Rates of Return on Investments in Common Stocks

The 50th Year Anniversary Issue

1964 - 2014
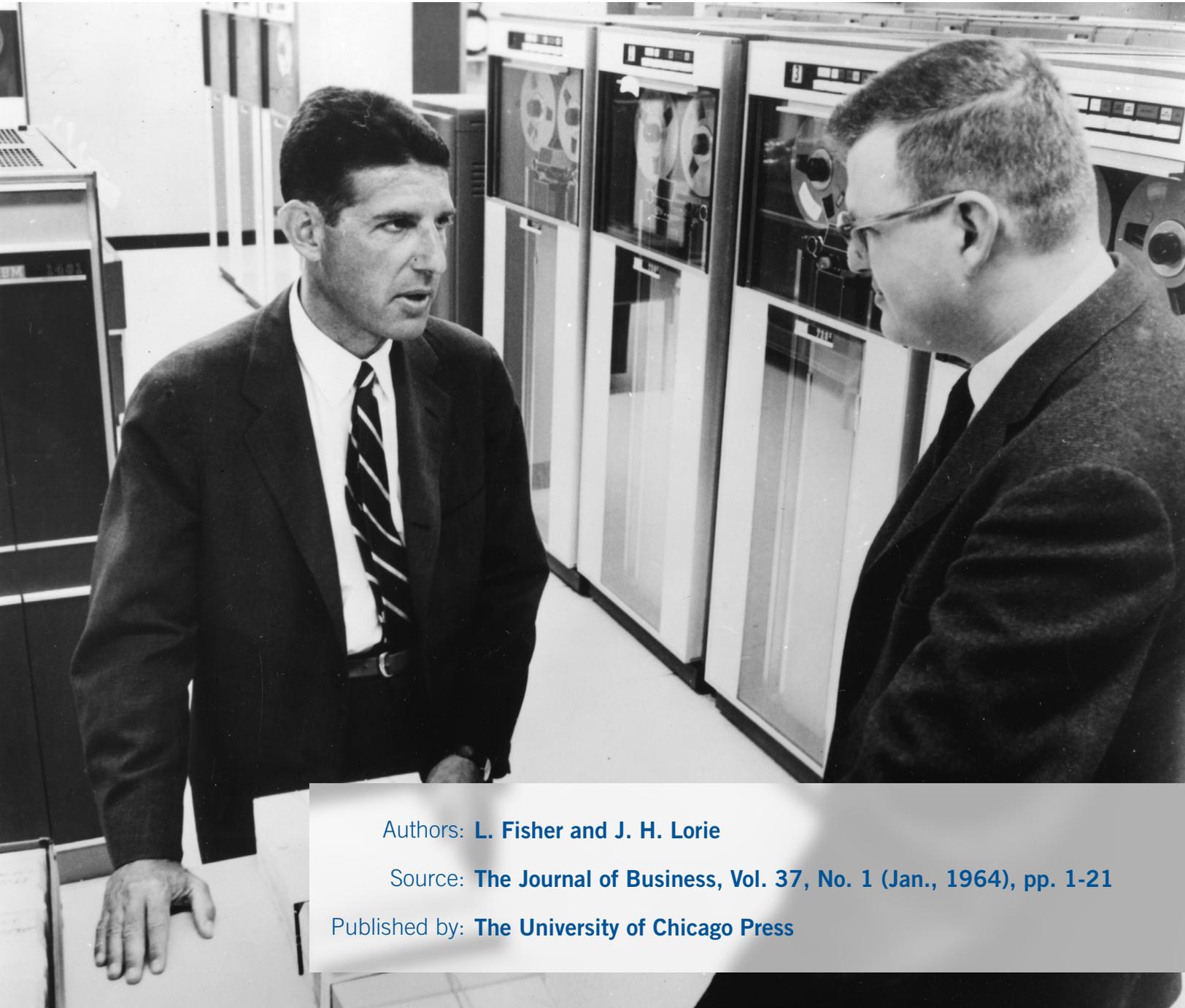
On December 2, 1963, at the Press Club in New York, Professors Fisher and Lorie announced the results of their 3-1/2 year data development and analysis project. The groundbreaking results presented in the study made headlines around the country.

THE WALL STREET JOURNAL.

December 2, 1963 – James H. Lorie, professor of business administration and director of the center, said the accumulation of this vast amount of detail, totaling between 2 and 3 million separate entries on the tapes, was essential not only for the center's initial project of measuring the return on stock investment but also for other planned future studies. Among them: A look at whether the price changes on the stock market follow any predictable pattern and a study of the relationship between changes and prices of stock and changes and other economic indicators, such as interest rates.

The New York Times

December 2, 1963 – The pioneering three-and-one-half-year research project, sponsored primarily by Merrill Lynch Pierce, Fenner & Smith, Inc, was designed to answer the seemingly simply question posed by Louis Engel, a vice president of the financial house: "Can you tell me what the experience has been of the average investor over the long haul in common stock? Has he made or lost money on his investments?"

Chicago Tribune

December 2, 1963 – All of the data compiled in the course of the study is on magnetic tape, which if unreeled, would stretch for 3-1/2 miles. Lorie and Fisher, whom Lorie credits with most of the work…, ran into immense unforeseen difficulties. It developed, for example, that there are 39 different kinds of capital adjustments that can affect investment values. There are seven different possible tax treatments for dividends. There are 50 different kinds of security, said Lorie, which may be described as common stock.

Los Angeles Times

December 2, 1963 – "The task was staggering," Dr. Lorie said. "Here was a sand dune and we were asked to trace the history of every grain of every shift and change over 35 years. It took 3-1/2 years to assemble the data. It took the computer 20 minutes to make something like 300-400 calculations and come up with answers."

Table of Contents

January 2014

In 1960, when Jim Lorie and Larry Fisher, then both faculty members at Chicago Booth, partnered with Louis Engel of Merrill Lynch to develop CRSP, it is hard to imagine that even they could have foreseen the enduring legacy of their work. It is not an overstatement to suggest that academics, individual investors, and investment practitioners have all been profoundly impacted by the development of CRSP.

With a pioneering spirit, CRSP and Chicago Booth applied the earliest available technology in creating machine readable databases. In 1964, with the publication of their seminal research paper entitled, Rates of Return, Fisher and Lorie for the first time in the history of finance established the long-term "total market return." With more than half a century of meticulous and continuous data curation, led by the efforts of Fisher and Lorie, CRSP serves multiple generations of scholarly researchers with research-quality data.

The results of their research and the data underpinning the analysis continue to serve as the foundations for quantitative advancements in finance, economics and related disciplines. Today, nearly 450 leading academic institutions rely on CRSP data to underpin research and teaching. CRSP is also widely used in the commercial market for backtesting and performance benchmarking. Moreover, our recent development of the investable CRSP Indexes demonstrates CRSP's commitment to preserving the legacy of Fisher and Lorie and building on their achievements.

We present this 50th year commemorative edition of the Rates of Return paper and the stories of Fisher, Lorie and Engel in recognition of these industry luminaries.

David K. Barclay
Chief Operating Officer
CRSP | Center for Research in Security Prices
Chicago Booth

RATES OF RETURN ON INVESTMENTS IN COMMON STOCKS*

L. FISHER† AND J. H. LORIE‡

THIS article presents data on rates of return on investments in common stocks. It answers the question of how much gain or loss an individual investor might have realized if he had bought all New York Stock Exchange common stocks—at five different dates and held them for varying lengths of time during the thirty-five years from 1926 through 1960, a total of twenty-two time periods. This work is the first to emerge from the Center for Research in Security Prices (sponsored by Merrill Lynch, Pierce, Fenner & Smith Inc.). For that reason we shall describe the facilities and plans of the Center before presenting the results and methods of its first research.

THE CENTER FOR RESEARCH IN SECURITY PRICES

PURPOSE

The sole purpose of the Center for Research in Security Prices is to conduct research and to disseminate the results throughout the academic and financial communities. This is what Merrill Lynch, Pierce, Fenner & Smith Inc. had in mind when they provided the funds to establish the Center, and this is what the Graduate School of Business had in mind when it sought support.

Despite the enormous quantities of available data on security prices and related things, the amount of scientific research based on these data is inadequate for many purposes. As a result, a number of controversies and unresolved issues about security prices have continued. For example, there is controversy about whether successive changes in the prices of common stocks are statistically independent or are serially correlated. If the former is true, technical analysis such as that based on the Dow Theory is not likely to be helpful; if the latter is true, regularities in the patterns of change can be detected by appropriate analysis.

Another controversy revolves about the question of what effect, if any, dividends exert on stock prices. Some feel

1

that only earnings matter and that pay-out ratios are therefore of no consequence. Others feel that the declaration of a dollar dividend affects the price of a stock differently from a dollar of earnings retained. This matter is of great consequence to American corporations, but the issue remains unresolved.

Another difficult problem is the subject of stock-market averages. Averages are a lot like the weather in that everyone talks about them—usually with dissatisfaction—but no one does anything. It is time that a serious effort be made to devise averages which do more than report descriptively the movement of some stock prices.

Many people would like to have more detailed and refined information regarding the relationship of movements in stock prices to changes of other sorts in the economy. More work needs to be done on the relationships between earnings and prices and on the relevance of balance-sheet data in understanding and predicting the financial fate of corporations.

The Center for Research in Security Prices will work on all of the problems mentioned above and many others. It is intended to be a permanent institution which will work as rapidly as financial resources and the imagination and abilities of research workers will permit. The facilities of the Center will be available not only to the faculty and graduate students of the University of Chicago but also to persons at other educational institutions and in the financial community who have serious problems for which the Center's facilities can be helpful. The Center's facilities will be described below.

FACILITIES

In March, 1960, Merrill Lynch, Pierce, Fenner & Smith Inc., made an initial grant of $50,000 to the University of Chicago to establish the Center for Research in Security Prices. The initial funds and additional resources equivalent to more than $150,000 have been used to establish a "laboratory" consisting of a very large volume of data on security prices and related things appropriately arranged and stored on computer tape. These data together with appropriate computer programs, make possible a large variety and number of analyses of security prices, of their determinants, and of their relationships to other things in the economy.

Security prices.—Monthly closing prices of all common stocks on the New York Stock Exchange from January, 1926, through December, 1960, have been placed on tape. Their accuracy has been appraised in several ways, and, after three years of checking and rechecking, it is estimated that the incidence of error is extremely low and that remaining errors do not bias the results.

Further, there are on tape—though at present in less useful form—daily high, low, and closing prices of all common stocks on the New York Stock Exchange from July 1, 1960, to the present. This information is being kept up to date, but has not yet been integrated with the file on monthly prices; and its accuracy has not been tested adequately.

Capital changes.—In order for the information on security prices to be useful for refined analyses, it is essential that information be recorded on all types of capital changes. This has been done. Such changes include, for example, cash dividends, stock dividends, stock splits, rights, exchanges of shares, etc. One way to summarize the data that have been placed on tape is to say that almost all pertinent information in the Commerce Clearing House series, *Capital Changes Reporter*, has been coded, checked for possible error, and placed on tape. The

accuracy of these data has been exhaustively checked.

An additional comment is warranted on the treatment of cash dividends. Where available, four dates are recorded for each dividend: (1) the date of declaration, (2) the ex-dividend date, (3) the date of record, and (4) the date of payment. Further, the tax status for each dividend is recorded; there are seven different tax categories (see Appendix A, p. 10).

Other data.—A wide variety of national economic statistics has been coded and is ready to go on tape. These data cover such things as commodity prices, interest rates, industrial production, and national income.

For each period for which there is a security price, data on trading volume have also been recorded.

Data on earnings have been placed on IBM cards and ultimately will be transferred to tape. This information is in the form reported by the company and requires extensive adjustment before its use for refined analysis will be worthwhile.

PLANS FOR THE FUTURE

It is planned to keep the basic file of information up to date and to include new data of various kinds, depending on the projects that the Center undertakes. Undoubtedly there will be experimentation with balance-sheet data, and certainly an effort will be made to include data on securities other than common stocks listed on the New York Stock Exchange. Priority will be given to research of general interest, but other projects will also be undertaken.

RATES OF RETURN ON INVESTMENTS IN COMMON STOCKS

It is surprising to realize that there have been no measurements of the rates of return on investments in common stocks that could be considered accurate and definitive. There have been many efforts, but each has been deficient in at least one crucial respect. Some have lacked comprehensiveness, having dealt only with a selection of individual securities such as those in one of the popular stock market averages. This study embraces all common stock listed on the New York Stock Exchange during the time periods covered—some seventeen hundred of them. Again, earlier studies have dealt only with one or two brief time periods in contrast to the twenty-two time periods within a thirty-five-year span covered here. Finally, all other studies have been deficient because they have ignored taxes and transaction costs. In this study actual New York Stock Exchange round-lot commission rates, as they existed on all purchase, sales and re-investment dates, have been included in the calculations; and all federal taxes, as they applied to income from dividends or capital gains at specific times in selected tax brackets, have been taken into consideration. As can be seen from Tables 1 and 2, taxes and commissions can have large effects on rates of return. This study has shortcomings, but we believe that they are less pronounced than those of previous work.

RATES OF RETURN WITH REINVESTMENT OF DIVIDENDS

Table 1 shows the results of investing an equal sum of money in each company having one or more issues of common stock listed on the New York Stock Exchange at the beginning of each period *and* of reinvesting dividends as received throughout the periods in the stock of the company making the payment. The results referred to thus include dividends and capital appreciation. Stocks listed at the beginning of any time period receive

equal weight. A stock listed after the beginning and before the end of a particular time period is included whenever holders of other stocks receive its shares and also to whatever extent funds for reinvestment become available through the delisting of stocks that still possess value. To be more specific, when a stock is delisted, it is sold over the counter and the proceeds are spread evenly over all stocks listed on that date.

It should be emphasized that Table 1 represents the rates of return—or rates of capital appreciation—that result from the adoption of only one of a large num-

ber of possible investment policies. An illustrative alternative would be the allocation of investment funds in proportion to the value of shares outstanding.

Except for the large amount of double-counting involved, this latter policy, in contrast to the one underlying Tables 1 and 2, would indicate the rates of return available to all the investors considered together.[1] That is, new investors as a

[1] By "double counting" we mean the ownership by one company of stock in another listed company. For example, Du Pont owns large amounts of stock in General Motors. Allocation of investment funds in proportion to the value of Du Pont and General Motors stock outstanding involves counting some of the General Motors shares twice.

TABLE 1—RATES OF RETURN ON INVESTMENT IN COMMON STOCKS LISTED ON THE
NEW YORK STOCK EXCHANGE WITH REINVESTMENT OF DIVIDENDS*

(Per Cent per Annum Compounded Annually)

| | INCOME CLASS | | | | | |
| PERIOD | Tax Exempt | | $10,000 in 1960 | | $50,000 in 1960 | |
	Cash-to-Portfolio†	Cash-to-Cash‡	Cash-to-Portfolio	Cash-to-Cash	Cash-to-Portfolio	Cash-to-Cash
1/26–12/60	9.0	9.0	8.4	8.2	7.4	6.8
1/26– 9/29	20.4	20.3	20.4	20.3	20.3	19.4
1/26– 6/32	−16.5	−16.8	−16.5	−16.8	−16.5	−13.4
1/26–12/40	2.4	2.3	2.4	2.4	2.2	2.4
1/26–12/50	6.8	6.8	6.3	6.1	5.5	5.1
9/29– 6/32	−48.4	−48.8	−48.4	−48.8	−48.2	−40.6
9/29–12/40	− 3.0	− 3.1	− 3.0	− 2.9	− 3.0	− 2.3
9/29–12/50	4.9	4.8	4.3	4.1	3.5	3.2
9/29–12/60	7.7	7.7	7.0	6.8	5.9	5.4
6/32–12/40	21.3	21.1	21.2	20.8	20.8	19.5
6/32–12/50	18.6	18.6	17.8	17.5	16.5	15.5
6/32–12/60	17.4	17.3	16.5	16.2	15.0	14.1
12/50–12/52	12.5	12.0	11.1	10.0	9.0	7.1
12/50–12/54	17.9	17.6	16.6	15.3	14.4	11.6
12/50–12/56	17.0	16.8	15.8	14.8	13.7	11.4
12/50–12/58	16.5	16.4	15.4	14.6	13.4	11.4
12/50–12/60	14.8	14.7	13.8	13.1	12.0	10.3
12/55–12/56	6.4	5.4	5.7	4.6	4.0	2.8
12/55–12/57	− 3.7	− 4.2	− 4.4	− 4.0	− 6.0	− 4.2
12/55–12/58	13.0	12.6	12.2	11.1	10.5	8.2
12/55–12/59	14.0	13.7	13.3	12.2	11.6	9.2
12/55–12/60	11.2	10.9	10.5	9.6	8.9	7.2

* The data underlying this table have been exhaustively checked. We are confident that any subsequent refinement or adjustment will do no more than change an occasional figure after the decimal point.

