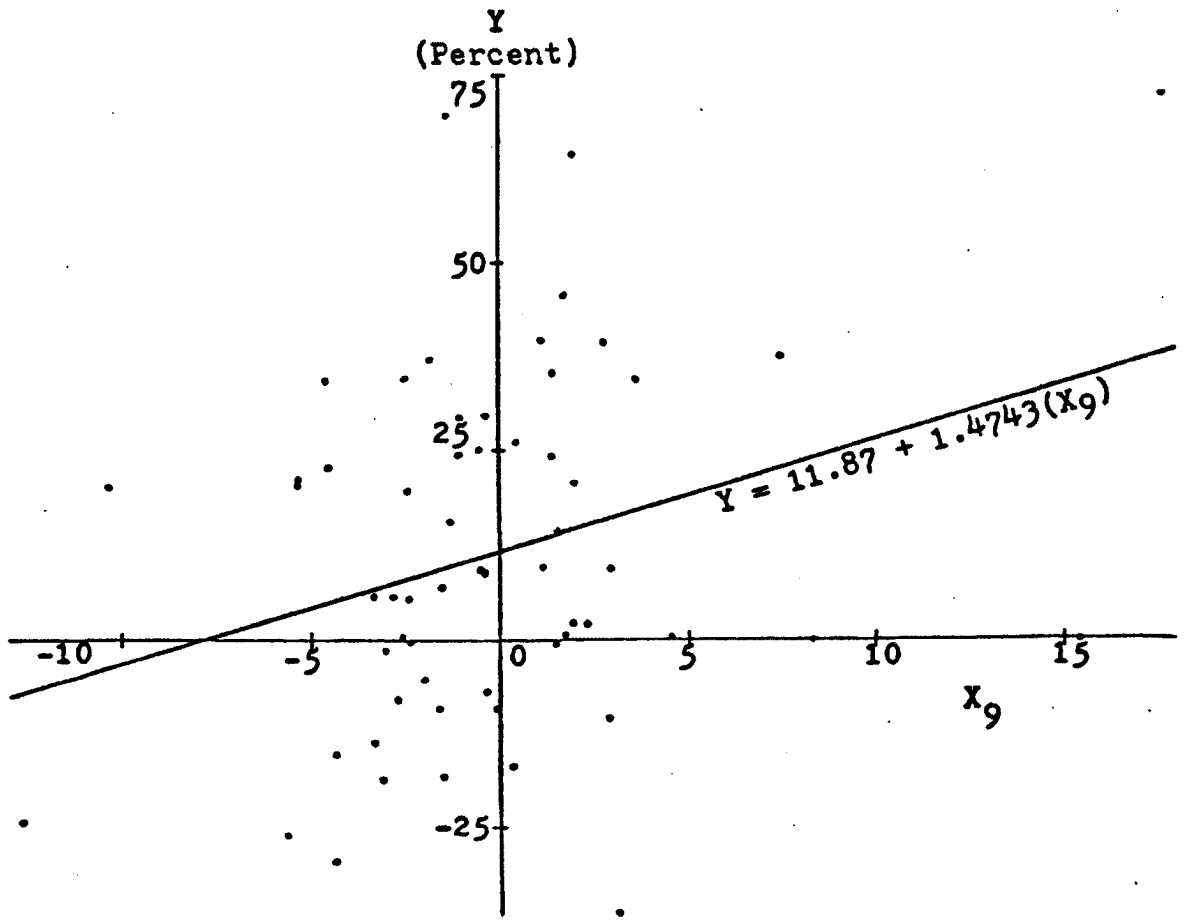

Fig. 4.--Scatter Diagram, Change in Common Stock Market Price (Y) versus The Difference Between the Current Price Earnings Ratio and the Median Price Earnings Ratio for the Previous 15-Year Period (X_9), 1971 Data for 60 Companies.

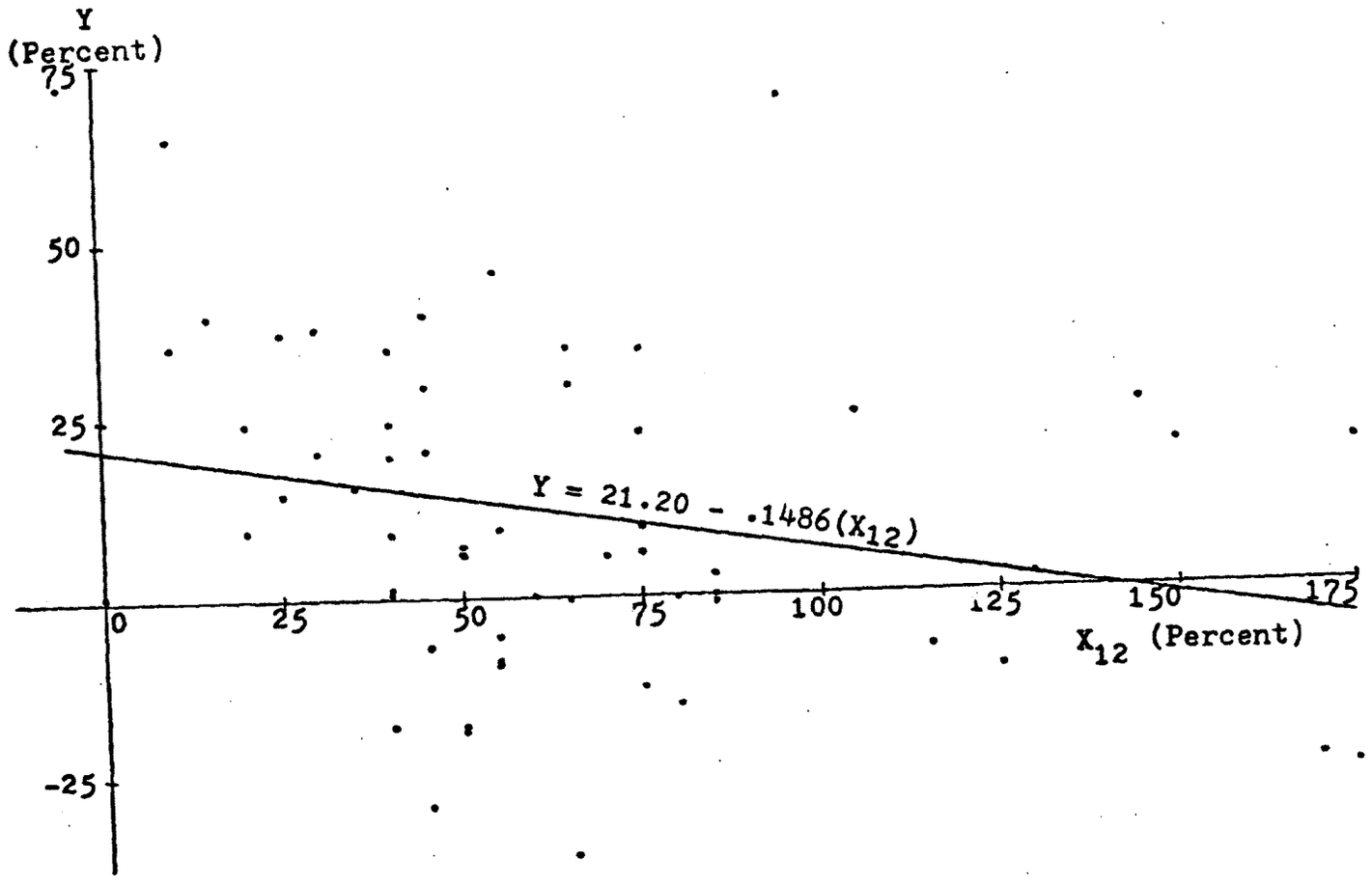

Fig. 5.--Scatter Diagram, Change in Common Stock Market Price (Y) versus Estimated Price Appreciation for the Next 3-5 Year Period (X_{12}), 1971 Data for 60 Companies.

price) is associated with the Y axis, and the variable to be used for making the prediction is plotted on the X axis. Each point on a scatter-diagram refers to a pair of observations pertaining to a particular company. On the basis of the patterns shown on the 5 scatter diagrams, it was decided to use the equation $Y = a + bX$ for forecasting purposes. Although the graphs indicate considerable dispersion of the points, the relationship between the variables can best be described as a linear association.

