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 as follows:

"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."

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.

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. There-fore, 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 mag-netic 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 with-out 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 proce-dures for quality control. So that these proce-dures 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 prob-lem 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 Ex-change at the time we began the collection of data in mid-1960 and consulted the records of the stock exchange for listings and delistings be-tween 1926 and 1960. For each month that a com-pany was listed, we prepared a prepunched card for use as a coding form. This coding form con-tained 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 accu-rate and that our keypunch operator would some-times make mistakes. But we also knew that in printing approximately 400,000 prices the publish-ers 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

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.

YIELDS ON DIRECT PLACEMENTS SINCE 1951
(Abstract)

Avery B. Cohan, University of North Carolina
and National Bureau of Economic Research

This paper is a progress report on the National Bureau's study of direct placements (which, in turn, is part of the Bureau's larger study of interest rates).

The paper, first, describes the prominent place direct placements have come to occupy in the market for long-term funds. Then, after discussing sources of data, it describes statistical procedures in some detail. The problem to which these procedures respond is, essentially, the problem of constructing a price index when the underlying or intrinsic nature of the "commodity" is known to be changing. In this connection, the quality problem is discussed in detail.

The paper then presents tentative series on yields on direct placements and compares that series with comparable series on public offerings.

The paper concludes with a description of plans for future study.