† "Cash-to-Portfolio" means the net rate of return which would have been realized after paying commissions and taxes (if any) on each transaction but continuing to hold the portfolio at the end of each period.

‡ "Cash-to-Cash" means the net return which would have been realized after paying commissions and taxes (if any) on each transaction including the sale of the portfolio at the end of each period.

group might be able to allocate funds among securities in proportion to the value of shares outstanding but could not invest equal amounts in each security without changing many prices by a substantial amount.

The decision to invest equal amounts in each company with common stock listed on the Exchange was based on the desire to calculate rates of return that would on the average be available to the individual investor who selected stocks at random with equal probabilities of selection—that is, exercised no judgment. A policy of allocating funds in proportion

to shares outstanding or according to any other criterion implies less neutrality of judgment in making investments.[2]

Since dividends constitute income, the results of reinvesting dividends would

[2] We made an estimate of the rates of return available from investing in proportion to the value of shares outstanding by using the Standard and Poor's Index of Common Stocks on the New York Stock Exchange—an index weighted according to the value of shares outstanding—and the Standard and Poor's dividend series. For three of the five starting dates covered in Table 1, the Standard and Poor's rate of return was generally higher; for two of the starting dates, the yield was lower. No systematic differences were detectable, and on the average the yields based on the two different investment policies were similar.

TABLE 2—RATES OF RETURN ON INVESTMENT IN COMMON STOCKS LISTED ON THE NEW YORK STOCK EXCHANGE WITHOUT REINVESTMENT OF DIVIDENDS*

(Per Cent per Annum Compounded Annually)

| | INCOME CLASS | | | | | |
| | Tax Exempt | | $10,000 in 1960 | | $50,000 in 1960 | |
PERIOD	Cash-to-Portfolio†	Cash-to-Cash‡	Cash-to-Portfolio	Cash-to-Cash	Cash-to-Portfolio	Cash-to-Cash
1/26–12/60	6.9	6.9	6.7	6.5	6.2	5.8
1/26– 9/29	19.8	19.7	19.8	19.7	19.8	18.9
1/26– 6/32	−13.2	−13.4	−13.2	−13.4	−13.2	−10.8
1/26–12/40	1.6	1.6	1.6	1.7	1.6	1.8
1/26–12/50	5.1	5.1	4.9	4.7	4.5	4.2
9/29– 6/32	−48.2	−48.6	−48.2	−48.6	−48.1	−40.4
9/29–12/40	− 4.9	− 5.0	− 4.9	− 4.7	− 4.8	− 3.8
9/29–12/50	2.3	2.2	2.0	2.0	1.8	1.7
9/29–12/60	4.9	4.9	4.6	4.4	4.3	3.8
6/32–12/40	24.5	24.4	24.5	24.1	23.8	22.5
6/32–12/50	21.4	21.4	20.8	20.5	19.3	18.5
6/32–12/60	20.5	20.5	19.7	19.5	17.8	17.3
12/50–12/52	12.6	12.1	11.2	10.1	9.1	7.2
12/50–12/54	17.3	17.1	16.1	14.9	14.1	11.5
12/50–12/56	16.6	16.5	15.5	14.6	13.6	11.3
12/50–12/58	16.2	16.1	15.2	14.4	13.3	11.3
12/50–12/60	15.0	14.9	14.0	18.2	12.1	10.4
12/55–12/56	6.5	5.4	5.7	4.6	4.0	2.8
12/55–12/57	− 3.3	− 3.8	− 4.1	− 3.7	− 5.8	− 4.1
12/55–12/58	12.6	12.2	11.9	10.8	10.3	8.1
12/55–12/59	13.7	13.4	13.0	12.0	11.4	9.1
12/55–12/60	11.1	10.9	10.4	9.6	8.9	7.2

* The data underlying this table have been exhaustively checked. We are confident that any subsequent refinement or adjustment will do no more than change an occasional figure after the decimal point.

† "Cash-to-Portfolio" means the net rate of return which would have been realized after paying commissions and taxes (if any) on each transaction but continuing to hold the portfolio at the end of each period.

‡ "Cash-to-Cash" means the net return which would have been realized after paying commissions and taxes (if any) on each transaction including the sale of the portfolio at the end of each period.

obviously vary with the individual investor's tax bracket. For the purposes of this study, three brackets were chosen, although any other set of tax rates could easily be substituted for those used.

The first category, "Tax Exempt," shows the rate of return that might have been realized by a tax-exempt institution.

Results in the $10,000 income class were computed in the following manner:

1. The marginal income tax rate as of 1960 was figured for a married man with standard deductions and an adjusted gross income of $10,000.
2. For each previous year, a tax rate was determined for a man who held the same relative position in the pattern of income distribution existing in those years as the man with $10,000 income in 1960.
3. The appropriate tax rate for such a married man was then applied in all preceding years.

A precisely analogous procedure was used in computing results in the third category, the $50,000 income class.[3]

For each period in each income class, two rates of return are shown. The first, "Cash-to-Portfolio," shows the after-tax rate of interest, compounded annually, which an individual would have had to get on a sum equal to the gross purchase price of the portfolio at the beginning of a period in order to equal the value of the portfolio at the end of the period. The second figure, "Cash-to-Cash," indicates the rate of interest, compounded annually, that would be required on the value of the initial portfolio in order to equal the value of the sum at the terminal date after selling the portfolio and paying the commissions and the capital gains tax applicable.

[3] For example, men with incomes of $10,000 or $50,000 in 1960 would have had incomes of $1,780 or $8,950, respectively, in 1933. The ranges of marginal tax rates were from 0 to 25 per cent for dividends for the $10,000 man and from 0 to 62 per cent for the $50,000 man. The capital gains tax rates ranged from 0 to 13 per cent for the $10,000 man and from 5 to 42 per cent for the $50,000 man.

These rates of return speak for themselves and require little comment. The periods were chosen for obvious reasons. The period from 1926 to 1960 is a long span with booms and depressions—prime examples of each!—and war and peace. The periods beginning in September, 1929, were included to indicate the experience of those who invested at the height of the stock-market boom of the 1920's. The periods beginning in June, 1932, were included to show the results of investing at the nadir of this country's worst depression. The numerous brief, recent periods were included to bring details of postwar experience into sharp focus.

Aside from most periods ending in 1932 or 1940, the rates of return are surprisingly high. For half the twenty-two periods, the rates are above 10 per cent per annum compounded annually; and for two-thirds of the periods, the rates exceed 6 per cent.

RATES OF RETURN WITHOUT REINVESTMENT OF DIVIDENDS

Table 2 presents the rates of return on common stocks listed on the New York Stock Exchange with no reinvestment of dividends, that is, under the assumption that dividends are spent for consumers goods and services at the time they are received. The method of calculation is analogous to that used to compute the yield-to-maturity of a bond. The cost of the stock at the beginning of a time period is analogous to the purchase price of the bond; dividends are analogous to interest payments; and the value of the stock at the end of the period is analogous to the sum received by the holder of a bond when it matures. The rate of return is the rate of discounting which makes the stream of after-tax cash flows have a present value of zero.

At first, one might expect the rates in Table 2 to be significantly lower than in Table 1, whereas in fact the rates are quite similar for most of the periods. This similarity merely reflects the fact that the rate of appreciation in the prices of stocks after the receipt of dividends was on the average about the same as the rate of appreciation before the receipt of dividends. When the latter rate was higher, the rates in Table 2 are lower than those in Table 1 and vice versa. The individual who reinvests his dividends is obviously wealthier at an ending date than is the individual who spends them. On the average he also has a great deal more invested in his portfolio. Thus, some rates in Table 1 are higher than the corresponding rates in Table 2; some are lower; and several are the same.

Table 3 shows rate of capital appreciation or gain, ignoring dividends.

COMPARISON WITH OTHER INVESTMENT MEDIA

Comparable data for other investment media are not available, but there is some information on realized rates of return on a *before-tax* basis. The most nearly complete data are for savings in commercial

TABLE 3

RATES OF CHANGE IN VALUE OF INVESTMENT IN COMMON STOCKS LISTED
ON THE NEW YORK STOCK EXCHANGE, IGNORING DIVIDENDS*

(Per Cent per Annum Compounded Annually)

| | INCOME CLASS | | | | | |
| PERIOD | Tax Exempt | | $10,000 in 1960 | | $50,000 in 1960 | |
	Cash to Portfolio†	Cash to Cash‡	Cash to Portfolio†	Cash to Cash‡	Cash to Portfolio†	Cash to Cash‡
1/26–12/60.......	3.9	3.8	3.9	3.6	4.0	3.3
1/26– 9/29.......	15.7	15.6	15.7	15.6	15.7	14.8
1/26– 6/32.......	−21.0	−21.2	−21.0	−21.2	−20.9	−17.8
1/26–12/40.......	− 2.8	− 2.9	− 2.8	− 2.8	− 2.7	− 2.4
1/26–12/50.......	1.3	1.3	1.4	1.2	1.5	1.1
9/29– 6/32.......	−51.5	−51.9	−51.5	−51.9	−51.3	−43.4
9/29–12/40.......	− 8.0	− 8.0	− 8.0	− 7.7	− 7.6	− 6.4
9/29–12/50.......	− .5	− .6	− .5	− .6	− .2	− .3
9/29–12/60.......	2.5	2.4	2.5	2.2	2.7	2.1
6/32–12/40.......	16.9	16.8	16.9	16.4	16.8	15.1
6/32–12/50.......	13.1	13.0	13.0	12.5	13.0	11.7
6/32–12/60.......	11.9	11.9	11.9	11.4	11.8	10.8
12/50–12/52.......	5.9	5.4	5.9	4.8	5.9	4.0
12/50–12/54.......	11.4	11.1	11.3	10.1	11.3	8.6
12/50–12/56.......	10.8	10.7	10.8	9.8	10.8	8.4
12/50–12/58.......	10.7	10.6	10.7	9.8	10.7	8.5
12/50–12/60.......	9.6	9.4	9.5	8.7	9.5	7.6
12/55–12/56.......	1.5	.4	1.4	.4	1.4	.2
12/55–12/57.......	− 8.3	− 8.8	− 8.3	− 7.9	− 8.3	− 6.6
12/55–12/58.......	8.0	7.6	8.0	6.9	8.0	5.8
12/55–12/59.......	9.3	9.0	9.3	8.2	9.3	6.9
12/55–12/60.......	6.7	6.5	6.7	5.8	6.7	5.0

* The data underlying this table have been exhaustively checked. We are confident that any subsequent refinement or adjustment will do no more than change an occasional figure after the decimal point.

† "Cash-to-portfolio" means the net rate of change which would have been experienced after paying commissions and taxes (if any) on each transaction but continuing to hold the portfolio at the end of each period.

‡ "Cash-to-cash" means the net change which would have been experienced after paying commissions and taxes (if any) on each transaction including the sale of the portfolio at the end of each period.

banks, mutual savings banks, and savings and loan associations. Savings in these institutions never earned as much as 6 per cent per annum for any of the twenty-two time periods listed in Table 1 and for most of the period 1926–60 earned less than 4 per cent.[4]

Data on mortgage loans made by commercial banks and life insurance companies on non-farm homes are available for 1920–47 and for some subperiods.[5] When all types of mortgages are considered together, realized yields never exceeded 6 per cent, never fell below 4 per cent, and averaged about 5 per cent.

Realized yields on municipal and U.S. government bonds, as indicated by Standard and Poor indexes, ranged from −7.0 to +7.8 per cent and averaged less than 4 per cent during the twenty-two time periods listed in Table 2.

Average realized yields on large issues of corporate bonds are available for 1900–58.[6] These yields ranged from −6 per cent to just over 15 per cent during periods from 1920–43. The very high yields were achieved during the recovery from the depression of 1929–32 when prices of industrial bonds, in particular, advanced sharply in price. During most periods, yields varied between 5 and 8 per cent and in recent periods have been lower.

The fact that many persons choose investments with a substantially lower average rate of return than that available on common stocks suggests the essentially conservative nature of those investors and the extent of their concern about the risk of loss inherent in common stocks. And yet their experience with mortgage foreclosures during the 1930's and the substantial rate of default on bonds during the same period shows that even such "conservative" investments carry considerable risks.

The Center for Research in Security Prices is currently at work on another study that will focus more sharply on this question of risk in common-stock ownership. While the present study is concerned with rates of return on common stocks for various periods, the above-mentioned study will seek to answer the questions of how often and how much the investor might have gained or lost on every stock listed on the New York Stock Exchange from 1926 to 1960.

[4] Based on (1927–60) figures of U.S. Savings and Loan League *Fact Book*, 1955 and 1962 editions, whose sources are as follows: (1) (savings and loan associations) data of members of Federal Home Loan Bank System, (2) (mutual savings banks) Association of Mutual Savings Banks, (3) (commercial banks) data of Board of Governors of the Federal Reserve System and Federal Deposit Insurance Corporation. The yields for these years were computed by dividing the total cash and credited dividends or interst by the average of deposits at the beginning and end of each year.

1926 Figures: (savings and loan associations) R. W. Goldsmith, *A Study of Saving in the United States* (Princeton, N.J.: Princeton University Press, 1955), Vol. I, p. 447; and (savings and commercial banks) J. V. Lintner, *Mutual Savings Banks in the Savings and Mortgage Markets* (Boston, Mass.: Division of Research, Graduate School of Business Administration, Harvard University, 1948), p. 477.

[5] J. E. Morton, *Urban Mortgage Lending: Comparative Markets and Experience* (Princeton, N.J.: Princeton University Press, 1956), p. 114.

[6] W. B. Hickman, *Statistical Measures of Corporate Bond Financing Since 1900* (Princeton, N.J.: Princeton University Press, 1960), pp. 291, 298–301. Hickman's sample included all straight corporate bond issues of $5 million or more.

See also David Durand, "A Quarterly Series of Corporate Basic Yields, 1952–57, and Some Attendant Reservations," *Journal of Finance*, Vol. XIII, No. 3 (September, 1958).

Basic yields, together with some discussion of the basic-yield concept, appeared in David Durand, *Basic Yields of Corporate Bonds, 1900–1942* (Technical Paper No. 3 [New York: National Bureau of Economic Research, 1942]). The series was later brought up to date, with additional discussion, in David Durand and Willis J. Winn, *Basic Yields of Bonds, 1926–1947: Their Measurement and Pattern* (Technical Paper No. 6 [New York: National Bureau of Economic Research, 1947]). Later basic-yield figures have appeared in the *Economic Almanac* and the *Statistical Abstract of the United States*.

CONCLUSIONS

During the entire thirty-five-year period, 1926–60, the rates of return, compounded annually, on common stocks listed on the New York Stock Exchange, with reinvestment of dividends, were 9.0 per cent for tax exempt institutions; 8.2 per cent for persons in the $10,000 income class; and 6.8 per cent for persons in the $50,000 income class.[7] These rates are substantially higher than for alternative investment media for which data are available. It is probably worth noting here that a dollar earning 9.0 per cent per annum, compounded annually, would be worth over $20 in 35 years.

The rates for the postwar periods are substantially higher than for the periods prior to the war except during the period of recovery after 1932. It will perhaps be surprising to many that the rates have consistently been so high. For the postwar period (1950–60), as a whole, rates have exceeded 10 per cent for all tax brackets considered, even after payment of capital gains taxes, and this has been true for most subperiods as well.

Rates of return without reinvestment of dividends varied from period to period above and below rates with reinvestments, but on the whole the rates were similar. This merely reflects the fact that, on the average, rates of appreciation in the prices of stocks after the receipt of dividends were similar to rates before the receipt of dividends.

APPENDIX A: METHODOLOGY

DATA

COVERAGE OF THE DATA

Definition of common stock.—We have used a rather strict definition of common stock. To be considered a common stock, an issue must have residual rights to dividends. However, where the payment of a dividend is a liability for another company, which is the case for most guaranteed stocks, we do not consider the security to be a common stock. We ignore the matter of the voting power of the stock.