The mathematically computed line of regression, $Y=a+bX$, will be used in this study in an effort to forecast stock market prices. The a and b symbols are numerical constants and once they are known one may calculate a predicted value of Y for any given value of X by direct substitution. The b value represents the "slope" of the regression line. The amount of change in Y, associated with a given change in X, is indicated by the slope. The formulas required for the computation of a and b may be stated as follows[18]:

$$b = \frac{n\,(EXY) - (EX)\,(EY)}{n\,(EX^2) - (EX)^2}$$

$$a = \frac{EY - b \cdot EX}{n}$$

---

[18]Freund and Williams, Modern Business Statistics, p. 292.

In the above formulas, n is the number of pairs of observations, EX and EY are the sums of the given X's, and Y's, EX^2 is the sum of the squares of the X's, and EXY is the sum of products of the corresponding X's and Y's. All of these values were used earlier in computing the coefficient of correlation (r).

Forecasting Stock Prices

The regression formula, Y = a +bX, was computed for each independent variable (X_5, X_6, X_8, X_9, and X_{12}) showing a relatively high positive or negative correlation. The data for the initial 60 companies, as shown in Appendix F, were used in these calculations. The determination of b and a values for each of the X variables resulted in the following regression equations:

X_5 : $Y = 10.30 + .5598\ X$

X_6 : $Y = 10.93 + .1279\ X$

X_8 : $Y = -9.33 + 1.3792\ X$

X_9 : $Y = 11.87 + 1.4743\ X$

X_{12}: $Y = 21.20 - .1486\ X$

The straight lines expressed by each of the above formulas are shown on the scatter diagrams in Figures 1 through 5. Substitution of data for an independent variable in any of these equations results in a computed stock price, or Y value, which falls on the respective regression line. The regression equations may, therefore, prove useful in the fore-

casting of future stock prices.

A new group of companies was selected to test the five regression formulas determined above. Fifty firms were picked using accepted procedures for choosing random numbers.[19] These companies, together with their associated independent variable data, are listed in Appendix G. The regression formulas were used to compute estimated stock price appreciation (Y) values for each company for each of the five X variables. The 250 computed Y values were tabulated for comparison with the actual stock price changes in Appendix H. The purpose of this excercise was to discover the percentage of occurrences in which the regression equations accurately predicted the direction taken by the market price of the stock. Interestingly, the percentages for each of the X variables were consistently high and similar to one another. The proportion of occurrences in which the computed Y correctly predicted the direction taken by the actual stock price change and the resulting percentages are shown below:

X_5	35/50	70%
X_6	36/50	72%
X_8	36/50	72%
X_9	36/50	72%
X_{12}	37/50	74%

---

[19]Selby, Standard Mathematical Tables, pp. 594-98.

Analysis of certain individual X variables serves to highlight the positive and negative aspects of the results described above. In the case of X_5, 5-Year Annual Earnings Per Share Growth, the direction of stock price movement was correctly anticipated in 35 out of 50 instances; however, the regression formula failed to indicate direction of change for 14 stocks which decreased in value. Furthermore, one stock which the formula indicated would go down, Lykes-Youngstown, in fact increased in price. On the positive side, the regression equation appears to be a better indicator than merely observing the direction of 5-year earnings per share growth and judging that the price change would move accordingly. Using this approach, one would have correctly predicted only 27 out of 50, or 54%. The same comments apply with regard to other variables used. In the case of X_6, predictions based on only the 5-year price gain would have yielded correct results in 42 percent of the instances. For X_9, if one bought or did not buy a stock based on whether the current P/E was greater or less than the historical median, he would have been right in 60% of the cases.

Increasing Accuracy in Forecasting Prices

The analysis conducted thus far, while interesting, did not produce results which are highly conclusive. A means of utilizing the data to better advantage in forecasting stock price movement is provided by a further examination of the computed Y values in Appendix H. For each

of the 50 firms there are 5 computed Y values. It may be observed in studying these values that the computed Y values for the stocks which depreciate in value have something in common. In the great majority of these cases, one or more of the predicted stock prices, although positive, is a relatively low positive number. Based on this determination, the following principle will be established for the purchase of stocks using regression analysis: Only those stocks will be bought where all five of our computed Y values are greater than +10.0%. If this were done with the 50 companies under study, a total of 13 stocks would have been purchased. Twelve of these appreciated in value, while only one decreased in price during the course of the year. The average appreciation in value was 34.77%. The firms, with their percentage stock price change, are shown below.

Leaseway Transport	+136.7%
Kendall	+55.8
Delta Airlines	+45.4
Pepsico	+32.5
ARA Services	+30.4
Chesebrough-Pond's	+28.6
Broadway-Hale	+25.7
Marshall Field	+25.6
Sybron	+22.0
Emerson Electric	+21.6
H J Heinz	+20.1

Singer	+18.8
Amerada Hess	-11.2
Average Change	+34.8%

The rule of buying only those stocks for which all 5 computed Y values are above +10.0% was further tested with 20 additional companies. From all those stocks listed on the New York Stock Exchange which had X variable data available, the 10 with the greatest appreciation and the 10 with the greatest decrease in market value for 1971 were chosen. These companies, together with their independent variable data, are listed in Appendix I. The actual and computed Y values are shown in Appendix J. Using the rule described above, 3 of these stocks would have been purchased with a combined average appreciation of +64.7%. These stocks, with their percentage market value change, are as follows:

First Natl Stores	-37.1%
Genuine Parts	+94.5
Leaseway	+136.7
Average Change	+64.7%

CHAPTER V

MULTIPLE REGRESSION ANALYSIS

Description

Although there are many situations in which fairly accurate predictions can be made of one variable in terms of another, it stands to reason that these predictions should be improved if they took into account additional relevant information. Among the many equations that can be used to express relationships between more than two variables, the most widely used in statistical work is the multiple linear regression equation of the form

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots$$

where Y is the variable which is to be predicted while X_1, X_2, X_3, . . ., are known variables on which the prediction is to be based. Of the latter, there may be two, three, four, or more, depending on the nature of each individual problem.[20]

The constants b_1, b_2, and b_3, which show, in this case, the average increase in Y associated with unit increases in X_1, X_2, and X_3, are termed partial regression coefficients.

---

[20]R. Clay Sprowls, Elementary Statistics, (New York, N.Y.: McGraw-Hill Book Company, 1955), p. 257.

The constant b_1, for example, is termed the partial regression of Y on X_1, holding X_2 and X_3 constant, and b_2 is termed the partial regression of Y on X_2, holding X_1 and X_3 constant. Therefore, b_1, for example, measures the average change observed in Y with unit changes in X_1, determined while simultaneously eliminating from Y any variation accompanying changes in $X_2$and X_3.

Multiple regression analysis enables one to measure the joint effect of any number of independent variables upon a dependent variable. The average relationship between these variables is described by the multiple regression equation and this relationship is used to predict the dependent variable. The concepts and techniques in this chapter, therefore, are just extensions of those in simple regression analysis. Measuring the simultaneous influence of several factors provides a more powerful and realistic tool of analysis than considering only one independent variable.

Predicting Stock Prices

The availability of regression formulas which describe the combined effect of several X variables on a Y value opens up the possibility of using this technique in forecasting the future market value of stocks. In this section of the study two independent variables in association with a dependent variable were analyzed. Analysis of several independent variables would be accomplished following generally the same format. Extension of this procedure to describe the effect

of many variables, however, creates highly complex and lengthy calculations necessitating the use of computer programming.

Using two independent variables, the multiple regression equation is expressed in the form

$$Y = a + b_1X_1 + b_2X_2$$

To obtain the values a, b_1, and b_2, several new formulas must be introduced. The value of a is given by the equation

$$a = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2$$

The b values are found by solution of the following formulas[21]:

$$b_1 = \frac{(Ex_2^2)(Ex_1y) - (Ex_1x_2)(Ex_2y)}{D}$$

and

$$b_2 = \frac{(Ex_1^2)(Ex_2y) - (Ex_1x_2)(Ex_1y)}{D}$$

where

$$D = (Ex_1^2)(Ex_2^2) - (Ex_1x_2)^2$$

---

[21]George W. Snedecor and William G. Cochran, Statistical Methods, (Ames, Iowa: The Iowa State University Press, 1967), p. 383.

In describing the use of multiple regression, the independent variables X_6 and X_8 have been used. In the computation of the coefficients of correlation (r) earlier in the study, these two variables had r values of +.360 and +.336, which were the highest of the 16 independent variables evaluated. It is appropriate at this point in the study to insert a caution in the use of multiple regression concerning colinearity. When the independent variables in a multiple regression are highly correlated with each other, the net (partial) regression coefficients may be unreliable. If X_6 and X_8 moved together, it would be difficult to distinguish their separate effects on Y. While colinearity affects the reliability of individual variables in the regression, it may not alter the predictive power of the total regression equation. The errors of the various regression coefficients tend to compensate for each other in the estimate of the dependent variable. In the situation under study, it was assumed that the variables X_6 and X_8, 5-Year Price Gain and Price Earnings Ratio respectively, were not correlated in such a way as to affect the forecasting of the Y variable, stock price change. In a more extensive analysis of multiple regression, appropriate statistical techniques would be used to account for colinearity between variables.[22]

---

[22]William A. Spurr and Charles P. Bonini, Statistical Analysis for Business Decisions, (Homewood, Ill.: Richard D. Irwin, Inc., 1967), p. 610.

To find the terms in the formulas for b_1 and b_2, the equations below will be used. Since it has been decided to use the variables X_6 and X_8, the elements x_1 and x_2 are indicated as x_6 and x_8 respectively.[23]

$$Ex_8^2 = EX_8^2 - n(\bar{X}_8)^2$$
$$Ex_6y = EX_6Y - n(\bar{X}_6)(\bar{Y})$$
$$Ex_6x_8 = EX_6X_8 - n(\bar{X}_6)(\bar{X}_8)$$
$$Ex_8y = EX_8Y - n(\bar{X}_8)(\bar{Y})$$
$$Ex_6^2 = EX_6^2 - n(\bar{X}_6)^2$$

The observations used here are the 60 companies described during the initial phase of this study - the 30 Dow-Jones Industrials plus the 30 randomly selected firms. The data necessary for computation of the equations above is tabulated in Appendix F. These calculations result in the following: $Ex_8^2 = 1908$, $Ex_6y = 32724$, $Ex_6x_8 = 11254$, $Ex_8y = 2631$, $Ex_6^2 = 255897$, $\bar{Y} = 11.533$, $\bar{X}_6 = 4.708$, and $\bar{X}_8 = 15.128$. Substitution of these values into a, b_1, and b_2, give one the following:

$$a = -1.65$$
$$b_1 = .0908$$
$$b_2 = .8434$$

---

[23]Taro Yamane, Statistics, An Introductory Analysis, (New York, N.Y.: Harper & Row, 1967), p. 756.

The resulting multiple regression equation is as follows:

$$Y = -1.65 + .0908(X_6) + .8434(X_8)$$

Substituting the X_6 and X_8 values found in Appendix G into the above equation results in the series of computed Y values which are listed in Appendix K. The relationships discovered using the multiple regression method correctly predicted the direction of stock price movement in 36 out of the 50 cases for a percentage of success of 72%. The results, therefore, in this limited analysis of multiple regression, were virtually the same as the results using simple linear regression. If must be emphasized, however, that the purpose of this section of the study was to introduce multiple regression into the calculations only on a limited basis to test the feasibility. The fact that the results were not especially noteworthy does not negate the possibility of successfully using this technique with more companies and a greater number of independent variables.

CHAPTER VI

SUMMARY AND CONCLUSION

The statistical data regarding a company's performance was analyzed in this paper using correlation and regression analysis in attempting to predict the direction of stock prices. First, several variables related to the price of common stock were compared with stock price appreciation to determine the degree of correlation. Those variables showing the highest correlation were selected for further analysis. The linear regression formula $Y = a + bX$ was computed using the data for two groups of companies. Estimated Y values (stock price change) were calculated using those independent variables previously selected. Through this procedure, it was possible to show the percentage of occurrences in which the regression equation accurately predicted the direction taken by the market price of the stock. The statistical data was further analyzed to increase the accuracy of the forecasting, and finally, the subject of multiple regression analysis was introduced to indicate its potential in helping to forecast stock prices.

The evaluation carried out in this paper was limited in nature. More elaborate studies could be conducted using many firms and numerous independent variables. The increased

utilization of multiple correlation and regression analysis using a computer would assist immeasurably in accomplishing the highly complex and lengthy calculations involved in this procedure. There are indications from the results of this study that useful data can be obtained using these techniques. As stated in the introduction, the information arrived at through correlation and regression analysis would never give one all the answers. Regardless of how extensive the investigation, the data obtained would be useful, not to replace, but to supplement the traditional guidance received from fundamental and technical sources.

APPENDIX A

Independent Variables Defined

X_1 5 - Year Return on Equity

X_2 Latest 12 - Month Return on Equity

X_3 5 - Year Return on Total Capital

X_4 5 - Year Annual Sales Growth

X_5 5 - Year Annual Earnings Per Share Growth

X_6 5 - Year Price Gain

X_7 Dividend Yield

X_8 Price Earnings Ratio

X_9 The difference between the current price-earnings ratio and the median price-earnings ratio for the previous 15 year period.

X_{10} The Value Line Quality (Dependability) rating. This rating is now called "Safety."