Initially we had the naïve belief that it would be easy to distinguish between common stocks and other securities by using the designations employed in the lists in the *Bank and Quotation Record*. But as the study progressed we found over fifty ambiguous designations of securities. It was necessary in each such instance to examine more detailed descriptions to determine the rights of the security-holder and, thus, whether the security was a common stock.

If the company had only one issue of stock, this was almost always found to be common stock; but there were exceptions. For example, Green Bay and Western Railroad has a capital structure consisting of "Class A Income De-

bentures," "Class B Income Debentures," and "common stock." Class A income debentures and the common stock have rights ordinarily ascribed to first and second non-cumulative preferred stocks. The Class B income debentures have the residual rights to the income of the corporation. In our files, we treat the Class B income debentures as the common stock.

Another example is General Gas and Electric's "Class A common stock," which was listed on the New York Stock Exchange from 1925 to 1933. Until 1929, the company also had outstanding dividend certificates, which were traded over the counter. These certificates were entitled to share equally in dividends after the Class A and Class B common stock had each received 5 per cent dividends. For our purposes, General Gas and Electric Company had no common stock listed until 1929.

Radio-Keith-Orpheum had both "Class A" and "Class B common stocks" outstanding. The Class B stock was entirely owned by the Radio Corporation of America. Until April, 1931, the Class A stock was in fact a participating preferred stock under our definition; but then Radio Corporation of America waived its rights to dividends. Under our definition the Class A became common stock at that time.

Sometimes a company has several securities simultaneously listed, each of which we consider

[7] The specification of tax brackets and treatment of taxes are discussed on pp. 3, 6, 16, and 17.

to be common stock. Where all of these issues had a continuous series of price quotations, we treat each of them as a separate common stock. Occassionally voting trust certificates and stock have been traded simultaneously for a relatively short period. In such cases we take as the common stock the issue which has the largest dollar volume of trading.

List of common stocks.—The common stocks included in the study are those listed and traded on the New York Stock Exchange during 1926–60. Our data cover stocks only during the period of their listing. Several issues which were seldom or never traded on the Exchange during the period have effectively been excluded from the study. Two of these issues were essentially worthless by 1926; one was the stock of a holding company whose only asset is included; the remainder are New York banks traded for a short time late in 1927 and early 1928 and delisted in 1928. Whenever a substantial amount of stock remained outstanding after a merger, this remainder is included under a different identification number as long as reasonably reliable price quotations are available.

Period of time.—Our study covers the period from January 30, 1926, through December 30, 1960. This period was selected for a variety of reasons. First, we wanted to include a substantial period of time prior to the great "bull" market of 1928–29. Second, we were limited in our efforts to go back to earlier periods by the difficulties of getting complete data for the period prior to 1926.

The data terminate in December, 1960, because of the amount of time needed to process the information included. We intend to bring the data up to date.

THE BASIC DATA

We have included all data necessary for the determination of realized rates of return. Therefore our file includes price quotations; records of dividends and other distributions of cash and property to common stock holders; information on mergers and exchange offers affecting common stock holders; and, in the case of companies which were delisted during the period of the study, a record of the over-the-counter price shortly after the delisting.

Price quotations.—The prices consist of the last price on the last business day of the month, whenever such prices are available. When there was no trading of a particular stock on the last business day of the month, we use the mean of the bid-and-asked price or the bid price or the asked price.

In our file we have a price for each month that a stock was listed from January, 1926, through December, 1960—a total of more than 350,000 quotations. In addition to each price, we have recorded the transaction volume and the round-lot commission applicable at that time. We have also indicated our estimate of the quality of the price information. The overwhelming majority of prices are sales prices which are deemed to have been "good." We suspect errors in approximately 0.5 per cent of the prices in our file. A substantial proportion of these errors are believed to be caused by errors in our sources. All errors are believed to be small and unbiased.

Data on distributions to stockholders and other capital changes.—We have included data on thirty-nine kinds of distributions to stockholders and other capital changes. These data consist of information on cash dividends, stock dividends and splits, subscription rights, distributions of other stocks, mergers, exchange offers, and reorganizations.

Approximately 90,000 cash dividends are included. For each dividend we have recorded the amount payable (and in the case of liquidating dividends, the percentage liquidation which the dividend constitutes), the payment data, the date of record, and where available, the ex-dividend date on the New York Stock Exchange, and the date of declaration.

These dividends are classified according to their tax status as follows:

1. Ordinary cash dividends, i.e., dividends subject to
   a) Surtax only (1926–35)
   b) The same tax as other income (1936–July 31, 1954)
   c) The 4 per cent dividend credit (August 1, 1954–date)
2. Dividends not entitled to the 4 per cent dividend credit
3. Dividends subject to both normal tax and surtax (1926–35)
4. Distributions of capital gains
6. "Dividends" which are return of capital
5. "Dividends" which are paid in partial liquidation of a company
7. "Dividends" which are reported capital gains retained by a regulated investment company

Dividends which are paid in property other than listed common stock are also included. For these dividends we show the estimated fair market value of the dividend and classify them in the same manner as cash dividends.

For distributions of subscription rights, we have recorded the following information: the number of rights required to subscribe to one additional share (or unit), the subscription price, the market value of one right, the percentage of cost basis for tax purposes which may be allocated to the right, and the record date for the distribution of the rights. We have also classified subscription rights according to whether they were rights to subscribe to the same issue of common stock or rights to subscribe to something else such as preferred stock, stock of another company, bonds, etc.[8]

For each stock dividend, split, and reverse split, we have recorded the percentage change in the number of shares outstanding as a result of the split, and the change of par value occurring at the time of the split. We have also recorded the date of the dividend, or effective date of the split, as well as the record and ex dates.

The information recorded on distribution of other listed common stocks depends on the taxability of the distribution. We have recorded the effective date of the distribution, ex and record dates, a number identifying the stock distributed and the number of shares receivable per share held. For distributions which are treated as non-taxable exchanges (spin-offs) we have also recorded the percentage of cost basis to be allocated to the stock received. For other distributions—those which are treated as dividends, taxable exchanges or returns of capital—we have recorded the fair market value of each share received.

Mergers in which the old and new securities are all listed are recorded at least twice. For the security or securities absorbed in the merger, the following information is recorded:

1. The effective date of the merger
2. A number identifying the successor security
3. The tax status of the merger, i.e., non-taxable exchange, taxable exchange, non-taxable exchange except that a portion of the proceeds may be taxable as income or as capital gain
4. The number of shares of the new security received for each share of the old
5. The dollar value of cash and other property (such as preferred stock and debentures)
6. The fair market value of each new common share except where the merger is a non-taxable exchange

7. Where the merger is a non-taxable exchange, the percentage of the cost basis to be allocated to the new common stock
8. Where a portion of the proceeds of the exchange of stock is taxable or may be taxable under a special rule, the maximum amount so taxable

Mergers are further classified into three categories: those which are probably deemed mergers in the legal sense; those effected through the voluntary exchange of stock where at least 50 per cent of the publicly held stock was exchanged under the terms of the offer; and exchange offers where less than half of publicly held stock is exchanged. Two other kinds of "merger" items are included in the file: reorganizations which are most easily handled by treating them as if they were mergers and occasional changes in security identification numbers to simplify the analytical programs.

For the emergent security, the date of the merger and the identity of the security or securities merged are recorded.

Data on other changes.—When a security ceases to be traded on the New York Stock Exchange for any reason other than by being absorbed by a merger into another listed common stock, we have recorded the mean of the bid-and-asked prices over the counter at the end of the month in which the delisting occurs.[9]

In order to facilitate reference to the file, we have recorded effective dates of changes of name and have generally noted reported reorganizations and bankruptcy proceedings.

Organization of the data.—The data have been arranged to facilitate the use of an electronic computer in making the calculations reported in the text. Our file is recorded on magnetic tape.[10] All data pertaining to a particular common stock are arranged chronologically within the records for that stock. The sequence of securities is essentially alphabetical within the restrictions imposed by two requirements: (1) stocks which

[8] Where the subscription rights apply to "units" including the same issue of stock, we arrive at the value subscription price per share by deducting the value of other securities making up the unit from the gross subscription price.

[9] In a few cases the price recorded is the price on some other stock exchange rather than the over-the-counter price. Very occasionally a security had substantial value at the time of delisting, but we could find no over-the-counter price until several months after the delisting. In these cases we have used the first reasonable price available.

[10] The number of reels of tape occupied by the file depends on the particular use to be made of it the next time we use the data, e.g., when arranged for printing, it requires eight reels of magnetic tape; when arranged for the computations described in the text, two reels.

have been delisted without being merged into another listed stock are at the beginning of the file arranged in the order of the dates of their delisting, and (2) stocks whose holders received another stock precede the other stock in the file. Each event—the price at the end of the month, the occurrence of the ex date of the dividend, the distribution of rights, etc.—constitutes an independent "record" which contains information pertinent to that event.

Unless the majority of the stock is exchanged, the data for a particular stock are kept together regardless of name changes. If the stock is still listed after an exchange, the remainder of the record for that stock constitutes information for another "security" in our file.[11]

METHODS OF COLLECTION AND CONTROL[12]

The methods of collection and control are complicated. We have organized the material in the following section approximately in the chronological order in which the project progressed.

Strenuous efforts were made to keep all clerical procedures as simple as possible. This effort was made in order to reduce the amount of error in the file by making much of the work within the capacity of clerks and leaving for the computer that work for which it had superior efficiency. The first step in carrying out this policy was to get for each month a list of companies with stocks listed on the New York Stock Exchange. This was done by reference to the files of the Exchange.

Collection of price data.—The list for each month was then arranged in alphabetical order and an IBM card for each listed company was punched and printed with the company name, the month and year, and an identifying number. These cards were believed to be in the same order as lines in the *Bank and Quotation Record*, which was the basic source of information on stock prices and volume of trading.[13] Clerks were instructed to write relevant data in prepared spaces on these specially designed punch cards. They were to fill out a card for each price quotation other than one which was specifically stated as being for preferred or preference stock. If there was a prepunched card it was to be used. If there was no prepunched card, a blank card was to be filled out. Prepunched cards for which no information was found were put in a special file. In addition to the information described in the preceding section, the clerk placed a code on each card designating the kind of stock (e.g., 1 for common stock: 2 for Class A; 3 for Class B; 4 for Class C; — for new or when-issued, etc.). Also recorded was the par value, which was useful as a quality-control device. Prices were recorded precisely as printed. When there was no sale price, the clerk recorded the bid price, the asked price, or if available, both. Conversions of fractions to decimals and computation of the means of bid and asked prices were later performed on the computer. A staff of approximately forty full-time and part-time clerks worked approximately 6 months to code these price data.

The punching of the partly prepunched cards was completed by one keypunch operator. Rather than check the accuracy of the keypunching alone by means of the usual methods of verification, the quality of both clerical recording of prices and the keypunching was checked by means of independent coding and punching a sample of 5,000 price quotations. We have relied on this sample check of the coding and punching and the consistency checks which will be discussed below.

The accuracy of the recording and punching of data was estimated from this sample, which had 100 clusters of 50 entries each. In all, 133 prices were found to be in error. Of these errors, approximately 79 per cent were of the sort which would be detected by the computer programs and corrected. Thus, approximately 0.5 per cent of the total number of prices would continue to be erroneous after all data-processing. By a coincidence which will not be described, it was found that the original data after correction through data-processing would be 99-44/100% pure. The errors which are not detect-

---

[11] Where we gathered information before discovering that a particular security is not a common stock for our purposes, we have nevertheless retained that information in our file and have placed special indicators on it. We have not, however, applied the error detection and correction procedures reported below to most of these data.

[12] Much of the material in this section is also presented in L. Fisher, "Use of Electronic Computers in the Quality Control of Financial Data" in *1963 Proceedings of the Business and Economics Section of the American Statistical Association* (Washington, D.C.: American Statistical Association, 1963), pp. 169–71.

[13] Prior to February, 1928, the basic reference was the predecessor of the *Bank and Quotation Record*, the "Bank and Quotation Section" of the *Commercial and Financial Chronicle*.

able by our data-processing procedures are small and unbiased.

We began with approximately 400,000 pre-punched cards. Of these the clerks filled out 380,000, leaving 20,000 blank. In addition, the clerks found about 10,000 prices for which they had no prepunched cards. Many of these additional prices were for stocks which were not in the list obtained from the files of the New York Stock Exchange. The prepunched cards that had either sales or both bid-and-asked prices for issues designated as common stock in the *Bank and Quotation Record* were sorted by identification numbers and date and placed on magnetic tape. This file was called the "main file." All other cards were sorted in the same manner and listed before they were also placed on magnetic tape. This file was the "hold-out file." From this list the decision was made as to which of the stocks among those not specifically designated as common in the *Bank and Quotation Record* were common stocks for our purposes.

Initial error detection and correction.—The first correction procedure applied was the detection prices which were invalid on their face. Both files were checked to make certain that fractions were valid, that is, that the numerator was an odd number smaller than the denominator and that the denominator was a power of 2. Where both a bid and an asked price appeared, they were compared to make sure that the bid was less than the asked price. The "mail file" was also checked to make sure that each entry in fact contained either a sale price or both bid and asked prices. The same additional test was applied to the "hold-out file" except that the existence of only a bid price or only an asked price was ignored if the volume of trading was less than one thousand shares for the month. All errors detected by these procedures were corrected.

Collection and initial correction of cash dividends.—Two different methods had to be used for the collection of cash dividends. For the period 1937–60, annual issues of *Moody's* or *Standard and Poor's Dividend Records* or the annual section of *Standard Corporation Records* were used, depending on which was in the University of Chicago Library. For the period 1926–36, *Moody's Quarterly Dividend Record* was used. Since the only known complete file of this last publication is in Moody's New York offices, the data from the earlier period were recorded by trainees working for the research division of Merrill Lynch, Pierce, Fenner & Smith Inc.

For each dividend the data described above were recorded, and also recorded was a code to indicate whether the dividend was annual, quarterly, monthly, or extra, etc. As a cross-check, annual total payments per share were taken from the appropriate *Moody's Manual.* The computer was then used to check the annual totals of the individual dividends taken from the dividend records against the annual totals from *Moody's Manuals.* All discrepancies were resolved.

Other distributions to stockholders, mergers, exchange offers, name changes, and reorganizations.—After preliminary experiments, data on other distributions to stockholders, mergers, exchange offers, name changes, and reorganizations were taken from the Commerce Clearing House Corporation's *Capital Changes Reporter.*

Nine specially designed printed forms were used for this purpose. After the initial coding and punching, a printed list was made, classified by company name. This list was compared with the original source to make sure that the interpretation of the entries was correct and that capital changes had been collected for each company. For companies for which a list of capital changes did not appear in the *Capital Changes Reporter* or was incomplete, reference was made to a variety of sources including Prentice-Hall's *Capital Adjustments, Moody's Manuals,* the *Commercial and Financial Chronicle,* and *Poor's Daily Digest of Corporate News.*

At the same time that the list of capital changes was being checked against the source, the information it contained about mergers and name changes was used in conjunction with our now-corrected information on dates of trading to assign a new number to each new security so as to place it in the appropriate order within the file. When the errors detected by this check had been corrected, the file of capital changes was recorded on magnetic tape and the security identification numbers were placed in the file.

Because there were thirty-nine different kinds of records in this file, a computer program was written which decoded the information to provide a verbal description of each item. These verbal descriptions were then checked against the sources of data, and more errors were detected and corrected.

Tax status of distribution of cash.—Since the existence of a special tax status for a cash dividend is generally not shown in any of the dividend records, it was necessary to find out about any special tax status from other sources. We

used the appropriate sections of the *Capital Changes Reporter* and *Capital Adjustments* for this purpose. However, we then found that our records of distributions of cash had come from three sources, each of which had only a portion of the information required. We therefore wrote and used computer programs to collate the information about the distributions from the three sources.

Editing and detecting additional errors in the price-volume files.—The security numbers were applied to the two files of price and volume information by means of a computer program. Fractions were converted to decimals; bid and asked prices were averaged; the commission applicable to each price of each stock was computed, and the sums of the price and commission were inserted into the file. After correction for prices which had not been assigned new identification numbers with the original program, these files were sorted according to the new number and merged together.

Until this time the computer we used was the UNIVAC I at the University of Chicago. The file of prices then consisted of thirty-three reels of magnetic tape. The University was installing an IBM 7090 computer (which has since been converted to an IBM 7094). At this time the entire file was printed, and the UNIVAC I tapes were copied onto IBM tapes through the data conversion facilities of the Reader's Digest Association.

The error detection program designed to make price data "99-44/100% pure" was now written for the IBM 7074 computer owned by Merrill Lynch, Pierce, Fenner & Smith Inc. and used. The program performed the following tasks:

1. Except for securities we do not regard as common stock, price quotations were placed into two categories:[14]
   a) Sole quotations for a security for a month
   b) One of several quotations for a security for a month
2. Quotations in category 1b were further broken down
   a) The quotation with the highest product of price and trading volume for the month
   b) All other

(If there was no trading the quotation with the highest price was placed in category 2a. If there was a tie, the first quotation in the file was placed in category 2a.)

[14] If there were identical quotations, all but one were deleted.

3. Quotations in categories 1a and 2a were further divided into three subcategories each:
   a) The magnitudes of the changes from both the preceding quotation in category 1a or 2a for this security to this quotation and from this quotation to the next such quotation exceed a point plus 10 per cent of the first quotation of the pair and the changes are of opposite sign—the price is "bad"
   b) The price is not bad but one of these changes exceeds a point plus 25 per cent—it is "suspicious"
   c) The price is neither bad nor suspicious—it is "good"
4. For each security, a one-to-three-line list of time periods for which a quotation falling into categories 1a or 2a is in the file for each month was prepared.
5. For each gap in the series of price quotations for a security, further action was taken. If the gap was for less than 12 months, dummy records were created. For longer gaps a special report was made.
6. A two-page summary of the price information for each common stock was prepared. For each month, January, 1926–December, 1960, this summary shows the category of the real or dummy quotation and the price.

The output from the program consisted of several files, each on one or more reels of magnetic tape:

1. The main file consisting of all good, suspicious, and duplicate (category 2b) quotations, and all quotations which were for other than common stock or which show a volume of trading but no price
2. A "bad and missing" file containing the dummy and bad quotations
3. A suspicious file[15]
4. A "multiple-price" file showing all quotations in category 2
5. A list of deletions

All of these files, except the main file, were printed and examined in order to make corrections and to revise our plans for the next steps in preparing the data. For January, 1926, and December, 1960, it was impossible for the program to call a price bad. Suspicious prices for these months were treated as if they were bad. A sample of the remaining suspicious items was checked. In almost every case the price agreed with a price in the *Bank and Quotation Record*.

For the calculations reported in the text we did not think it necessary to correct all the bad and missing items. We checked all quotations classified as bad and missing for the beginning

[15] Files 2 and 3 were sorted by date and company name as well as by security number and date.

or ending dates of the various periods studied. We further corrected all discrepancies for a number of other dates where, through printer's errors, prices had not been aligned with company names in the *Bank and Quotation Record* (in this case several consecutive prices had been attributed to wrong securities). In addition, all discrepancies were checked for a sample of 115 securities.

Construction of the combined file.—The separate files for prices, cash dividends, and ten groups of other events were rearranged to provide a convenient, common format and were merged together. The arrangement of the data is described in Appendix B. Critical discrepancies, discovered by checking the output of the error detection program just described were corrected and several thousand additions were made by means of special correction programs. This first application of the correction programs altered data for about 16,000 events.

Error detection procedure for the combined file.—The data in the combined file had come from several independent sources. Hence it was likely that inconsistencies would be found. For example, when a company's stock is split, this action must generally be approved by a meeting of the stockholders. The *Capital Changes Reporter* states the date the action was approved by the stockholders. Trading in the new (split) shares cannot take place until the stock exchange has been notified of the stockholders' action. Hence, the split may not actually be effective until the next month or even later. Thus it was necessary to search the file for further inconsistencies in the data.

A fairly elaborate program for the IBM 7094 was used for this purpose. This program, which we have called "Filter King," reports the following kinds of inconsistencies:

1. Some or all of the events for a security may be misplaced within the file if
   a) Price quotations are missing for more than 12 consecutive months
   b) Price quotations are available on the Exchange after the date at which we thought the stock had been exchanged through a merger or after the date for which we had collected an over-the-counter price
   c) The information for a security was above that of a company whose stockholders received this security through merger or other distribution
2. Prices or capital changes may have been misrecorded if a price quotation is inconsistent with previous and/or subsequent quotations, adjusted for dividends, stock dividends, rights, etc.

The program also performs other functions: (1) If a stock is being traded with a stock dividend on and also on a when-issued basis, the program chooses the price consistent with the date of the stock dividend. (2) Visual reference to the information in the file is not convenient. When printed, the file takes about 22,000 pages of 11 × 15-inch paper. Therefore, Filter King produces a 2,500-page summary of the file. This summary contains the most recent name and the identification number of each security and part of the information about each event, that is, for prices, the year and month, whether the price is consistent with previous and subsequent prices and the price itself; for other events, the year and month of the ex date, the kind of event, and the most useful number pertaining to the event (the amount of a dividend; the size of a stock dividend, split, or reverse split; the value of one right; the number of shares received in a spin-off or merger); but other information is omitted.

Filter King was used twice to test the entire file before we made the computations reported here. Between the first and second runs about 13,000 changes were made in the file and about 4,000 changes after the second run. All possible inconsistencies in the arrangement of the file were investigated and action taken to correct them. All inconsistencies in price quotations which were near capital changes in time or which suggested the existence of an unrecorded capital change were also investigated. All inconsistent or missing price quotations for dates of sale or purchase of a portfolio were checked again.

Other inconsistencies were left until later. We are now investigating and correcting them. The process will be completed about the time this article is published.

CALCULATION OF RATE OF RETURN

In this section we describe the methods we have used to calculate rates of return. Since the programs for calculating rates of return with and without reinvestment of dividends use nearly identical procedures, we shall first describe the calculation of rates of return with reinvestment of dividends and then indicate the changes required to compute rates of return without reinvestment.

The calculation of average rate of return for all listed common stocks over a particular period of time requires that we know how much was invested at the initial date and what the value

of the portfolio is at the final date. Since our investor is required to pay federal income taxes, we must also know what the cost basis for tax purposes (or simply "basis") is. The file is arranged by security and chronologically within each security. The computer program for calculating rate of return with reinvestment of dividends selects one item from the file at a time and then takes appropriate action. During the course of the computation several additional files are maintained internally:

1. A file of regular income tax rates—one entry for each month
2. A file of capital gains tax rates—one entry for each month
3. A file showing the current number of shares held, the cost of these shares, and the number of dollars available for reinvestment at the end of the current month. All of these items pertain to the security currently under consideration
4. A deferred file which shows the effective date, that is, payment date, of a distribution to stock holders or other capital change and its effect on the number of shares held, the cost of shares, and/or funds available for reinvestment. An entry is made in this file when the ex date for some distribution occurs and is maintained until the distribution date
5. A cross-reference file which contains the identification number of a security distributed as a result of a spin-off or merger, the identification of the security whose holders receive something as a result of a spin-off or merger, the effective date of the spin-off or merger, and the number of shares, their cost basis, and the amount of funds received. An entry is made in this file whenever a spin-off or merger occurs and is maintained until the program is processing information for the security received
6. A file showing the amount to be reinvested in each security each month through the proceeds of sales of securities over the counter
7. The running total of the initial investment in securities
8. Running totals of the value of the investment in each security at the ending date of the study. Three such totals are maintained: one showing the market value of the portfolio, one showing net proceeds from the sale of the portfolio before payment of the capital gains taxes, and one showing net proceeds after payment of capital gains taxes

An event may directly affect the number of shares held, the cost basis, or the funds available for investment. In general the computations are performed as follows: when the first event for a stock is read into the computer, the number of shares, basis, and funds are set to zero. When

the price for a beginning date of the study comes up, we add $1,000 to the initial investment, place $1,000 into the basis, and divide $1,000 by the price plus commission to compute the number of shares held initially.[16]

Most dividends and other distributions to stockholders are based on the number of shares held on the day prior to the ex-dividend date and not on the number of shares held on the distribution date. Therefore dividends, etc., have been placed in the file according to their ex dates. When a new event is perceived, its key date is noted, the deferred file is scanned, and processing of all items in the deferred file with effective dates less than the current key date is completed. Completion of the processing requires that the change in shares be added to the current number of shares held, change in basis be added to the current basis, change in funds available be added to current funds available, and the entry in the deferred file be destroyed.

When a price is found, the next step is to determine whether the price is for an ending date of the study. If it is for an ending date, processing of all remaining items in the deferred file is completed, and the appropriate computation of the value of the interest in the security is made for each of the three assumptions about valuation of the proceeds and added to the running totals. If the price is not for an ending date for the study, the funds currently available are added to the current basis, funds are divided by price plus commission and the quotient added to the current number of shares, and funds are set to zero.

When a dividend is perceived, a new item is created in the deferred file. If the dividend is a cash dividend, current number of shares is multiplied by the amount of the dividend per share, the product is reduced by the amount of tax liability and placed in "change in funds." If the dividend causes a change in basis, the change in basis is also recorded.

When a stock dividend, split, or reverse split is perceived, a new item in the deferred file is created and the change in the number of shares held is recorded.

Perception of the description of a distribution of subscription rights requires the following ac-

[16] If the company has more than one common stock at the beginning date, the $1,000 is divided equally among the common stock issues, i.e., if there are two common stocks the initial investment in each stock is $500.

tion. If the subscription rights are for the same security, it is assumed that enough rights are sold so that after taxes the remaining rights may be exercised. The changes in number of shares held and basis are added to current number of shares and current basis. If the rights are not to the same security, they are treated as though they were dividends with the appropriate tax status and the current basis and funds currently available are adjusted.

When a spin-off occurs, both a deferred item and an item in the cross-reference file are created. The deferred item shows the change in basis. The cross-reference item shows the number of shares to be received after allowing for taxes and cost basis of these shares.

When a merger is perceived, processing of all items in the deferred file is completed and a cross-reference item is created. The cross-reference item shows the after-tax figures for the number of shares received, basis, and the funds received in exchange for the old stock. To funds in the cross-reference item, currently available funds are added. Current shares, basis, and funds are then reduced to zero.

The notation of the receipt of shares and/or funds as the result of a spin-off by another security or the merger of another security causes the cross-reference file to be searched. When the appropriate item is found, the shares, basis, and funds shown there are added to the respective current items; and the item in the cross-reference file is destroyed.

When an over-the-counter price is perceived, processing of the remaining items in the deferred file is completed; the current number of shares is multiplied by the over-the-counter price; and the product, adjusted for tax liability on credit, is added to current funds. The current funds are then added to the file of funds received through over-the-counter sales.

When the end of the file has been reached and the computations for the last security completed, the discount rates on investment are calculated. The calculation is performed by finding the ratio of the value of the portfolio or cash at the ending date by the total initial investment, taking the natural logarithm of the ratio and dividing it by the length of the period to find the rate of return, compounding continuously. This rate is then converted to the equivalent rate of return, compounding annually.[17]

Rates of return without reinvestment of dividends.—The calculation of rates of return without reinvestment of dividends required several changes in the computer program. The deferred file must also show the amount of cash to be paid out. A file showing the amount of cash paid out each month must be maintained. Portions of the subprograms for handling cash dividends other than partial liquidations, rights to subscribe to other securities, mergers, and several housekeeping functions had to be altered to maintain the file of cash paid out. The major change, however, was in the actual computation of rate of return. The rate of return is the discount rate which makes the present value of the income stream purchased equal to the initial investment in the portfolio. This income stream consists of dividends received each month (assumed to be at midmonth) and the final value of the portfolio.

A new method of making successive approximations was used. This method requires about five to ten iterations, depending on the length of the period and the magnitude of the rate, to compute the rate of return to more than seven significant decimal digits. The program for making the computations shown in Table 2 required a total of 25 minutes on the IBM 7094.

[17] The actual computer program performs all of the calculations reported in Table 1 in one pass of the file. It can compute results for six starting dates, twelve ending dates, and three sets of tax rates (one of which is zero) in about 22 minutes.

APPENDIX B: ORGANIZATION OF THE FILE OF BASIC DATA

The master file of basic data is recorded by the IBM 7094 in BCD mode at 556 characters per inch on tapes written by IBM 729-II or 729-VI tape drives. Logical records, each 180 characters long, are in blocks of seven. Each logical record contains the information pertinent to one "event." For our purposes an event is the record of the price of a stock at the end of a month, the information describing a cash dividend, stock dividend, split, distribution of rights, spin-off of another stock, description of a merger or exchange offer in which the security is involved, a name change, or an explanatory note.

The logical records are organized as follows:

Characters 1–5—Security identification number (numeric)

Characters 6–13—A numeric key for sorting records within a security

Characters 6–9—Year and month, e.g., 3507 for July, 1935

For end of month (price-volume) records

Characters 10–13—Contain a code showing the quality of price information in terms of how nearly consistent the quotation is with previous and following quotations, and the amount of cross-checking we have done on it; whether the quotation is for a sale, a mean of bid and asked prices, only a bid or asked quotation, or shows only the value of trading for the month; and a sorting number if more than one quotation exists for the month

For other events

Characters 10–11—Show the day of the month on which the event occurred. For most such items the date used is the stock exchange ex date

Characters 12–13—Are reserved for sorting within the day

Character 14—Contains a mnemonic code for the name of the security, e.g., *A* for Class *A; F* for Certi*F*icate; *M* for A*M*erican shares

Character 15—Contains an alphabetic code whenever the ex date has been estimated from record or distribution dates

Characters 16–17—Contain a two-digit numeric code showing the kind of item described by the record:

  (−7)–(−5)—Receipt of shares through merger, spin-off or other distribution by another listed company

  00–09—Price-volume items

  10–19—"Cash" dividends

  20–24—Stock dividends, splits, and changes in par value

  30–42—Subscription rights

  51–60—Distributions of other listed common stocks

  70–79—Mergers and exchange offers

  80—Name changes

  83—Record of the over-the-counter price immediately after a security has been delisted on the stock exchange

  90–99—Notes

Character 18—For items of kind 70–77 contains a numeric code showing the taxability of the exchange

Character 19—For items of kind 30–79 contains an alphabetic code showing whether prices are sales or bid and asked averages

Character 20—For items of kind 10–19 contains a numeric code showing whether or not the dividend was paid in cash

Characters 21–26—Show the effective or distribution dates for items of kind 10–83

Characters 27–30—Are reserved for an industry code

Characters 31–80—Are divided into five 10-column decimal fields. The contents of these fields are shown in Table B1

Characters 80–86—Are used for events of kinds (−7)–(−5) and 51–79

  Characters 81–85—Contain the security identification number of the security whose holders receive shares (kinds −7, −6, −5) or whose shares are received (kinds 51–79) in a distribution or exchange of stock

  Character 86—Contains the mnemonic name code, described under character 14, for the other security

Characters 87–140—Are used in alphanumeric text. For price-volume items the field contains the name of the company issuing the stock, par value, and a "company number" which was used for the initial data collection of the project. For dividends, the field contains the name of the company and, if a "cash" dividend was really property such as preferred stock, a description of what was distributed. For a name change, the field contains a new name. For a note, the field contains the text of the note

Characters 141–46—Contain the declared date for a dividend

Characters 147–52—Contain the record date for a dividend or other distribution

Characters 153–58—Contain an "alpha" number used in a later stage of data collection (a single security may have more than one alpha number and, conversely, a particular alpha number may correspond to more than one security). The "alpha" number simplifies alphabetizing the file

Characters 159–64—Contain an alpha number corresponding to the security number in characters 81–85

Characters 165–80—Do not contain significant information. In effect they are padding so as to make the logical record length divisible by either 5 or 6 and thus allow the file to be used by a variety of computers

The assignment of security identification numbers was made in conformity with the following rules: securities which are common stocks under some definitions but not under the definition used by us are given numbers less than 26,000 if no listed stock of the company fits our definition of common stock, and a fifth digit of nine if the security under consideration falls into this category but at least one listed stock of the company does meet our definition. The numbers must be chosen so that, in items of kind 51–78 for the security, the security identification number (characters 1–5), will be less than the number in Characters 81–85. The security identification numbers have also been chosen so that the numbers for securities for which an item of kind 83 appears will be a monotonically increasing function of the year and month associated with an over-the-counter

price, and so that all of these security identification numbers are smaller than those assigned to securities which were still listed at the end of December, 1960.