X_{11} The Value Line 12 - month Performance rating

X_{12} The Value Line estimated price appreciation for the next 3 to 5 year period.

X_{13} The percentage change from previous year's sales to current year's sales.

X_{14} The percentage change from current year's sales to the following year's sales, as estimated by Value Line.

X_{15} The percentage change from previous year's earnings per share to current year's earnings per share.

X_{16} The percentage change from current year's earnings per share to the following year's earnings per share, as estimated by Value Line.

Dependent Variable Defined

Y The percentage change in common stock market price.

APPENDIX B

Data Tabulations: Dow-Jones 30 Industrials

		Y	X_1	X_2	X_3	X_4
1.	Allied Chemical	+20.7	9.2	6.7	6.5	1.9
2.	Aluminum Co. (Amer.)	-23.8	11.0	9.9	7.0	5.6
3.	Am. Brands	-7.2	14.1	14.8	11.6	14.5
4.	American Can	-16.0	10.9	9.3	7.7	7.2
5.	Am. Tel. & Tel.	-8.4	9.5	9.1	7.1	8.3
6.	Anaconda	-26.2	9.2	7.9	8.3	1.6
7.	Bethlehem Steel	+26.8	8.1	5.9	7.2	4.0
8.	Chrysler	+2.2	8.9	-	7.4	6.5
9.	DuPont	+8.7	15.7	13.8	13.9	4.3
10.	Eastman Kodak	+28.6	25.7	20.0	24.8	11.4
11.	General Electric	+33.6	13.1	7.1	10.7	5.2
12.	General Foods	-18.0	16.7	15.8	15.0	6.6
13.	General Motors	0.0	18.5	12.7	17.4	3.0
14.	Goodyear	+0.4	12.9	10.6	9.8	7.5
15.	Inter. Harvester	+8.6	7.3	5.4	6.4	2.5
16.	Inter. Nickel	-29.2	15.7	18.5	13.6	7.7
17.	Inter. Paper	-0.7	9.6	8.8	8.7	6.7
18.	Johns-Manville	+0.6	10.5	8.4	10.3	2.7
19.	Owens-Illinois	-18.4	11.6	11.1	7.0	11.0
20.	Procter & Gamble	+35.3	17.8	18.1	15.1	8.0
21.	Sears Roebuck	+34.4	14.5	13.0	12.8	7.7
22.	St'd Oil of Calif.	+5.3	11.5	10.0	10.5	9.6
23.	St'd Oil of N.J.	+0.5	12.7	12.0	11.5	7.1
24.	Swift	+19.5	5.7	7.5	4.8	2.9
25.	Texaco	-1.4	15.6	13.4	13.2	9.6
26.	Union Carbide	+6.0	11.3	9.5	8.7	5.2
27.	United Aircraft	-13.0	11.1	7.5	9.0	11.0
28.	U.S. Steel	-7.3	6.2	4.9	5.1	3.1
29.	Westinghouse Elec.	+37.5	11.4	10.4	9.3	8.9
30.	Woolworth	+23.6	9.7	8.8	8.1	11.1
	Sum	122.7	365.7	310.9	308.4	202.4
	$(Sum)^2$	15055	133736	96659	95111	40966

Note:

Variables X_1 through X_{16} and variable Y are defined in Appendix A.

APPENDIX B -- Continued

	X_5	X_6	X_7	X_8	X_9	X_{10}
1.	-12.2	-61.1	5.2	14.2	-5.3	4
2.	7.3	-28.2	3.5	11.3	-12.7	6
3.	6.7	15.9	5.0	10.7	-0.3	6
4.	1.2	-26.3	5.4	11.1	-4.4	8
5.	2.9	-22.6	4.9	13.4	-1.6	10
6.	0.5	-48.8	4.8	4.9	-5.6	4
7.	-4.8	-48.3	7.0	12.4	+0.4	6
8.	-	-44.7	2.2	12.3	+2.3	4
9.	-2.9	-47.8	3.8	21.0	-0.5	8
10.	9.4	20.8	2.0	27.9	-1.1	10
11.	-11.4	-24.9	3.0	19.9	-4.6	8
12.	4.7	0.3	3.1	17.5	-1.5	8
13.	-7.9	-25.2	6.1	21.8	+8.3	8
14.	3.9	28.4	2.7	16.2	+1.7	8
15.	-8.6	-40.3	6.2	14.9	+2.9	8
16.	6.2	28.1	3.6	14.5	-4.3	8
17.	1.4	10.2	4.1	18.5	+1.5	6
18.	-0.4	35.0	2.8	19.1	+4.6	6
19.	3.9	-22.2	2.4	14.8	-3.2	8
20.	11.3	70.5	2.5	20.9	+1.4	10
21.	7.1	12.3	2.1	25.6	+3.6	10
22.	2.5	-29.6	5.5	9.6	-2.4	8
23.	3.0	-10.6	5.4	11.4	-2.6	8
24.	4.0	7.3	2.5	12.6	-2.4	4
25.	4.2	-14.5	4.8	10.4	-3.1	8
26.	-5.6	-47.3	4.5	16.7	-3.3	8
27.	-2.3	-61.0	5.1	10.2	-3.3	6
28.	-5.9	-50.6	8.3	10.3	-2.7	6
29.	5.8	6.6	2.3	27.0	+7.5	6
30.	-0.2	13.4	3.2	15.4	+1.4	6
Sum	23.8	-405.2	124.0	466.5	-29.3	214
$(Sum)^2$	566.4	164187	15376	217622	858.5	45796

APPENDIX B -- Continued

	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}
1.	6	150	-5.0	+4.0	-36.7	+19.4
2.	6	170	-1.5	+5.8	-6.8	-1.0
3.	8	45	+0.4	+5.6	+9.3	+9.7
4.	6	80	+6.6	+6.1	+2.0	+4.2
5.	6	55	+8.5	+9.7	-0.8	+5.8
6.	4	175	-31.0	-2.5	-31.7	-3.2
7.	2	145	+0.2	+10.7	-42.4	+22.0
8.	6	130	-0.4	+10.3	-	-
9.	6	55	-0.1	+6.8	-8.0	+12.4
10.	6	45	+2.0	+5.4	+2.8	+7.8
11.	8	75	+3.3	+3.9	+17.8	+33.6
12.	6	50	+10.5	+7.1	+7.6	+6.2
13.	4	60	-17.7	+45.0	-64.6	+233.3
14.	2	40	-1.2	+10.2	-22.4	+32.4
15.	4	75	+2.2	+5.4	-16.5	+27.6
16.	8	45	+56.4	+9.3	+78.3	+12.5
17.	2	65	+3.6	+4.8	-28.8	+8.1
18.	4	0	-1.0	+9.9	-18.5	+23.8
19.	10	50	+8.3	+7.0	+4.2	+10.0
20.	6	10	+10.0	+9.1	+13.0	+13.5
21.	8	40	+4.1	+8.9	+4.9	+11.3
22.	8	70	+7.2	+7.3	-0.9	+5.7
23.	6	40	+6.8	+6.6	-0.5	+6.1
24.	6	40	-1.0	-0.8	+34.9	+13.8
25.	10	85	+7.4	+4.8	+6.7	+6.0
26.	6	75	+3.2	+5.8	-14.3	+9.8
27.	8	75	+4.2	-14.3	-10.9	-6.7
28.	4	115	+1.2	+10.6	-32.2	+28.7
29.	4	30	+22.9	+9.9	-18.4	+30.3
30.	2	20	+11.2	+10.8	+8.6	+9.1
Sum	172	2110	121.3	223.2	-164.3	592.2
$(Sum)^2$	29584	4452100	14714	49818	26994	350701

APPENDIX C

Summary Computations Used For Coefficient of Correlation (r) - Dow-Jones 30 Industrials

	EX	$(EX)^2$	EX^2	EXY
X_1	365.7	133736	4955.5	1794.1
X_2	310.9	96659	3767.5	1267.6
X_3	308.4	95111	3678.6	1690.4
X_4	202.4	40966	1676.2	816.2
X_5	23.8	566.4	1068	-227.4
X_6	-405.2	164187	36826.2	2746.6
X_7	124.0	15376	587.3	370.3
X_8	466.5	217622	8129.7	3834.8
X_9	-29.3	858.5	535.8	961.4
X_{10}	214	45796	1620	1023.6
X_{11}	172	29584	1128	478
X_{12}	2110	4452100	206950	1360.5
X_{13}	121.3	14714	3152.5	347.6
X_{14}	223.2	49818	3842.6	1386.4
X_{15}	-164.3	26994	20230.4	-1932.0
X_{16}	592.2	350701	62099	4418

Note:

$EY = 122.7$
$(EY)^2 = 15055$
$EY^2 = 11457.2$

APPENDIX D

Data Tabulations: 30 Firms Randomly Selected

		Y	X_1	X_2
1.	Distillers Corp-Seag.	+38.8	8.7	8.8
2.	Shell Oil	-0.3	12.8	9.2
3.	First Natl Stores	-37.1	1.6	3.7
4.	SS Kresge	+72.4	15.2	15.9
5.	Ethyl	+5.9	20.3	15.4
6.	McDonnell Douglas	+63.7	13.7	14.1
7.	Carborundum	+20.5	10.4	9.5
8.	Internatl Utilities	+20.2	10.3	10.4
9.	Dart Industries	+29.3	13.6	12.5
10.	Thrifty Drug	+7.0	14.7	13.5
11.	Kimberly-Clark	-8.5	9.5	8.2
12.	Deere & Co.	+25.0	9.3	7.1
13.	Coastal States Gas	+8.5	22.7	18.8
14.	Continental Can	-17.4	13.6	12.8
15.	Associated Dry Goods	+38.9	12.6	9.9
16.	Hercules	+22.6	14.5	12.0
17.	American Express	+45.3	11.7	13.4
18.	Genl American Trans	+23.6	11.9	11.7
19.	Fluor	-10.7	23.1	13.6
20.	RJ Reynolds Inds.	+8.2	18.8	21.6
21.	Southern RR.	+37.0	6.1	7.3
22.	Anchor Hocking	-5.7	16.0	15.5
23.	BF Goodrich	+2.3	7.9	4.2
24.	Crane	+34.6	7.3	6.6
25.	Bethlehem Steel	+26.8	8.1	5.9
26.	Arden-Mayfair	0	6.0	2.3
27.	Brown Shoe	+14.1	14.6	10.5
28.	Sperry Rand	+20.0	12.6	11.2
29.	Outboard Marine	+68.8	13.3	9.9
30.	Missouri Pacific RR	+15.4	5.0	4.4
	Sum	569.2	365.9	319.9
	$(\text{Sum})^2$	323989	133883	102336

Note:

Variables X_1 through X_{16} and variable Y are defined in APPENDIX A.

APPENDIX D -- Continued

	X_3	X_4	X_5	X_6	X_7
1.	7.0	7.5	6.6	38.9	2.4
2.	10.5	6.7	-0.8	-27.6	5.0
3.	1.6	3.7	8.3	7.2	3.1
4.	12.9	22.7	20.0	324.7	0.8
5.	9.8	11.3	3.2	-44.9	3.5
6.	10.3	4.8	16.4	-23.7	1.7
7.	8.9	8.5	2.0	13.1	3.1
8.	6.5	16.7	7.8	22.5	3.9
9.	11.0	11.8	3.6	-19.2	0.9
10.	12.4	7.7	3.3	-13.9	3.3
11.	7.9	6.8	0.9	11.9	3.9
12.	7.6	3.2	-6.2	-25.2	5.0
13.	9.0	21.4	11.8	190.2	0
14.	10.3	8.9	8.5	29.0	4.1
15.	10.0	7.9	1.5	3.6	2.9
16.	11.6	7.5	2.0	-2.7	3.0
17.	11.7	10.1	18.1	235.5	1.5
18.	6.0	-1.0	2.0	-12.4	3.7
19.	17.1	21.0	13.2	50.2	0
20.	16.7	7.3	5.9	16.2	4.5
21.	4.9	5.5	9.0	-0.2	5.2
22.	13.8	10.4	12.4	54.7	3.1
23.	6.7	5.0	-10.3	-31.0	3.6
24.	6.0	12.4	4.3	0.2	4.7
25.	7.2	4.0	-4.8	-48.3	7.0
26.	3.9	4.5	-30.7	-67.4	0
27.	13.6	8.4	-0.1	0.9	4.2
28.	9.7	5.9	15.5	13.6	2.4
29.	11.5	10.2	2.9	23.9	4.0
30.	2.9	5.1	-4.0	-32.1	7.7
Sum	279	265.9	122.3	687.7	98.2
$(Sum)^2$	77841	70703	14957	472931	9643

APPENDIX D -- Continued

	X_8	X_9	X_{10}	X_{11}	X_{12}
1.	14.2	+2.7	8	6	15
2.	12.6	-2.4	8	6	80
3.	15.1	+3.1	4	8	65
4.	32.2	+17.7	8	8	-5
5.	7.7	-2.8	8	10	50
6.	11.7	+2.0	6	4	10
7.	13.5	+2.0	4	6	45
8.	11.7	-5.3	8	10	30
9.	17.7	-0.3	6	4	65
10.	12.0	-1.5	6	8	50
11.	16.4	-0.1	6	2	55
12.	10.4	-0.6	8	6	105
13.	23.9	+1.2	8	6	20
14.	12.8	+0.3	8	4	40
15.	16.8	+1.1	8	6	45
16.	17.0	-4.5	8	8	75
17.	25.2	+1.7	8	8	55
18.	15.4	-1.1	6	6	40
19.	12.3	+2.8	4	4	125
20.	11.2	-0.3	8	8	40
21.	8.1	-1.8	6	10	25
22.	11.5	-2.0	8	6	55
23.	17.8	+1.8	4	4	85
24.	10.0	-2.5	6	4	65
25.	12.4	+0.4	6	2	145
26.	27.9	+15.4	2	4	85
27.	13.7	+1.5	8	6	25
28.	10.2	-10.3	4	8	175
29.	14.6	-1.4	2	8	95
30.	5.2	-1.3	4	8	35
Sum	441.2	15.5	188	188	1795
$(Sum)^2$	194657	240.2	35344	35344	3222025