A more detailed description of the "kind" numbers follows:

00–09 Price-Volume Items

00—The only price-volume item for this security this month and the item does not fall into the category receiving kinds 02, 05, 08, or 09

01—There is at least one other price quotation for this security this month. This is the "best" quotation available for making up a continuous series and does not fall into categories 03, 06, 08, or 09

02—An item which would have received 00, except that the price is somewhat out of line with previous and subsequent quotations

03—An item which would have received 01, except that the price is somewhat out of line with previous and subsequent quotations

04—There is another price quotation for this security this month which is "better," and the item does not fall within the category receiving 09

05—An item which would have received 00 or 02, but the price is very much out of line with previous or subsequent quotations

06—An item which would have received designation 01 or 03, but its price is very much out of line with previous or subsequent quotations

TABLE B1

CONTENTS OF THE FIVE "AMOUNT" FIELDS IN MASTER FILE OF INFORMATION
(All Amounts Are for One Share)

Kind	Amount 1	Amount 2	Amount 3	Amount 4	Amount 5
(−7)–(−5)	0	0	0	0	0
00–06	Price	Volume	Asked Price	Price+Commission	Weight
07	0	0	0	0	Weight
08	0	Volume	0	0	0
09	Price	Volume	Asked Price	Price+Commission	0
10–13, 15–19	Dollars per share	0	0	0	0
14	Dollars per share	Per cent of basis to dividend	0	0	0
20–23	Per cent change in shares	0	0	Old par	New par
24	Per cent change in shares	0	0	Value of 1 share	0
30–41	Value of 1 right	Per cent of basis to rights	Reciprocal of subscription ratio	Subscription price	0
42	Value of 1 right	0	Reciprocal of subscription ratio	Subscription price	0
51	No. of shares per share	Per cent of basis to new stock	0	0	0
52–60	No. of shares per share	0	0	Value of 1 new share	0
70–79 (tax status 0)	No. of shares per share	Per cent of basis to new stock	Total dollar value of cash and other property	Value of 1 new share	Portion of amount 3 which is a dividend
70–79 (tax status 1)	No. of shares per share	0	Total dollar value of cash and other property	Value of 1 new share	Portion of amount 3 which is a dividend
70–79 (tax status>1)	No. of shares per share	Per cent of basis to new stock	Total dollar value of cash and other property	Value of 1 new share	Maximum gain recognized
80, 90–99	0	0	0	0	0
83	Price	0	0	0	0

TABLE B2

Kind		No. Used in Computations
−7 to −5.............		*470*
−7...............	333	
−6...............	70	
−5...............	67	
−4.....................		*2,119*
00–06...................		*358,438*
00............	352,424	
01............	2,343	
02............	429	
03............	9	
05............	3,208	
06............	25	
07.....................		*2,772*
10–12..................		*90,572*
10............	413	
11............	88,731	
12............	1,698	
13.....................		*324*
14.....................		*157*
15.....................		*0**
16.....................		*1,712*
18.....................		*232*
19.....................		*227*
20–24..................		*3,157*
20.............	2,202	
21.............	876	
22.............	73	
24.............	6	
30–31..................		*867*
30.............	839	
31.............	28	
41–42..................		*349*
41.............	281	
42.............	68	
51–53..................		*67*
51.............	57	
52.............	2	
53.............	8	
60.....................		*70*
70–78..................		*333*
70.............	164	
71.............	50	
73.............	55	
74.............	12	
76.............	40	
77.............	7	
78.............	5	
83.....................		*268†*
Total..............		462,134

* Six of these events should have been present but had not yet been added to the file.

† Plus one dummy event where the over-the-counter price was omitted.

07—A dummy record created because there is no price quotation in our file for this month and we believe there should be

08—Contains a volume of trading but no price

09—A price quotation for a security which has been given either a security identification number less than 26000 or an identification number with a fifth digit of 9

10–19 Dividends Paid in Cash or Property Other Than Listed Common Stock

10–12—Dividends or portions of dividends taxed as ordinary cash dividends

13—Dividends taxed as ordinary income (not eligible for dividends received credit)

14—Partial liquidation

15—Undistributed capital gains

16—Return of capital

18—Distributed capital gains

19—Dividend subject to both normal tax and surtax when ordinary dividends were subject only to surtax

20–24 Stock dividends, splits, and reverse splits

20—Stock dividend

21—Stock split

22—Reverse split

23—Par change only

24—Optional stock dividend

30–42 Subscription rights

30—Rights to subscribe to additional shares of the same stock

31—Rights to subscribe to units which include the same stock

41—Non-taxable rights to subscribe to some other security

42—Taxable rights to subscribe to some other security

51–60 Distributions of Other Listed Common Stock (Old Continues To Exist)

51—Non-taxable distributions (spin-offs)

52–53—Distributions where capital gain or loss may or must be recognized

60—Distributions taxable as cash dividends

70–79 Mergers and exchanges of stock

70–71—Exchanges of stock through outright merger

73–74—Exchange offers accepted by a majority of the publicly held stock

76–77—Exchanges in reorganization

78—Record of change in security identification number to make identification numbers conform with rules stated above

79—Exchange offers where majority of publicly held stock was not exchanged or where the exchange did not include common stock and the security continued to be traded

80–99 Miscellaneous

80—Name change

83—First price quotation available (at the end of a month) following a delisting or suspension of

a security which was not accompanied by an item of a kind 70–77

90—Note that the amount of money or price shown in the initial coding of an adjacent item was not in U.S. dollars

91–92—Note that another security was merged into this one

93—Note that this security was merged into another and the description under the other security had to be consulted in our sources

94—Notation of receivership, liquidation, or worthlessness of stock

95–96—Notation of distribution of security to holders of senior securities of the company

97—Notation of issuance of convertible bonds or convertible preferred stock

98—Notation of call of convertible bonds or convertible preferred stock and redemption

99—Miscellaneous notes

−7, −6, −5 *Cross-reference Notes*

  −5—Security spin-off

  −6—Security distributed as a dividend

  −7—Security distributed in a merger or exchange of stock

A condensation of the file was used for studies reported here. The condensed file omits all items for securities which are not common stocks. Also omitted are items which are not needed for these studies, viz., kinds 04, 07–09, 23, 80, 90–99. Kind −4 is added to show the current name of security, taking its information from the first item for a security or from kind 80. The condensed file is recorded in binary form at a density of 800 characters per inch. Each item is 11 computer words long. Physical records are 1,100 words long.

The information included is taken from characters 1–12, 16–18, 21–26, and 31–85 of entries in the master file.

The count of each kind of item is shown in Table B2.

In the future, the master file will be kept in another condensed binary form, which will allow all the data to be contained on a fifth as much magnetic tape as the present (BCD) master file.
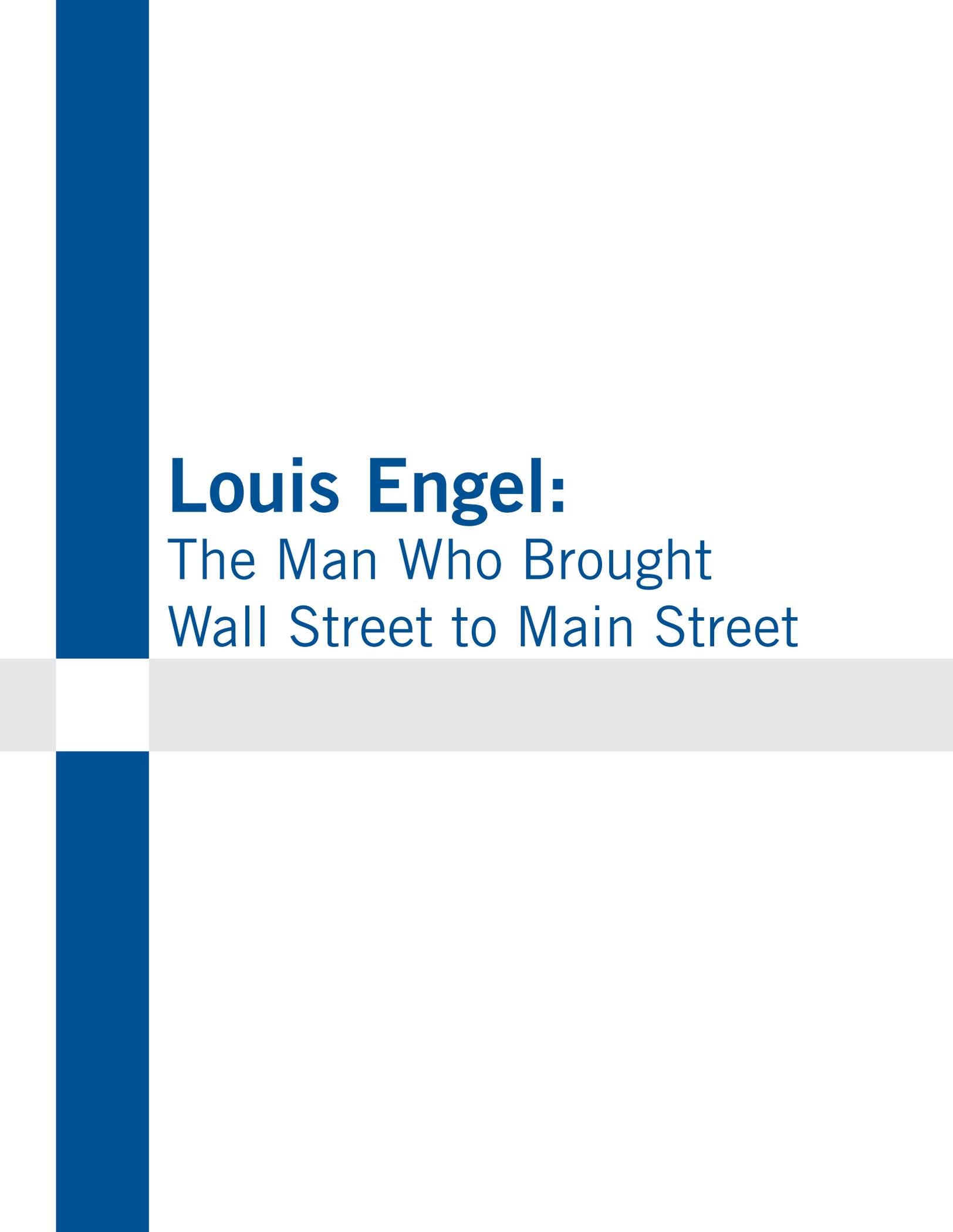
Louis Engel:
The Man Who Brought Wall Street to Main Street

Louis H. Engel, Jr. was born on November 27, 1909 in Jacksonville, IL. He attended the University of Chicago and was graduated in the spring of 1930 with a Ph.B., or Bachelors of Philosophy Degree. While at Chicago, Engel served as the Managing Editor, and later as Chairman of the Editorial Board, for the school's newspaper, the Daily Maroon. He was also active in several other organizations where he honed his public relations and writing skills.

From 1930 to 1932, Engel worked as a staff member

at the University of Chicago Press. In 1932, he moved to New York to take a $35.00 a week job writing for the magazine, *Advertising and Selling*. In 1933, Engel was appointed

Managing Editor. He joined *Business Week* as a reporter in 1934 and in 1936, at age 27, he was made the Managing Editor. In 1940, family friend Ted Braun tried to recruit Engel to the newly merged firms of Merrill Lynch and E.A. Pierce. Braun, who worked as a consultant on the Merrill deal, believed that Charlie Merrill would be putting greater emphasis on marketing and advertising in the combined firms and thought Engel to be a good fit. Engel declined Braun's invitation and continued on at *Business Week*. In 1946, management changes at *Business Week* forced Engel to seek new opportunities. That summer, Engel interviewed with Winthrop H. Smith, a founding partner at Merrill, but ultimately decided to pass on the opportunity presented to him. Following that decision, Engel signed on as Editor and correspondent for the newly launched Kiplinger publication – *Changing Times*. Within weeks of starting his new position, Engel determined that his new role at Kiplinger was not going to work for him. In the fall of 1946, Engel called Win Smith at Merrill to inquire whether the Advertising and Sales Promotion Manager position was still available. During that phone call, Engel accepted an offer from Smith and he joined Merrill on November 15, 1946.

Edwin J. Perkins, author of the book *Wall Street to Main Street: Charles Merrill and Middle Class Investors* provides this excerpt on Engel's influence on Merrill's promotional efforts.

> Engel designed the type of aggressive advertising program that Charlie had been advocating for years. Imaginatively using his own naiveté to create more genuinely informative advertisements, Engel wrote purposefully unsophisticated copy that addressed the questions and concerns of persons geographically and intellectually distant from Wall Street. He presumed that most readers glancing at a brokerage ad knew as little – or even less – about stocks and bonds as he himself had known prior to joining Merrill Lynch. In his copy he tried to explain investing principles in plain and simple language.
>
> Engel's unorthodox ideas and his desire for creative independence soon produced conflicts with George Hyslop, his immediate superior. Hyslop, a partner

inherited from the Pierce organization, had been assigned primary responsibility for advertising and public relations soon after the merger in 1940. Engel found he could not make a move without checking every little detail with Hyslop, and the policy irritated him to no end. "He insisted on seeing every little two-by-two ad even if it was for hog bellies in the National Butcher magazine," Engel jokingly recalled. Charlie and Smith soon realized that Hyslop's managerial philosophy had become an obstacle to progress, and he was eased out of the partnership. After Robert Magowan assumed most of Hyslop's duties, the advertising department functioned more harmoniously, since Magowan granted Engel far greater latitude in decision making. Magowan was also a supportive ally in internal conflicts with other departments.

One technique that Engel favored was inviting potential customers to respond promptly through the mail – either by writing a personal letter or by filling out a coupon requesting free and objective information. At one point he told the partners: "We have a research department to analyze people's portfolios and you never advertise it;…it's the best thing you've got to see, for God's sake." Given the green light, Engel launched a major campaign to inform the public about the availability of portfolio analysts on the Merrill Lynch staff. When magazine and newspaper readers responded eagerly to ads inviting them to mail a list of their holdings to the New York office for analysis, the personnel in the research department complained bitterly about the extra workload. An overworked staff argued that Engel's initiatives had gone too far – another case of "too much of a good thing." They pressured Magowan to force Engel to scale back the placement of ads proclaiming their services – services for which neither they nor the partners were receiving any direct monetary compensation.

With Charlie's blessing, Magowan backed Engel to the hilt, and the ads ran on schedule. Magowan told the research department that coping with the heavy inflow of mail was an internal problem that required an internal solution. The advertising department, he reminded the complainers, was producing remarkable results in generating thousands of solid leads for Merrill Lynch brokers nationwide. Engel had demonstrated beyond all doubt that, with the proper advertising, a brokerage firm could attract the attention of thousands of households that were interested in planning for their financial security – households that were expressing a new willingness to consider something other than whole-life insurance policies with modest rates of return. When the onslaught of letters continued, the partners authorized the head of the research department to add more staff to handle the deluge. Charlie and Smith concluded that the maintenance of a large research department was another inescapable high cost of doing business if they wanted to provide reliable services on a mass scale in a thoroughly professional manner.

Engel's most innovative advertisement, one that often turns up on lists of the one hundred most influential ads in the nation's history, appeared in the fall of 1948, approximately two years after he had joined Merrill Lynch. Titled "What Everybody Ought to Know about This Stock and Bond Business," the ad consisted of six thousand words of very small print squeezed onto a full-size newspaper page.

The copy was informational and educational – and textbook dry in tone. There were no explicit references to the firm's own brokerage services in the entire text, but at the bottom right of the page was a small calling card that identified Merrill Lynch as the sponsor and invited readers to request free reprints of the ad in pamphlet form. In the history of print media, no single advertisement with so much seemingly boring copy had ever been published for any product or service.

At first, Charlie, Smith, and Magowan were only lukewarm about this questionable marketing concept. At five thousand dollars, the cost of running the full-page ad in the New York Times was extremely high relative to the size of the advertising budget (2 percent

of the annual appropriation, to be exact). How many people, the partners wondered, would actually make the effort to read the text and how many, in turn, would respond by asking for reprints of a text that could easily be clipped right out of the newspaper? The partners also debated whether to spend so much money for an essentially generic advertisement that was likely to benefit rival firms. Despite his superiors' reservations, Engel insisted on gambling a portion of his annual budget on the innovative concept.

A 6,450 word ad written by Engel as published in the New York Times on October 19, 1948.