APPENDIX D -- Continued

	X_{13}	X_{14}	X_{15}	X_{16}
1.	+7.0	+4.4	+6.2	+3.7
2.	+1.5	+8.6	-19.0	+11.4
3.	+10.2	+8.8	+31.7	+24.5
4.	+16.7	+13.7	+11.8	+14.7
5.	+9.4	+8.6	+8.9	+4.6
6.	-30.6	-4.8	-30.7	-39.3
7.	+4.1	+9.3	-4.2	+8.8
8.	+7.8	+7.6	+20.5	+38.3
9.	-0.7	+9.5	-13.6	+14.1
10.	+8.2	+7.9	+8.8	+9.4
11.	+3.7	+8.5	-25.8	+18.9
12.	+9.1	+5.4	-15.2	+20.2
13.	+11.1	+18.8	+13.6	+22.8
14.	+9.6	+7.7	-5.8	0
15.	+2.0	+4.4	-8.9	+11.1
16.	+7.1	+9.5	+12.6	+13.5
17.	+14.0	+10.7	+24.0	+10.0
18.	+7.7	+4.5	+13.8	+3.6
19.	+16.6	+9.2	-13.6	+24.2
20.	+10.1	+4.8	+20.4	+8.7
21.	+5.4	+1.7	+24.2	+3.2
22.	+5.2	+12.3	-11.0	+12.7
23.	+0.1	+11.8	-59.8	+109.1
24.	+24.3	+2.9	-14.4	0
25.	+0.2	+10.7	-42.4	+22.0
26.	+8.0	+4.7	-45.9	+75.0
27.	+8.9	+1.2	+4.2	+12.0
28.	+0.8	+6.5	-15.6	+17.5
29.	-7.0	+18.4	-21.3	+29.5
30.	+9.7	+9.6	+0.7	+11.3
Sum	180.2	236.9	-145.8	515.5
$(Sum)^2$	32472	56122	21258	265740

APPENDIX E

Summary Computations Used For Coefficient of Correlation (r) - 30 Firms Randomly Selected

	EX	$(EX)^2$	EX^2	EXY
X_1	365.9	133883	5184.1	7005.9
X_2	319.9	102336	3994.3	6521.9
X_3	279	77841	2994.7	5616.9
X_4	265.9	70703	3229	5582.7
X_5	122.3	14957	3268.2	4160
X_6	687.7	472931	220401	33235
X_7	98.2	9643	422.7	1782.8
X_8	441.2	194657	7509	9263.8
X_9	15.5	240.2	783.6	820
X_{10}	188	35344	1288	3672
X_{11}	188	35344	1320	3836
X_{12}	1795	3222025	153525	27872
X_{13}	180.2	32472	3537	1395
X_{14}	236.9	56122	2531.4	4230.3
X_{15}	-145.8	21258	15096	-2788.8
X_{16}	515.5	265740	26353.5	4487

Note:

$EY = 569.2$
$(EY)^2 = 323989$
$EY^2 = 28730.6$

APPENDIX F

Summary Computations Used For Coefficient of Correlation (r)
Combined Observations For Dow-Jones 30 Industrials
and 30 Randomly Selected Stocks

	EX	$(EX)^2$	EX^2	EXY
X_5	146.1	21345.2	4336.2	3932.6
X_6	282.5	79806	257227	35982
X_8	907.7	823919	15639	13099
X_9	-13.8	190.4	1319.4	1781.4
X_{12}	3905	15249025	360475	29232.5
X_{13}	301.5	90902	6689.5	1742.6
X_{16}	1107.7	1226999	88452.5	8905

Note:

EY = 692

$(EY)^2$ = 478726

EY^2 = 40188

APPENDIX G

Data Tabulations: 50 Firms Randomly Selected

		X_5	X_6	X_8	X_9	X_{12}
1.	Anchor Hocking	12.4	54.7	11.5	-2.0	55
2.	Foremost-McKesson	2.8	-25.1	13.5	-1.0	40
3.	Archer-Daniels-Midland	12.9	71.0	9.6	-4.4	55
4.	Rohm & Haas	-5.4	-37.2	21.5	-1.5	75
5.	Amerada Hess	9.7	193.8	14.5	+2.5	65
6.	Gen Tel & Elec	2.0	-38.3	14.2	-5.3	85
7.	Crown Zellerbach	-1.8	-11.1	15.6	-1.9	50
8.	Chesebrough-Pond's	10.4	64.9	25.6	+3.8	5
9.	Broadway-Hale	7.1	21.7	15.7	+1.2	70
10.	Crowell-Collier	-2.4	-34.2	17.1	+1.1	165
11.	Kendall	2.4	36.6	22.0	+8.5	-10
12.	Lykes-Youngstown	-58.3	-45.5	22.0	+13.4	175
13.	Pennwalt	-10.4	-52.1	21.4	+0.4	90
14.	Kellogg	6.3	-2.1	16.1	-2.4	70
15.	Singer	1.6	-6.0	19.3	+3.8	25
16.	Std Oil (Ind)	7.0	9.8	12.8	+0.3	15
17.	HJ Heinz	10.3	66.3	15.1	+2.6	-5
18.	Emerson Elect	12.8	133.8	29.7	+10.7	5
19.	Castle & Cooke	7.9	51.4	12.5	-1.0	100
20.	Borman's	-1.3	-36.3	10.2	-3.3	175
21.	Universal Leaf Tobac	2.2	29.5	10.0	+0.5	65
22.	Pepsico	9.2	28.8	19.5	0	60
23.	Clark Oil & Refining	9.1	116.0	12.1	+0.1	70
24.	Rohr	11.8	-36.1	12.6	+2.6	130
25.	Natl Distil & Chem.	0.5	-15.9	11.5	-2.6	95

Note:

See Appendix A for Definitions of X Variables.

APPENDIX G -- Continued

		X_5	X_6	X_8	X_9	X_{12}
26.	Allis-Chalmers	-7.7	-56.7	25.8	+10.3	170
27.	Union Electric	6.3	-27.6	12.4	-6.1	70
28.	American Standard	-1.6	39.2	20.0	+7.5	50
29.	Corning Glass	-0.7	-25.0	39.6	+2.6	65
30.	Miles Laboratories	7.4	2.2	20.4	+1.4	90
31.	Leasway Transport	8.4	-0.4	20.6	+5.0	75
32.	General Cigar	13.8	-3.9	9.5	-5.0	100
33.	Smith Kline & French	1.0	-37.7	18.7	-3.3	55
34.	Sybron	4.8	43.5	18.1	+3.6	60
35.	Uniroyal	-2.4	2.1	22.7	+10.7	40
36.	Marshall Field	0.2	-5.2	15.0	+2.0	45
37.	Otis Elevator	-1.7	-30.2	13.5	-2.0	35
38.	United States Gypsum	-11.3	-10.7	23.3	+5.8	50
39.	Gen Tire & Rubber	-1.2	-26.4	11.3	-0.7	60
40.	National Steel	-3.6	-34.5	11.0	+0.5	65
41.	Federal-Mogul	-6.9	-37.5	12.6	+0.1	80
42.	Consolidated Freight	-2.4	27.7	15.8	+6.3	45
43.	Dana	4.3	5.4	12.4	+1.9	20
44.	Metromedia	-7.0	-16.9	15.9	-0.1	100
45.	US Freight	-3.6	-17.2	16.1	-1.9	90
46.	Ara Services	10.7	127.8	32.8	+2.3	0
47.	Niagara Mohawk Power	-0.1	-40.4	11.9	-4.6	80
48.	Allied Stores	-7.6	-28.0	11.3	-0.5	95
49.	EF MacDonald	-7.8	-48.1	9.8	-6.7	290
50.	Delta Airlines	13.0	31.1	15.2	+4.7	45