Finally, a compromise was reached. The partners agreed to allow Engel to run a trial advertisement in the Cleveland Plain Dealer, where the cost of space was much lower than in New York. If the ad bombed, they could drop the whole idea without wasting another nickel of the advertising budget. When the public response to the Cleveland experiment proved encouraging, Engel received permission to test the ad in the New York Times. During the week or so after publication, the firm received more than five thousand requests for pamphlet reprints. "What was most amazing," Engel recalled, "was that we got hundreds and hundreds of long and thoughtful letters." Some

respondents were profusely appreciative. One person wrote: "God bless Merrill Lynch;…I have been wanting to know this all my life;…I owned stocks and bonds and I never really knew what I owned." The firm ran the same advertisement, or slightly revised versions, in newspapers across the country, not only during the next few months, but, indeed, for years thereafter. The total number of responses exceeded three million, and those returns translated into millions of prospective customers for the firm's eager brokers. With that one concept alone, Engel proved himself a promotional genius. His subsequent aggressive campaigns, which were typically both educational and eye-catching, set new standards for brokerage firms and other enterprises in the financial services sector.

As a result of the highly visible Merrill ad campaign, Engel was approached by the publisher Little, Brown and Co. to write a book on the topic. Six weeks later Engel finished the manuscript for "How to Buy Stocks." The first edition of the book was published in May 1953. It became a best-selling investment primer that was regularly revised in new editions for the next several decades. The eighth edition of the book was published in 1994 by Engel's co-author and Merrill colleague, Henry Hecht. In total, the book has sold over 7 million copies across its eight editions.

In 1959 Engel had a "silly idea." He wondered what the chances of making money would be if you threw a dart at the list of New York Stock Exchange listed companies and you bought the issue the dart hit. He approached Columbia University School of Business and asked if such a calculation could be made. "The school replied, in essence, Great idea. But we only have one lifetime."

Engel then turned to his alma mater, the University of Chicago, and in March 1959 he called Jim Lorie, then Professor of Finance at the Graduate School of Business (now Booth School of Business). The underlying question raised by Engel was, relative to other asset classes, how have the returns of stocks performed over the long term. In an early article based on the Merrill ads, Engel had written about and compared the returns of certain stocks relative to the returns of savings accounts and life

insurance investments. What Engel sought to quantify was the average return of the market "over the long haul."

Professor Lorie was intrigued by Engel's question from an academic perspective and he proposed that Merrill fund a study to compile the historical data needed to calculate the returns. The Center for Research in Security Prices at the University of Chicago was initially funded by a $50,000 grant from Merrill. The project was publicly announced in August 1960. It was initially projected by the CRSP team that it would take 6 months to research, compile and cleanse the data. The actual time required to complete the database was 3-1/2 years and the research grants received by CRSP from Merrill during this period were approximately $150,000.

In December 1963, the landmark results of the 3-1/2 year project were announced. Fisher and Lorie published the results in the January 1964 issue of *The Journal of Business*. In keeping with his earlier full page ad promotion, Engel decided to publish the entire Rates of Return on Investments in Common Stock paper in the Wall Street Journal. Like his 1948 "what everybody ought to know.." campaign, this ad also caught the attention of the investing public, as well as investment professionals and academic researchers.

Merrill continued to promote the Fisher/Lorie research results by offering free copies of the article for the next several years. Merrill also funded other smaller research projects at CRSP following the development of the master data files in 1964.

Louis Engel retired from the firm as Vice President and Partner in 1969. He died in 1982 at the age of 72. You can read more about Louis Engel in One Hundred Minds that Made the Market by Kenneth L. Fisher.

Reprint of the "Rates of Return on Investments in Common Stocks"
Wall Street Journal June 2, 1964

The computer first used in creating the CRSP database – UNIVAC 1 - System #42 – was delivered to the school in November 1957. System #42 was the next to last commercial installation of the UNIVAC 1.

Lawrence Fisher:

Created the Benchmark for Research-Quality Data

Lawrence Fisher was born in Los Angeles, CA on October 19, 1929. He attended Pomona College and earned a BA in Economics in 1951. He later attended the University of Chicago earning an AM in 1955 and a Ph.D. in 1956, both in Economics.

Fisher served as Assistant Professor, Associate Professor and Professor of Finance at the University of Chicago, Graduate School of Business (now Booth School of Business) from 1957 until 1978. He also served as the Associate Director of the Center for Research in Security Prices (CRSP) from 1960 until 1978. He then joined Rutgers University Business School, Faculty of Management, Department of Finance and Economics. In 1990, he was named the First Fidelity Bank Research Professor of Finance. He was professor emeritus at the time of his death in 2008 at the age of 78.

Professor Fisher's research interests were in problems related to the behavior of financial markets, including the measurement of risk and return (e.g., indexes, betas, computer algorithms, and construction of data bases); portfolio theory; option theory; market efficiency; duration, risk, and immunization of fixed-income securities; information; capital structure; and effects of taxes and transaction costs. He served as author, co-author and contributor to many research papers and books on these topics.

William L. Fouse, indexing pioneer and co-founder of Mellon Capital described the role of Lawrence Fisher in the CRSP project in the quote below:

In September 1963, Professor Fisher spoke at the Annual Meeting of the American Statistical Association held in Cleveland, Ohio. His presentation to the Business and Economic Statistics Section was titled *Use of Electronic Computers in the Quality of Financial Data*. The presentation discusses the many challenges and the novel methods used to develop the CRSP stock file, one of the first research--quality financial databases. The full the text of the speech follows on the next page.

CRSP thanks Mrs. Lois A. Fisher for allowing us to reprint and share this paper.

> " With respect to the Center for Research in Security Prices, Jim Lorie might be viewed as the maître d'. But Larry Fisher was the chef. He deserves the credit for the actual work: the production, specificity, accuracy – and cleanliness, if you will – of the data."

USE OF ELECTRONIC COMPUTERS IN THE QUALITY CONTROL OF FINANCIAL DATA*

Lawrence Fisher, University of Chicago

At the Center for Research in Security Prices at the University of Chicago we have collected a file of information on common stocks listed on the New York Stock Exchange. At the present time this file covers all such common stocks for the periods in which they were listed during the years 1926 through 1960. It contains information on about 1700 stocks.

For each stock we have the price at the end of each month and all other information needed to find the consequences of making an investment in a particular stock at one date and holding that stock until some later date. In addition to price quotations - we have 390,000 of them - we collected data on 90,000 cash dividends and on 10,000 other events - events such as distributions of subscription rights, either to more shares of the same stock, some other stock such as preferred stock, some other kind of security such as a convertible bond, or something else such as a case of Old Guckenheimer Whiskey. Other events are stock dividends, splits, and reverse splits; spin-offs; reorganizations; mergers; exchange offers; and changes of corporate name.

Computations of rates of return have been made before. In our collection of data we have had two objectives: to collect a library which could also be used for other kinds of studies and to make rate-of-return calculations which could, in fact, provide benchmarks against which the efficiency of various methods of making investment decisions could be measured.

The investor in stock traded on a stock exchange must pay broker's commissions whenever he buys or sells. He must pay income taxes when he receives dividends or realizes capital gains. Therefore each record of the price of a share of stock also includes the round-lot commission to be added to or subtracted from the price. Therefore, for example, a cash dividend is not just a cash dividend but, according to the way it was taxed, one of seven kinds of cash dividend.

The information we have is stored on magnetic tape. It is arranged by security and chronologically within the security.

Late last winter when I agreed to be here today, we fully expected to have published our benchmark studies before summer, and I expected to report to you some additional information. Indeed, if the accuracy of the data we had in our files had been sufficiently good, I would be doing that now. But, even though we intend to estimate

---

average rates of return to investment in common stocks listed on the New York Stock Exchange by taking the largest sample which is feasible, the whole population of common stocks listed on that exchange, to have computed such an estimate without at least all of our applications in quality control would have given us a large chance of making a large error.

Today I shall describe some of our procedures for quality control. So that these procedures will appear to be meaningful, let us first consider our sources of data and our methods of collecting these data.

Each item in our file describes either the price for a month or one other event, such as the payment of one dividend, one distribution of rights, etc.

Financial information about most companies listed on the New York Stock Exchange is readily available. So readily available, in fact, that one could barely read all of it. Our first problem was to find sources of information which could be consulted periodically. We found that, for our purposes, the most convenient sources of information were The Bank and Quotation Record for prices, Moody's and Standard and Poor's Annual Dividend Guides for most cash dividends, and Prentice-Hall's Capital Adjustments and the Commerce Clearing House Capital Changes Reporter for other events.

Price Information

In order to record prices, we obtained a list of stocks traded on the New York Stock Exchange at the time we began the collection of data in mid-1960 and consulted the records of the stock exchange for listings and delistings between 1926 and 1960. For each month that a company was listed, we prepared a prepunched card for use as a coding form. This coding form contained the name of the company, the date, and two numbers - a "company number" which referenced our information on listing and delisting, and an "alpha number" to aid in the alphabetizing of the cards.

We got office space, hired clerks, checked The Bank and Quotation Record out of our library and set the clerks to work copying numbers from The Bank and Quotation Record. We knew that the copying of the numbers would not always be accurate and that our keypunch operator would sometimes make mistakes. But we also knew that in printing approximately 400,000 prices the publishers of The Bank and Quotation Record would also make some mistakes.

Rather than coding and punching all prices twice and then resolving discrepancies manually we found a better procedure. We know that the change in the price of a stock during one month

is very nearly independent of its change during the next month. Therefore if a price changes a large amount from one date to a second date and by a similar amount in the opposite direction from the second date to a third, there is reason to believe that at the second date, the price was misrecorded. A "large change" was rather arbitrarily taken to mean a change in magnitude of more than 10 per cent of the previous price plus a dollar.

To see whether this method of finding errors would be successful and to test the accuracy of the original coding and punching of cards and their recording on magnetic tape, a random sample of 100 clusters of 50 prices each was coded a second time, punched, and recorded as first recorded. This test revealed 132 errors in price in our original data collection. Of these errors, 72 were caused by failure to find any price for the stock that month or were prices which were invalid on their face because the bid price was higher than the ask or because the fractional part of the price was impossible, i.e., 7/7 and 3/1. The remaining 58 erroneous prices had face validity. Of these 58, 30 were in error by more than 10 per cent plus a dollar and 28 were not. Of these 28, 14 were too high and 14 were too low. The average magnitude of the error was 2½ per cent of the price and the mean error was -3/4 per cent. Thus the check we planned appeared to be satisfactory.

Computer programs were written which among other things, checked the validity of the fractions (before converting them to decimals), made sure wherever both bid and asked quotations (rather than sales prices) appeared that the bid was less than the ask; looked for missing prices quotations, and finally made the comparison of consecutive prices described.

Dividends

In collecting prices we could reasonably expect to find approximately one price for each month a security was listed. But in collecting data on cash dividends there was no way to predict the frequency of dividends for each company.

Annual dividend guides that list publicly held companies in alphabetical order and that describe each dividend paid during the year are available for the period beginning in 1937. For earlier periods, quarterly guides are available.

To collect the data, clerks were given cards with a coding form printed on them, a list of names and code numbers of listed companies, and a dividend guide. They filled out as many cards as there were cash dividends for listed companies. This information was then punched into the cards and the data were transcribed onto magnetic tape. For the last years of the study, the annual guides note the exchanges on which a stock was listed. For the earlier years they do not. Because it was so easy to make clerical errors, our method of collection then could not be expected to produce a very complete list of dividends.

To check on the dividends, we turned to Moody's Manuals, which show annual dividends per share for each security described. We recorded these totals to nearest cent, punched them into cards, and placed them on magnetic tape. A computer program was written which compared the sum of each company's dividends for a year, found by adding dividends copied from a dividend guide with the total for the year as reported in Moody's Manual. Whenever a discrepancy was found, a report was printed. This report showed the individual dividends in question, their total and the discrepancy.

The appropriate dividend guide or manual or both were then consulted to resolve the discrepancy, and the error in the file of dividends or annual totals corrected.

This process was repeated several times until there were no more discrepancies.

Other Events

The other events in our files are usually called capital changes. We coded these changes from Commerce Clearing House Capital Changes. These reports are prepared primarily for use by accountants in filling out income tax returns. They are in loose-leaf form so that complete information for a company is in one or at most two volumes. However, information about companies liquidated or reorganized many years ago is frequently omitted. Where no information was shown for a company before the time at which we are first interested in it, we checked contemporary sources.

Since there is a large variety of capital changes, most such changes had to be punched into two cards in order to obtain a standard, legible format. These cards were listed and this printed copy was compared with the Capital Changes Reports. After errors were corrected, the cards were placed on magnetic tape, using a somewhat different format. To this file we added an over-the-counter price for securities which had been delisted.

As a further check, a computer program was written to take the coded information on each of these stock dividends, splits, rights, mergers, etc. and decode it to form a verbal description. The verbal descriptions were then compared with the Capital Changes Reports. As a result of this comparison approximately 2,000 errors were found and corrected. Most of these errors had been caused by failure to keep the two cards on which the event had been recorded together or by failure in some of the computer programs which had been used to manipulate the data.

We then had three files of information, each of which came from a different primary source. These files were in reasonably good condition. They were then merged into a single file.

Checking the Combined File

But further checks were necessary. For

is very nearly independent of its change during the next month. Therefore if a price changes a large amount from one date to a second date and by a similar amount in the opposite direction from the second date to a third, there is reason to believe that at the second date, the price was misrecorded. A "large change" was rather arbitrarily taken to mean a change in magnitude of more than 10 per cent of the previous price plus a dollar.

To see whether this method of finding errors would be successful and to test the accuracy of the original coding and punching of cards and their recording on magnetic tape, a random sample of 100 clusters of 50 prices each was coded a second time, punched, and recorded as first recorded. This test revealed 132 errors in price in our original data collection. Of these errors, 72 were caused by failure to find any price for the stock that month or were prices which were invalid on their face because the bid price was higher than the ask or because the fractional part of the price was impossible, i.e., 7/7 and 3/1. The remaining 58 erroneous prices had face validity. Of these 58, 30 were in error by more than 10 per cent plus a dollar and 28 were not. Of these 28, 14 were too high and 14 were too low. The average magnitude of the error was 2½ per cent of the price and the mean error was -3/4 per cent. Thus the check we planned appeared to be satisfactory.

Computer programs were written which among other things, checked the validity of the fractions (before converting them to decimals), made sure wherever both bid and asked quotations (rather than sales prices) appeared that the bid was less than the ask; looked for missing prices quotations, and finally made the comparison of consecutive prices described.

Dividends

In collecting prices we could reasonably expect to find approximately one price for each month a security was listed. But in collecting data on cash dividends there was no way to predict the frequency of dividends for each company.

Annual dividend guides that list publicly held companies in alphabetical order and that describe each dividend paid during the year are available for the period beginning in 1937. For earlier periods, quarterly guides are available.

To collect the data, clerks were given cards with a coding form printed on them, a list of names and code numbers of listed companies, and a dividend guide. They filled out as many cards as there were cash dividends for listed companies. This information was then punched into the cards and the data were transcribed onto magnetic tape. For the last years of the study, the annual guides note the exchanges on which a stock was listed. For the earlier years they do not. Because it was so easy to make clerical errors, our method of collection then could not be expected to produce a very complete list of dividends.

To check on the dividends, we turned to Moody's _Manuals_, which show annual dividends per share for each security described. We recorded these totals to nearest cent, punched them into cards, and placed them on magnetic tape. A computer program was written which compared the sum of each company's dividends for a year, found by adding dividends copied from a dividend guide with the total for the year as reported in Moody's _Manual_. Whenever a discrepancy was found, a report was printed. This report showed the individual dividends in question, their total and the discrepancy.

The appropriate dividend guide or manual or both were then consulted to resolve the discrepancy, and the error in the file of dividends or annual totals corrected.

This process was repeated several times until there were no more discrepancies.

Other Events

The other events in our files are usually called capital changes. We coded these changes from Commerce Clearing House _Capital Changes_. These reports are prepared primarily for use by accountants in filling out income tax returns. They are in loose-leaf form so that complete information for a company is in one or at most two volumes. However, information about companies liquidated or reorganized many years ago is frequently omitted. Where no information was shown for a company before the time at which we are first interested in it, we checked contemporary sources.

Since there is a large variety of capital changes, most such changes had to be punched into two cards in order to obtain a standard, legible format. These cards were listed and this printed copy was compared with the _Capital Changes Reports_. After errors were corrected, the cards were placed on magnetic tape, using a somewhat different format. To this file we added an over-the-counter price for securities which had been delisted.

As a further check, a computer program was written to take the coded information on each of these stock dividends, splits, rights, mergers, etc. and decode it to form a verbal description. The verbal descriptions were then compared with the _Capital Changes Reports_. As a result of this comparison approximately 2,000 errors were found and corrected. Most of these errors had been caused by failure to keep the two cards on which the event had been recorded together or by failure in some of the computer programs which had been used to manipulate the data.

We then had three files of information, each of which came from a different primary source. These files were in reasonably good condition. They were then merged into a single file.

Checking the Combined File

But further checks were necessary. For

example, when a dividend is about to be paid or some other distribution is about to be made, the stock usually begins to trade without the dividend or property to be distributed (goes ex dividend or ex distribution) several days before the record date, which in turn is weeks or months before the actual date of payment. When the dividend or distribution is large, however, a special rule is frequently adopted by the stock exchange. The ex date may not be until after the actual distribution has been made, and the delay may be of any length. We found one case where a stock was listed until 15 months after a split but old stock rather than new was traded until the company was delisted. Hence it was necessary to check the ex dates of such distributions as inferred from the record or distribution date with the price series. Furthermore, when a capital change is made, both old and new stock may be traded in simultaneously. Where we had more than one price, the appropriate one had to be chosen. A further source of error lay in the methods used in checking capital changes.

I wrote still another computer program for checking the consistency of the file. In addition to raising questions about "large" successive price changes of opposite sign, it reports price changes of still larger magnitude that are not explained by dividends or other events and instances where a capital change implied a large change in price but none was shown. It also reports inconsistencies in the organization of the file. This program has now been run twice and the data are essentially ready for use in making estimates of rates of return. Many suspicious entries, however, have been flagged but final decisions about them have yet not been made. For about 10 per cent of the 1700 stocks in the file all discrepancies in the file have been resolved. In resolving these discrepancies we have found that a substantial number of them were caused by errors in our sources of data.

By use of these methods of quality control, we expect to have a file of data which is not only in convenient form, but which for most purposes is at least as accurate as the sources from which it was drawn.

James Lorie:

Recognized the Importance of CRSP for Future Research
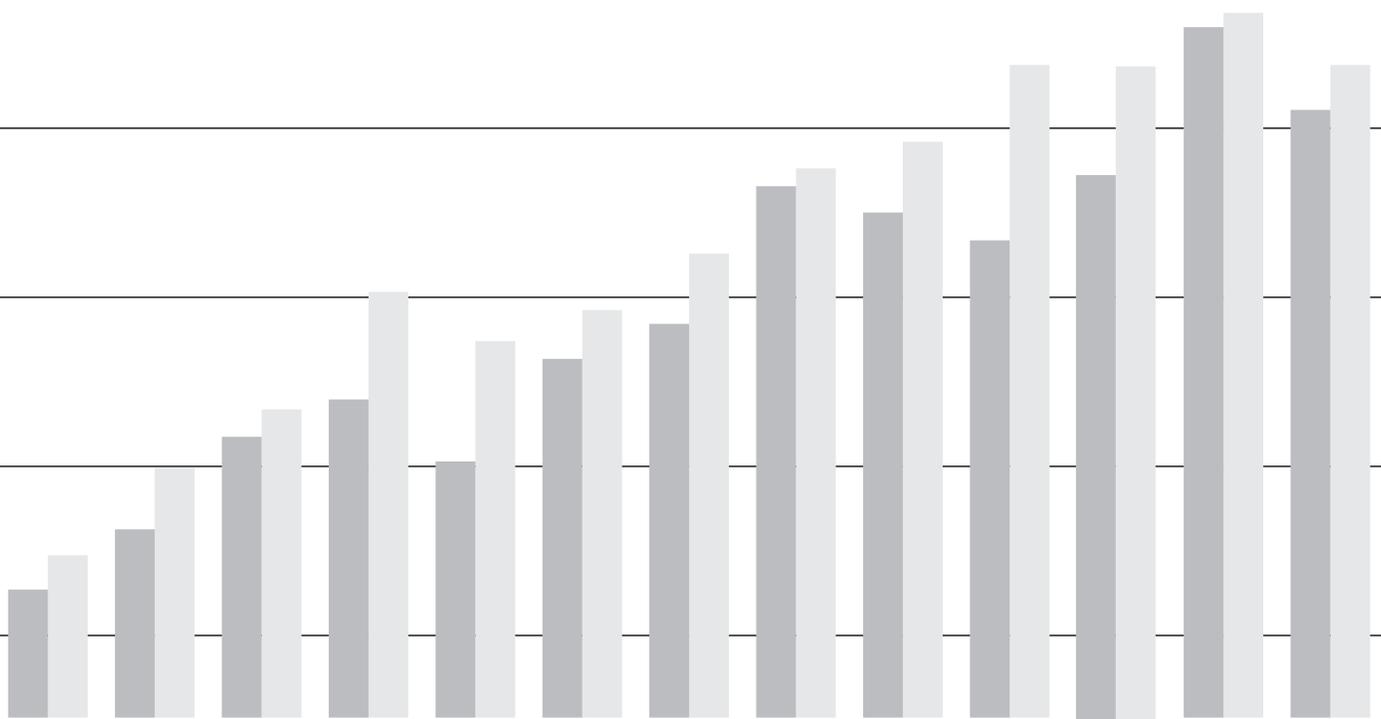
James Hirsch Lorie was born on February 23, 1922 in Kansas City, Missouri. He received his A.B. in 1942 and an A.M. in 1945, both from Cornell University. Professor Lorie earned his Ph.D. from the University of Chicago in 1947.

Lorie was a research assistant at Cornell University (1944-1945) and a member at a staff seminar on American Civilization in Salzburg, Austria (1947). He joined the faculty of the University Of Chicago Graduate School of Business (now Booth School of Business) the same year he received his Ph.D. He was the Eli B. and Harriet B. Williams Professor of Business Administration and also served as the Associate Dean of the Business School from 1956 to 1961. As an associate dean, Lorie, along with Dean W. Allen Wallis, helped introduce the Chicago Approach to business management education. The approach incorporates sociology and anthropology and other disciplines to provide a framework for understanding complex marketplace dynamics.

In addition to his teaching duties, which spanned from 1947 to 1992, he served as a consultant for the Division of Research and Statistics, Board of Governors, of the Federal Reserve System (1950-1952); and as consultant for the U.S. Treasury Department. He served as a director on several boards, including Square D. Company, Acorn

Fund of New York, Fundamerica of Japan, Vulcan Materials Company, Merrill Lynch & Co., and Sealy. Lorie was also a founding director of the Chicago Board Options Exchange, and a Director of the National Association of Securities Dealers.

Lorie's publications include *Basic Methods of Marketing Research* (1951), *A Half Century of Returns on Stocks and Bonds* (with Lawrence Fisher, University Press, 1977), *Modern Developments in Investment Management: A Book of Readings* (1978), and *The Stock Market: Theories and Evidence* (1985).

Lorie founded the Center for Research in Security Prices (CRSP) at the Graduate School of Business in 1960, and served as its director until 1975. "He shepherded the Center through its difficult early years and was rewarded by seeing it produce an explosion of empirical work that is a large part of the knowledge base of modern finance," said Eugene Fama, the Robert R. McCormick Distinguished Service Professor of Finance at Chicago Booth. Professor Lorie died in 2005 at age 83.

In September 1965, Professor Lorie spoke at the annual meeting of the American Statistical Association. His presentation touched on the results of the CRSP rates of return research, market behaviors and on the controversies of the period, including randomness and active management. The presentation, known as the "Philadelphia Talk", has been reprinted below.

Selected Papers No. 20:
Current Controversies on the Stock Market

By
James H. Lorie
Graduate School of Business
University of Chicago

James H. Lorie is Professor of Business Administration, Director of the Center for Research in Security Prices (sponsored by Merrill Lynch, Pierce, Fenner and Smith Inc.), and Director of Research, Graduate School of Business, University of Chicago. He received the A.B. and A.M. degrees from Cornell University and the Ph.D. degree from the University of Chicago. He has been on the faculty

of the Graduate School of Business since 1947, and was for several years Associate Dean of the School. Professor Lorie served as Consultant to the Board of Governors of the Federal Reserve System and as Senior Consultant for Joel Dean Associates, a management consulting firm. He is a Director of Gaertner Scientific Corporation, Standard Shares, Inc., Wallace Business Forms, and Hyde Park Federal Savings and Loan Association; and a member of the Board of Trustees of the Foundation for Economic Education, Inc. Professor Lorie has written and spoken widely about marketing, consumer spending, and business finance. At the beginning of 1964, Professors Lorie and Lawrence Fisher published the first detailed and comprehensive study of rates of return on all common stocks listed on the New York Stock Exchange. This report, based on the work of the Center for Research in Security Prices, received national and international attention. In a talk delivered by Professor Lorie at the annual meeting of the American Statistical Association in Philadelphia on September 8, 1965, he described the findings of this study and of related studies of the behavior of stock prices. This Selected Paper is based upon Professor Lorie's Philadelphia talk.

Current Controversies on the Stock Market

I am very pleased to be here this evening, but I must confess to some surprise at having been invited. I am not a statistician, nor am I rich enough or influential enough to be the object of flattery by the head of a large, privately supported university. In reflecting on possible reasons for Allen Wallis' invitation, I concluded that it sprang from a sense of brotherhood that he and I feel as members of the tiny group whose statistical training has been profoundly influenced by Mark Twain. Twain is perhaps better known as a novelist and humorist than as a statistician, but Allen and I and a few others know that a perceptive reading of his writings reveals a foreshadowing of the work of such diverse statisticians as Fisher, David Wallace and Mosteller, Hansen, Hurwitz, and Alfred Cowles. (As you can see, I'm gradually converging on the stock market.) In some early work on experimental design, Twain noted that a cat which had once jumped on a hot stove never

jumped on a hot stove again-or a cold stove either, for that matter. Fisher, of course, with the benefit of more powerful analytical tools, would have jumped on a cold stove. Twain wrote a fascinating book with the simple title, Christian Science. In it he tries to understand Mary Baker Eddy and the theory and mechanism of the Christian Science Church. He also tries to determine the authorship of the *Key to the Scriptures*. By comparing word patterns and vocabulary in that work with those in other writings known to be by Mrs. Eddy, Twain persuasively concluded that she did not write the *Key to the Scriptures*. By similar means, he concluded that the plays generally attributed to Shakespeare were really by another man of the same name. These early efforts of Twain, though not so costly as later work on the Federalist Papers, were more entertaining. Twain also worked as a demographer. He discovered through imaginative analysis of mortality statistics that it's impossible to live to be eleven, but that those beyond eleven are immortal. This seems obvious to us now-for instance, all of us here are over eleven or else not yet eleven-but in Twain's day his finding was received with incredulity. Finally, I would like to comment on Twain's work on the stock market as a transition to the rest of my talk which is on more recent efforts in that field. Twain's work was elaborate and is worth your attention, but I shall mention here only his major conclusion: April is a dangerous month in which to speculate in Stocks; other dangerous months are October, June, March, November, January, August, February, May, December, September, and July.

Fluctuations Matter

Some people say that sex is not as important as Freud thought; and as I get older, I am increasingly inclined to agree with them. Others deny that money is as important as the Socialists say. They may be right. Nevertheless, sex and money are undoubtedly both popular and even important. Twenty million Americans and their families own about 600 billion dollars' worth of stock, and fluctuations in the value of stocks matter. I am going to talk about fluctuations in the value of the most important group of stocks in the world, those listed on the New York Stock Exchange. There has been a very large amount of

statistical work on stock prices-as one would expect when the relevant data are so freely available and the prizes for original, correct work are so large, tangible, negotiable, and automatically bestowed. Until recently almost all of this work was by persons who knew a great deal about the stock market and very little about statistics. While this combination of knowledge and ignorance is not so likely to be sterile as the reverse-that is, statistical sophistication coupled with ignorance of the field of application-it nevertheless failed to produce much of value. The major enduring empirical work before World War II was Alfred Cowles' study of the rates of return on a substantial group of stocks for the period 1871-1940.

Rates of Return

Recently, scientific quantitative research has become much more voluminous and new results of importance have emerged, though some of the most interesting are still controversial. This upsurge of scientific labor has been facilitated by the availability of high speed computers and by the creation of two large files of tape of basic information on stocks. The first file to be completed and used-and the one with which I shall deal primarily in this talk-is of stock prices and it was created by the Center for Research in Security Prices of the Graduate School of Business of the University of Chicago. The second file is called "Compustat" and contains about 60 kinds of information found on corporate balance sheets and income statements. The data are available for about 1,000 firms for about 15 years. Compustat tapes are sold by the Standard Statistics Corporation and have been given to a number of universities. The Center for Research in Security Prices hereafter referred to as CRISP-was started in March, 1960, by a grant of $50,000 from Merrill Lynch, Pierce, Fenner and Smith Inc., to the University of Chicago in order to answer a basic question-what has been the average rate of return on investments in common stocks? We hoped to answer this question better than it had been answered before for $50,000 in one year. We spent $250,000 and took four years. Our optimistic naiveté may have been inexcusable but it can be explained, and the explanation will hopefully be of some interest. The results which I will discuss later have attracted an almost incredible amount of attention with unknown but probably substantial

practical consequences. We decided to deal with all the common stocks on the New York Stock Exchange. These stocks account for over 85 per cent of the value of all common stock outstanding in this country and the data on New York Stock Exchange stocks are relatively complete and accurate. This exchange is by far the largest in the world, its listed securities being worth over four times those of the second largest exchange -London.

All Stocks Included

A sample of these stocks would have been adequate for many though not all reasonable purposes-for example, a study of optimum industry groupings in the construction of index numbers-but, curiously, we concluded that it would be more costly to achieve a satisfactory level of accuracy for an adequate sample than for the entire population. Experience in a pilot study indicated a sharp rise in the incidence of clerical error if a sample of stocks were selected from available comprehensive lists. Further, some efficient methods of quality control of the clerical processes would not be available if a sample were used. We recorded monthly closing prices of these stocks for the 35 years beginning in January, 1926. We dealt with about 1,700 stocks and recorded all information necessary to compute rates of return. This information is voluminous and complex. It includes data on 39 different types of distributions of cash and property to shareholders-e.g., shares of stock, rights to buy stock, warehouse receipts for whiskey-the dates of distribution and the tax status. Each dividend, for example, fell in one or more of seven different tax categories. Information was required on mergers, spin-offs, exchange offers, commission rates on the purchase and sale of shares, tax rates on income and capital gains for individuals with different incomes, name changes, etc. Although the interpretation and coding of much of this information clearly required highly trained personnel, it was our original belief that the raw prices themselves could be recorded adequately by untrained clerks. Even this hope proved unfounded. The main difficulty was in deciding what was a common stock. We generalized from the work of Gertrude Stein, who, you may recall, said that a rose is a rose is a rose. We thought that a common stock is a common stock is a common stock; but it isn't. Further, some things not called common stocks are. Securities with over 50 different types of designations

proved to be common stocks-e.g., American Depository Receipts, certificates old, certificates new, certificates black, certificates blue, preferred stock, and even debentures. That is, securities with such designations were residual claimants to the income of corporations and were therefore, by our definition, common stocks. On the other hand, the common stock of the Green Bay and Western Railroad, for example, had preferential rights to corporate income and was therefore not a common stock. Of our almost 400,000 price quotations, over 30,000 required more than clerical attention.

Refinement, Accuracy

The man largely responsible for the work of CRISP, Lawrence Fisher, was fanatical in his desire for refinement and accuracy in measuring the rate of return on investments in common stocks, and relatively indifferent to the rate of return on the investment in making the measurement. He aspired to make the data on our tapes more accurate than the sources from which they came-a possibly laudable and assuredly extravagant ambition. Since we feel that he succeeded, our methods of quality control may be of interest. My account is taken in large part from a paper of Fisher's, "Use of Electronic Computers in the Quality Control of Financial Data." Two principles proved useful: (1) Recording data as found in the sources without adjustment; (2) Using the computer to identify "suspicious," inconsistent, or impossible items. In accord with these principles, the following procedures were used. For each month that a company was listed, we prepared a prepunched card for use as a coding form. This coding form contained the name of the company, the date, and two numbers-a "company number" which referenced our information on listing and delisting, and an "alpha number" to aid in the alphabetizing of the cards. Rather than coding and punching all prices twice and then resolving discrepancies manually, we found a better procedure. We know that the change in the price of a stock during one month is very nearly independent of its change during the next month. Therefore, if a price changes a large amount from one date to a second date, and by a similar amount in the opposite direction from the second date to a third, there is a reason to believe that at the second date the price was misrecorded. A "large change" was rather arbitrarily taken to mean a change in magnitude of more than 10 per cent of the previous price plus a dollar.