APPENDIX H

Comparison of Computed Y Values Versus Actual Y Values for 50 Randomly Selected Firms

	Actual Y Values	Computed Y Values				
		X_5	X_6	X_8	X_9	X_{12}
1.	-5.7	+17.2	+17.9	+6.5	+8.9	+13.0
2.	+14.5	+11.9	+7.7	+9.3	+10.4	+15.3
3.	+13.1	+17.5	+20.0	+3.9	+5.4	+13.0
4.	+25.9	+7.3	+6.2	+20.3	+9.7	+10.1
5.	-11.2	+15.7	+35.7	+10.7	+15.6	+11.5
6.	+2.9	+11.4	+6.0	+10.2	+4.1	+8.6
7.	+10.1	+9.3	+9.5	+12.2	+9.1	+13.8
8.	+28.6	+16.1	+19.2	+26.0	+17.5	+20.5
9.	+25.7	+14.3	+13.7	+12.3	+13.6	+10.8
10.	-9.6	+9.0	+6.6	+14.2	+13.5	-3.3
11.	+55.8	+11.6	+15.6	+21.0	+24.4	+22.7
12.	+23.2	-22.3	+5.1	+21.0	+31.6	-4.8
13.	+3.4	+4.5	+4.3	+20.2	+12.5	+7.8
14.	+12.5	+13.8	+10.7	+12.9	+8.3	+10.8
15.	+18.8	+11.2	+10.2	+17.3	+17.5	+17.5
16.	+32.1	+14.2	+12.2	+8.3	+12.3	+19.0
17.	+20.1	+16.1	+19.4	+11.5	+15.7	+21.9
18.	+21.6	+17.5	+28.0	+31.6	+27.6	+20.5
19.	-28.1	+14.7	+17.5	+7.9	+10.4	+6.3
20.	-26.7	+9.6	+6.3	+4.7	+7.0	-4.8
21.	-7.0	+11.5	+14.7	+4.5	+12.6	+11.5
22.	+32.5	+15.4	+14.6	+17.6	+11.9	+12.3
23.	-19.0	+15.4	+25.8	+7.4	+12.0	+10.8
24.	+5.6	+16.9	+6.3	+8.0	+15.7	+1.9
25.	-0.8	+10.6	+8.9	+6.5	+8.0	+7.1

Note:
See Appendix A for definitions of X and Y variables.

APPENDIX H -- Continued

	Actual Y Values	Computed Y Values X_5	X_6	X_8	X_9	X_{12}
26.	-18.8	+6.0	+3.7	+26.2	+27.0	-4.1
27.	-11.4	+13.8	+7.4	+7.8	+2.9	+10.8
28.	-49.6	+9.4	+15.9	+18.2	+22.9	+13.8
29.	+5.0	+9.9	+7.7	+45.3	+15.7	+11.5
30.	+1.4	+14.4	+11.2	+18.8	+13.9	+7.8
31.	+136.7	+15.0	+10.9	+19.1	+19.2	+10.1
32.	-23.6	+18.0	+10.4	+3.8	+4.5	+6.3
33.	+12.3	+10.9	+6.1	+16.5	+7.0	+13.0
34.	+22.0	+13.0	+16.5	+15.6	+17.2	+12.3
35.	-15.6	+9.0	+11.2	+22.0	+27.6	+15.3
36.	+25.6	+10.4	+10.3	+11.4	+14.8	+14.5
37.	-7.8	+9.4	+7.1	+9.3	+8.9	+16.0
38.	+2.4	+4.0	+9.6	+22.8	+20.4	+13.8
39.	+23.1	+9.6	+7.6	+6.2	+10.8	+12.3
40.	+4.4	+8.3	+6.5	+5.8	+12.6	+11.5
41.	+0.5	+6.4	+6.1	+8.0	+12.0	+9.3
42.	+78.5	+9.0	+14.5	+12.5	+21.2	+14.5
43.	+20.1	+12.7	+11.6	+7.8	+14.7	+18.2
44.	+55.0	+6.4	+8.8	+12.6	+11.7	+6.3
45.	+26.5	+8.3	+8.7	+12.9	+9.1	+7.8
46.	+30.4	+16.3	+27.2	+35.9	+15.3	+21.2
47.	+3.1	+10.2	+5.8	+7.1	+5.1	+9.3
48.	+24.1	+6.0	+7.4	+6.2	+11.1	+7.1
49.	+20.0	+5.9	+4.8	+4.2	+2.0	-21.9
50.	+45.4	+17.6	+14.9	+11.6	+18.8	+14.5

APPENDIX I

Data Tabulations for 10 Stocks which Increased Significantly In Market Value in 1971 and 10 Stocks which Decreased Significantly in Market Value in 1971.

	X_5	X_6	X_8	X_9	X_{12}
Stock Price Increase					
1. Maremont	-16.5	-61.4	9.9	+2.4	100
2. Leaseway	8.4	-0.4	20.6	+5.8	75
3. Amer Broadcasting	3.0	-50.4	11.2	-6.3	170
4. Genuine Parts	11.0	130.9	19.5	+6.0	-5
5. Thiokol Chemical	-0.6	-58.2	11.8	-7.7	140
6. Howard Johnson	3.6	-3.3	18.8	-4.2	45
7. Northwest Airlines	2.3	-39.7	45.3	+34.8	180
8. Cont. Freight	-2.4	27.7	15.8	+6.3	45
9. Jefferson-Pilot	4.9	-47.3	9.2	-3.3	80
10. Bell & Howell	3.9	-21.0	16.9	-8.6	120
Stock Price Decrease					
1. GAC Corp	7.7	0.6	12.2	+0.2	130
2. Kais Aluminum	3.7	-17.1	12.5	-7.5	115
3. Midland Ross	-10.6	-4.9	13.4	+2.4	35
4. Kennecott	15.8	-14.1	6.1	-6.9	190
5. First Natl. Stores	8.3	7.2	15.1	+3.1	65
6. Revere Copper	-9.2	-24.0	10.3	-2.2	130
7. National Can	14.8	76.6	13.1	+1.6	35
8. Scott Paper	0.9	-38.9	14.4	-9.6	110
9. Reynolds Metals	1.1	-50.9	8.8	-8.7	185
10. Cook United	11.8	93.1	10.6	-1.9	75

APPENDIX J

Comparison of Computed Y Values Versus Actual Y Values for 10 Stocks which Increased Significantly in Market Value in 1971 and 10 Stocks which Decreased Significantly in Market Value in 1971

	Actual Y Values	Computed Y Values X_5	X_6	X_8	X_9	X_{12}
Increase						
1.	+212.2	+1.1	+3.1	+4.3	+15.4	+6.3
2.	+136.7	+15.0	+10.9	+19.1	+20.4	+10.1
3.	+103.4	+12.0	+4.5	+6.1	+2.6	-4.1
4.	+94.5	+16.5	+27.7	+17.6	+20.7	+21.9
5.	+85.9	+10.0	+3.5	+6.9	+0.5	+0.4
6.	+85.6	+12.3	+10.5	+16.6	+5.7	+14.5
7.	+81.8	+11.6	+5.8	+53.2	+63.2	-5.6
8.	+78.5	+9.0	+14.5	+12.5	+21.2	+14.5
9.	+76.1	+13.0	+4.9	+3.4	+7.0	+9.3
10.	+75.5	+12.5	+8.2	+14.0	-0.8	+3.4
Decrease						
1.	-51.6	+14.6	+11.0	+7.5	+12.2	+1.9
2.	-42.5	+12.4	+8.7	+7.9	+0.8	+4.1
3.	-39.0	+4.4	+10.3	+9.2	+15.4	+16.0
4.	-38.1	+19.1	+9.1	-0.9	+1.7	-7.0
5.	-37.1	+14.9	+11.8	+11.5	+16.4	+11.5
6.	-36.4	+5.2	+7.9	+4.9	+8.6	+1.9
7.	-35.1	+18.6	+20.7	+8.7	+14.2	+16.0
8.	-33.3	+10.8	+6.0	+10.5	-2.3	+4.8
9.	-32.9	+10.9	+4.4	+2.8	-1.0	-6.3
10.	-32.5	+16.9	+22.8	+5.3	+9.1	+10.1

APPENDIX K

Computed Y Values for 50 Randomly Selected Firms Using Multiple Regression Analysis with Independent Variables X_6 and X_8

1.	13.0	26.	15.0
2.	7.4	27.	6.3
3.	12.9	28.	18.8
4.	13.1	29.	29.5
5.	28.2	30.	15.8
6.	6.8	31.	15.7
7.	10.5	32.	6.0
8.	25.8	33.	10.7
9.	13.6	34.	17.6
10.	9.7	35.	17.7
11.	23.6	36.	10.5
12.	12.8	37.	7.0
13.	11.7	38.	17.0
14.	11.7	39.	5.5
15.	14.1	40.	4.5
16.	10.0	41.	5.6
17.	17.1	42.	14.2
18.	35.5	43.	9.3
19.	18.2	44.	10.2
20.	3.7	45.	10.4
21.	9.5	46.	37.6
22.	17.4	47.	4.7
23.	19.1	48.	5.3
24.	5.7	49.	2.2
25.	6.6	50.	14.0

Note:

The names of the 50 companies listed above are found in Appendix G.

BIBLIOGRAPHY

BOOKS

Amling, Frederick. Investments, An Introduction to Analysis and Management. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1965.

Bellmore, Douglas A., and Ritchie, John C., Jr. Investments-Principles, Practices, Analysis. Cincinnati, Ohio: South-Western Publishing Company, 1969.

Bonini, Charles P., and Spurr, William A. Statistical Analysis for Business Decisions. Homewood, Ill.: Richard D. Irwin, Inc., 1967.

Brabb, George J., and Hanson, Kermit O. Managerial Statistics. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1961.

Brealey, Richard A. An Introduction to Risk and Return from Common Stocks. Cambridge, Mass.: The MIT Press, 1969.

Brealey, Richard A. Security Prices In A Competitive Market: More About Risk and Return from Common Stocks. Cambridge, Mass.: The MIT Press, 1971.

Cochran, William G., and Snedecor, George W. Statistical Methods. Ames, Iowa: The Iowa State University Press, 1967.

Cohen, Jerome B., and Zinbarg, Edward D. Investment Analysis And Portfolio Management. Homewood, Ill.: Richard D. Irwin, Inc., 1967.

Cootner, Paul H. The Random Character of Stock Market Prices. Cambridge, Mass.: The MIT Press, 1964.

Cottle, Sidney; Dodd, David L.; and Graham, Benjamin. Security Analysis, Principles and Technique. New York: McGraw-Hill Book Company, 1962.

Crary, David T., and Lishan, John M. The Investment Process. Scranton, Pa.: International Textbook Company, 1970.

Dixon, Wilfrid J., and Massey, Frank J., Jr. Introduction to Statistical Analysis. New York: McGraw-Hill Book Company, 1969.

Ezekiel, Mordecai, and Fox, Karl A. Methods of Correlation and Regression Analysis. New York: John Wiley & Sons, Inc., 1959.

Farrell, Maurice L. The Dow Jones Investor's Handbook. Princeton, N.J.: Dow Jones & Company, Inc., 1972.

Fredrikson, E. Bruce. Frontiers of Investment Analysis. Scranton, Pa.: International Textbook Company, 1971.

Freund, John E., and Williams, Frank J. Modern Business Statistics. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1958.

Gillett, Clarence H. Stock Market Timing, A New Approach. Larchmont, N.Y.: Investors Intelligence, 1965.

Gordon, William. The Stock Market Indicators (As A Guide To Market Timing). Palisades Park, N.J.: Investors' Press, Inc., 1968.

Greiner, Perry P., and Whitcomb, Hale C. The Dow Theory and the Seventy-Year Forecast Record. Larchmont, N.Y.: Investors Intelligence, Inc., 1969.

Jureen, Lars, and Wold, Herman. Demand Analysis. New York: John Wiley & Sons, Inc., 1953.

Krow, Harvey A. Stock Market Behavior, The Technical Approach to Understanding Wall Street. New York: Random House, Inc., 1969.

Loeb, Gerald M. The Battle For Investment Survival. New York: Simon and Schuster, 1965.

Loeb, Gerald M. The Battle For Stock Market Profits. New York: Simon and Schuster, 1971.

Markowitz, Harry M. Portfolio Selection, Efficient Diversification of Investments. New Haven, Conn.: Yale University Press, 1959.

Neter, John, and Wasserman, William. Fundamental Statistics for Business and Economics. Boston: Allyn and Bacon, Inc., 1956.

Roscoe, John T. Fundamental Research Statistics for the Behavioral Sciences. New York: Holt, Rinehart and Winston, Inc., 1969.

Rosenberg, Claude N., Jr. Psycho-Cybernetics and the Stock Market. Chicago, Ill.: Playboy Press, 1971.

Sauvain, Harry. Investment Management. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1967.

Selby, Samuel M. Standard Mathematical Tables. Cleveland, Ohio: The Chemical Rubber Co., 1968.

Shade, Philip A. Common Stocks, A Plan for Intelligent Investing. Homewood, Ill.: Richard D. Irwin, Inc., 1971.

Siegelman, Louis, and Spencer, Milton H. Managerial Economics, Decision Making and Forward Planning. Homewood, Ill.: Richard D. Irwin, Inc., 1964.

Sprowls, R. Clay. Elementary Statistics. New York: McGraw-Hill Book Company, 1955.

Vaughn, Donald E. Survey of Investments. New York: Holt, Rinehart and Winston, Inc., 1967.

Williamson, J. Peter. Investments, New Analytic Techniques. New York: Praeger Publishers, 1971.

PERIODICALS

Forbes, "23rd Annual Report on American Industry," January 1, 1971.

Forbes, "24th Annual Report on American Industry," January 1, 1972.

Las Vegas Review-Journal. Annual Financial Tabulations, January 2, 1972.

Moody's Industrial Manual. New York: Moody's Investors Service, Inc., 1971.

Standard & Poor's Corporation Records. New York: Standard & Poor's Corporation, 1971.

Standard Stock Reports. New York: Standard & Poor's Corporation, 1971.

The Value Line Investment Survey. New York: Arnold Bernhard & Co., Inc., 1971.