Test Sample

To see whether this method of finding errors would be successful and to test the accuracy of the original coding and punching of cards and their recording on magnetic tape, a random sample of 100 clusters of 50 prices each was coded a second time, punched, and recorded as first recorded. This test revealed 132 errors in price in our original data collection. Of these errors, 72 were caused by failure to find any price for the stock that month, or were prices which were invalid on their face because the bid price was higher than the asked, or because the fractional part of the price was impossible, e.g., 7/7 and 3/l. The remaining 58 erroneous prices had face validity. Of these 58, 30 were in error by more than 10 per cent plus a dollar and 28 were not. Of these 28, 14 were too high and 14 were too low. The average magnitude of the error was 2 1/2 per cent of the price and the mean error was - 3/4 per cent. Thus the check we planned appeared to be satisfactory in that all errors were small and the process was unbiased. Computer programs were written which, among other things, checked the validity of the fractions (before converting them to decimals), made sure wherever both bid and asked quotations (rather than sales prices) appeared that the bid was less than the asked, looked for missing price quotations, and finally made the comparison of consecutive prices described. In collecting prices we could reasonably expect to find approximately one price for each month a security was listed. But in collecting data on cash dividends there was no way to predict the frequency of dividends for each company.

Dividend Guides

Annual dividend guides that list publicly held companies in alphabetical order and that describe each dividend paid during the year are available for the period beginning in 1937. For earlier periods, quarterly guides are available. To collect the data, clerks were given cards with a coding form printed on them, a list of names and code numbers of listed companies, and a dividend guide. They filled out as many cards as there were cash dividends for listed companies. This information was then punched into the cards and the data were transcribed onto magnetic tape. For the last years of the study, the annual guides note the exchanges on which a stock was listed. For the earlier years they do not. Because

it was so easy to make clerical errors, our method of collection could not be expected to produce a very complete list of dividends. To check on the dividends, we turned to *Moody's Manuals*, which show annual dividends per share for each security described. We recorded these totals to the nearest cent, punched them into cards, and placed them on magnetic tape. A computer program was written which compared the sum of each company's dividends for a year, found by adding dividends copied from a dividend guide with the total for the year as reported in *Moody's Manual*. Whenever a discrepancy was found, a report was printed. This report showed the individual dividends in question, their total and the discrepancy. The appropriate dividend guide or manual or both were then consulted to resolve the discrepancy, and the error in the file of dividends or annual totals corrected. This process was repeated several times until there were no more discrepancies.

Capital Changes

The other events in our files are usually called capital changes. Since there are a large variety of capital changes, most such changes had to be punched into two cards in order to obtain a standard, legible format. These cards were listed and this printed copy was compared with the *Capital Changes Reports*. After errors were corrected, the cards were placed on magnetic tape, using a somewhat different format. To this file we added an over-the-counter price for securities which had been delisted. A computer program was written to take the coded information on each of these stock dividends, splits, rights, mergers, etc., and decode it to form a verbal description. The verbal descriptions were then compared with the *Capital Changes Reports*. As a result of this comparison approximately 2,000 errors were found and corrected.

Results of Study

Our results show the rates of return for 22 time periods between 1926 and 1960, with and without reinvestment of dividends, for persons in three different tax brackets, and with and without liquidation of the final portfolio and payment of the capital gains tax. Other time periods and tax brackets could easily and cheaply be added. Assuming equal initial investments in each company with

one or more common stocks listed on the New York Stock Exchange, the rate of return for a tax exempt institution which reinvested dividends for the period 1926-1960 was 9 per cent per year compounded annually. The comparable rate was 7.7 per cent if the investments were made at the height of the bull market in 1929 and the securities were held till the end of 1960. Since 1950, the rates were over 10 per cent. Incidentally, our work also showed that it paid to be tax exempt. If you had been exempt in 1926, an initial investment of $1,000 would have been worth about $20,000 in December, 1960. If you had an income of $50,000 in 1960 and comparable incomes in earlier years and were not tax exempt, your original $1,000 would have grown to only about $11,000. Our results were distributed to over 700,000 individuals, were reprinted in a full page in the Wall Street Journal, and were presented orally to audiences from the financial communities of London, Geneva, New York, Boston, Philadelphia, Chicago, Miami, Dallas, Los Angeles, and San Francisco. Why the interest? For all long periods and most short periods reasonably defined-the rates are higher, often far higher, than for other types of financial investments for which we have data. This disturbed individuals whose savings were in bonds and savings accounts and even seemed to have some impact on the trustees of private pension funds, whose assets exceeded 60 billion dollars, and on state legislatures which in most states have legally prohibited the investment of the assets of state employee pension funds in common stocks. One of our insights which we shared widely was that the cost of providing a given level of benefits is many times greater if assets earn 3 per cent rather than . Also interested and disturbed were managers of mutual funds-assets more than 30 billion dollars-since on the average the returns to investors in such funds were slightly less than from investment in randomly selected portfolios.

Economists Interested

Academic economists were interested because of the persistence over long periods of time in such large differences in rates of return in different financial media. The standard explanation was and is that stocks are riskier than other investments and that higher rates are therefore necessary to induce investment in stocks. That is reasonable and probably true. Merrill Lynch and we were interested in some measure of this

riskiness. Fisher in an article, "Outcomes for 'Random' Investments in Common Stocks Listed on the New York Stock Exchange," provided one measure of risk. He did a relatively simple thing, namely, computed rates of return on an annual basis and compounded annually for all possible combinations of purchases and sales at the ends of months for all common stocks listed on the New York Stock Exchange for the 35 years beginning in January, 1926. A simple stock listed for 420 months can be bought and sold at approximately 88,000 combinations of dates. Fisher calculated a frequency distribution based on about 57,000,000 rates of return. It would have taken him longer than it did, if he had not used a computer. We do not know how much longer, but we have authorized IBM to say that their computer speeded the work. This frequency distribution shows the results of random selection of stocks and of the timing of purchases and sales. The median rate was 9.8 per cent. Seventy-eight per cent of the transactions yielded a positive return, even after allowing for transaction costs. The interquartile range was from approximately 2 per cent to approximately 17 per cent. Over two thirds of the time the rate exceeded 5 per cent. Nearly one-fifth of the time the rate exceeded 20 per cent. Five times out of 100,000, the investor suffered a total loss, and 2 times out of a million, on the average, he earned money at the rate of a trillion per cent per annum-as would result from a stock's rising from r/s to 7/8 in a month. Fisher also calculated frequencies for purchases and sales during the 16 business expansions and the 16 contractions, as defined by the National Bureau of Economic Research, for the period 1926-1960. The major lesson of this exercise is that generally it doesn't pay to try to be clever in timing one's purchases. That is, delaying the purchase of stocks did not on the average result in superior yields, as the improvement from guessing the cycle was about offset by failure to profit from the strong secular rise during the period under study.

Comments on Findings

Some general comments on Fisher's work are in order. The variability of rates of return is much greater for short-term investments than for long-term and, consequently, the probability of gain for long-term investments was much greater than 78 per cent. Further, because the

rates of return are positively skewed, the holding of groups of stocks rather than a single stock at a time would have led, on the average, to a positive return more than 78 per cent of the time and to returns greater than 9.8 per cent per annum more than half the time. A fortiori, holding groups of stocks for long periods of time would have resulted in a relatively small probability of loss and a relatively high probability of gains greater than from alternative investment media. None of this should be interpreted as a recommendation to buy stocks--I have scrupulously avoided prophecy- but it is surprising to me at least that the superior returns from stocks in the past have been associated with such little risk. Keep in mind that banks were known to fail in the 1930's and not all mortgages turned out as well as the lenders hoped. Now let me turn to some closely related matters. The precise measurement of rates of return from all stocks on the New York Stock Exchange-and by implication from randomly selected portfolios-has caused renewed scrutiny of the performance of mutual funds and investment trusts. These organizations are in the business of investing funds, primarily in stocks. As I suggested earlier, returns to investors in such organizations on the average appear to have been slightly less than from direct investment in randomly selected portfolios. How can this result be explained? The managers of the funds controlled by these organizations are competent, responsible professionals whose careers depend in large part on success in selecting securities and in timing their purchase and sale, yet throwing darts at lists of stocks and dates is on the average as satisfactory a method of making investments as is reliance on competent professional judgment. I have said this before and have been incorrectly interpreted as derogating the ability of the managers of such funds. While my remarks do not constitute extravagant praise, they are not an indictment of the competence of individuals or even of the usefulness of mutual funds and investment trusts.

Possible Reasons

To cast light on what may seem to be a paradox, let's seek an explanation of the apparent inability of these funds and trusts to outperform the market. One part of the explanation is that institutions-mutual funds, trusts, pension funds, etc.-themselves are an important

influence on stock prices. Institutions now own over 20 per cent of New York Stock Exchange stocks and the percentage is growing. Clearly, if institutions have important influence on prices and their analysts are of approximately equal ability, however great, the stocks owned by such institutions will behave much as the market as whole and even individual institutions will have difficulty in showing superior performance. Further, such institutions as a matter of policy or law hold widely diversified portfolios. The law requires extensive diversification among issues by mutual funds, and the size of many funds makes even more extensive diversification essential. We have found in varied and extensive work on index numbers that it is difficult to pick a substantial sample of stocks at any time which, on the average, performs much differently from the market as a whole. Work by Benjamin King, Jr., for example, shows that on the average about 50 per cent of the variance in the prices of individual stocks is accounted for by movements in the market as a whole. Fisher has constructed indexes in which each stock receives equal weight-in marked contrast to the Dow-Jones and Standard and Poor's indexes which are heavily dominated by a few large companies-and found that his indexes have long term movements very similar to those of Dow-Jones and Standard and Poor's. Further, except for 1929 when stocks of small companies turn down several months before those of large companies, the cyclical turning points in the various indexes have been virtually identical in time. Thus, competent people competing with other competent people in selecting groups of stocks largely influenced in the same way by the same set of factors have great difficulty in being consistently superior. Before leaving this subject, I wish to tie up three loose ends. I said earlier that returns to investors in mutual funds would on the average probably have been slightly less than returns from investment in randomly selected portfolios. Why less? There are three reasons. First, such funds frequently charge 8 per cent for buying their shares. Second, management fees typically are .5 per cent per year. Third, such funds almost never are continuously fully invested in common stocks and the portion of their assets not so invested on the average yields a lower rate of return than the portion in stocks-hence the lower average yield to investors in such funds.

Funds Provide Services

The second loose end is my statement that such funds and similar institutions can be and almost certainly are-useful. The simplest and most comprehensive affirmative evidence to a believer in free competitive markets is the very rapid growth of such funds. The funds provide valuable services. They persuade many individuals to invest in stocks rather than other things which in the past have been less profitable than stocks. They provide valuable bookkeeping and custodial services, a relatively efficient means of achieving diversification and associated risk reduction for small in investors, and reduction in the agony of choice and responsibility. The third loose end was implicit. I have said that on the average mutual funds-and by implication, other institutionally managed funds -have selected stocks which have performed about the same as all stocks or randomly selected groups. So far, I have said nothing explicit about variability among funds in any given period or variability from period to period. What I have said would lead you to infer -correctly-that in any given year the common stocks for about half the funds do slightly better than all stocks and half do slightly worse. This is also obviously true for periods of 5 or 10 years. What accounts for the variability? Is it more than the result that could be expected from a random sampling process? William Sharpe has a plausible explanation which will be published in January 1966 in the Journal of Business. He finds that much of the variance among funds in rates of return from investment in their shares is explained by risk measured by variance in the net asset value per share-and by costs of management. The correlation between risk and rate of return is, as would be expected, positive. The correlation between costs of management and rates of return is negative, a result which no longer surprises you, I hope. Before moving on to the final section of my talk, which is on the great random walk controversy, I would like to pause for a moment of pontification. It is worse than useless to do investment research which is conventional in method and speed, since such research costs money and results in decisions only as profitable as random selection.

Random Walk Controversy

The great random walk controversy has aroused passion and occasionally bitter acrimony. It is an unusual controversy in that those on one side only are passionate. The angry ones are those who sell investment advice based on "technical analysis," which is analysis designed to predict price movements in stocks on the basis of interpretations-often allegedly objective or scientific-of recent movements in the level of prices or indexes and of trading volume. *Business and Financial Weekly*, an old and respected publication, had in its August 30, 1965, issue offers to sell advice on investments by 21 different technical analysts. Although these persons may not believe that you can buy happiness, they believe or say they believe that you can buy dollars or wealth at a great discount. For a few dollars you are offered allegedly reliable information about future movements of individual stocks or the market as a whole. The technical analysts believe that there are recurrent, discernible patterns in stock prices or prices and trading volume and that such movements are assuredly not random. The bemused and detached parties to the controversy are typically academic economists and statisticians, unembarrassed by the question, "If you're smart, why aren't you rich?," who present strong though not definitive evidence of the statistical independence or randomness of successive changes in stock prices. If such randomness exists, most technical analysis is silly and the advice of many persons is revealed to have no value. So far, the random walkers have dealt extensively only with stock prices and not with prices and trading volume considered together. The first random walker is believed to have been Louis Bachelier who first presented evidence in 1900, *La Thtorie de Speculation*. His work was seminal but the gestation period was long. Only within the last 10 years has his work been rediscovered by persons interested in testing it and extending it with other data.

Test for Randomness

Those walking randomly through the financial community include Cootner, Fama, Roberts, Granger, Clive and Morgenstern, Arnold, Moore, and others. Several have tested for serial correlation in successive price changes and in all instances the coefficients were extremely close to zero. Fama used a runs test and Morgenstern et al. used spectral analysis. All conconcluded that the evidence was consistent with randomness. The chartists are unimpressed by this conventional statistical evidence and they buttress their skepticism with arguments that the models underlying the statistical tests used are too simple to identify the complicated patterns which exist and can be perceived and used with profit. One non-believer in randomness, Sydney Alexander, took another tack. He proposed an objective decision rule for investing which he claimed yielded profits far greater than a simple policy of buying and holding stocks. His device, called the "filter-technique," was designed to time purchases and sales and was based on persistence or trends in prices-allegedly profitable departures from randomness. His scheme, which I will divulge only if you promise not to use it till tomorrow, worked as follows:

(1) After a stock has risen X per cent buy-

(2) Hold till it has declined x per cent and then sell short-

(3) Repeat ad nauseum or bankruptcy.

Alexander presented evidence for filters of many sizes, many of which for the periods under study yielded profits greater than could be obtained by buying and holding the same securities. Fama, the most energetic and prolific randomist, redid Alexander's work, taking into account transaction costs and the fact that dividends are a cost rather than a benefit when stocks are sold short. These details revealed that all filters are extremely unprofitable, compared to buying and holding, except to the broker. So far, the randomists are unscathed and generally poor. There will be more debate and more attacks on the randomist stronghold, but there is a haunting fear that those with the best arguments are silently sunning and swimming at St. Tropez.

Lorie, James. Papers, Box 1, Folder 1, Special Collections Research Center, University of Chicago Library

You can read more about the Jim Lorie and his role with CRSP in:

Capital Ideas: the improbable origins of Modern Wall Street, by Peter L. Bernstein

Fischer Black and the Revolutionary Idea of Finance, by Perry Mehrling

Perspectives on Equity Indexing, Frank J Fabozzi Editor, Robert Paul Molay, Consultant

About Center for Research in Security Prices (CRSP)

In 1960, Chicago Booth embarked on development of the world's first comprehensive database for historical security prices and returns information. The research-quality data created by this transformational project has spawned a vast amount of scholarly research from several generations of academics. Today, CRSP databases underpin research and teaching at nearly 450 leading academic institutions in 35 countries.

For more than 50 years CRSP has maintained the level of quality and industry standards introduced by Professors Lawrence Fisher and James Lorie in the formation of CRSP. Today, with a staff of nearly 80 professionals, CRSP continues its commitment to providing research data for the most rigorous tests in academic research and in backtesting applications by practitioners. In addition, CRSP has extended its core products with the introduction of the CRSP Indexes.

CRSP's new series of transparent and investable indexes provide the foundation for new areas of research and serve as benchmarks for investment vehicles, such as ETFs. In October, 2012, Vanguard announced that it would adopt 16 of the CRSP Indexes as benchmarks for certain Vanguard ETFs. The indexes demonstrate not only CRSP's innovative thinking, but also the depth of our commitment to positively influence practices in the financial arena.