# DERIVATION OF PROBABILITY DENSITY FUNCTIONS FOR THE RELATIVE DIFFERENCES IN THE STANDARD AND POOR'S 100 STOCK INDEX OVER VARIOUS INTERVALS OF TIME

DISSERTATION

Presented to the Graduate Council of the
University of North Texas in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

R. C. Bunger, B.B.A., M.B.A.
Denton, Texas

August, 1988

DRW

Bunger, Robert C., Derivation of Probability Density Functions for the Relative Differences in the Standard and Poor's 100 Stock Index over Various Intervals of Time. Doctor of Philosophy (Management Science), August, 1988, 38 pp., 4 tables, bibliography, 40 titles.

In this study a two-part mixed probability density function was derived which described the relative changes in the Standard and Poor's 100 Stock Index over various intervals of time. The density function is a mixture of two different halves of normal distributions. Optimal values for the standard deviations for the two halves and the mean are given. Also, a general form of the function is given which uses linear regression models to estimate the standard deviations and the means.

The density functions allow stock market participants trading index options and futures contracts on the S & P 100 Stock Index to determine probabilities of success or failure of trades involving price movements of certain magnitudes in given lengths of time.

TABLE OF CONTENTS

LIST OF TABLES . . . . . . . . . . . . . . . . . . v

Chapter

1. INTRODUCTION . . . . . . . . . . . . . . . . . 1

   A Historical Perspective

   Statement of the Problem

   Purpose of the Study

   Significance of the Study

2. LITERATURE REVIEW . . . . . . . . . . . . . 8

3. DATA DESCRIPTION . . . . . . . . . . . . . . 13

   Identifying the Index of Choice

   Specifying the Data

   Creating the Data Sets

4. METHODOLOGY AND RESULTS . . . . . . . . . . . 21

   Tests for Normality

   Tests for Normality After Adjusting Variances

   Tests for Normality Using Linear Regression Estimates

   Tests Against the Weibull Distribution

5. SUMMARY AND CONCLUSIONS . . . . . . . . . . . 30

   Possible Applications of the Density Function

   Suggestions for Further Study

REFERENCE LIST . . . . . . . . . . . . . . . . . . 35

## LIST OF TABLES


| Table | | Page |
|---|---|---|
| 1. | Statistics On Distributions of Relative Changes | 20 |
| 2. | Goodness-of-fit Test Results Using Unadjusted Parameters | 22 |
| 3. | Chi-Square Goodness-of-fit Tests Using Adjusted Standard Deviations | 25 |
| 4. | Goodness-of-fit Test Results (Regression Estimates Model) | 27 |

# CHAPTER 1

# INTRODUCTION

## A Historical Perspective

Trading in securities in America is reported by some historians (Teweles and Bradley 1982) to have begun as early as the 1720's when commodities and other goods were traded at an auction market which developed in lower New York at the lower end of Wall Street. Apparently the market in securities was unimportant and even obscure until the 1790's when the Continental Congress approved the issuance of Revolutionary War Bonds which were traded on the growing auction market.

A formal organization to trade securities was established in May of 1792 when a group of securities brokers signed an agreement to trade securities only among themselves and established the rate of commission which they would charge. This agreement became known as the "Buttonwood Tree Agreement" as it was common for these brokers to meet at noon each day under a particular buttonwood tree on lower Wall Street to conduct their transactions.

In 1817 the organization was further formally organized when the group copied the very formal securities trading organization which had been formed by a group of rival brokers in the city of Philadelphia in 1790. The newly formalized organization in New York was known as the New York Stock and Exchange Board and later became known as the New York Stock Exchange.

Over the next two centuries the scope of the stock exchanges and the amount of business they conducted grew explosively as the nation grew and the number of businesses with shares and other security instruments to trade increased dramatically. Regional exchanges were formed in various parts of the nation and other organizations were created to provide markets for the thousands of securities that resulted from the developing businesses.

As the stock market grew in scope and trading volume, participants desired and needed a way to refer to the value of the market in general as well as to changes in its value. The earliest successful tools to place value on the market were the Dow Jones Averages, the first of which was developed by Charles Dow in 1884. Eventually, Dow developed three averages which have come to be known as the Dow-Jones Industrial Average, the Dow-Jones Transportation Average, and the Dow-Jones Utility Average. The Industrial Average, which now consists of 30 stocks, is the best known of the three averages and is commonly cited by news media as a measure of the level of the stock market.

In time, other averages and indices were developed for the same purpose of referencing the level of the stock market. The most notable of these are: the Standard and Poor's 500 Stock Index, the New York Stock Exchange Index, the Wilshire 5000 Index, the NASDAQ Index, the Value-Line Composite Average, and the Standard and Poor's 100 Stock Index.

In the 1970's and 1980's numerous new trading and speculative instruments were created as adjuncts to the stock market. For some time there had been trading in options on stocks through dealers using a somewhat cumbersome mechanism of auctioning them by wire and telephone. A more extensive and proficient mechanism for trading stock options came into existence with the formation of the Chicago Board Options Exchange (CBOE). Its success led to the later formation of the American, Pacific, and Philadelphia option exchanges.

Soon after creation of the CBOE there was extensive trading in stock options, futures contracts on popular indices such as the Standard and Poor's 500 Index, options on the futures contract on the Standard and Poor's 500 Index, options on the Value Line Composite Average, and options on the Standard and Poor's 100 Index.

Portfolio managers of huge pension and mutual funds, insurance companies, and banking institutions discovered that with the availability of the new options and futures contracts they had methods at their disposal to hedge their enormous portfolios. Huge amounts of foreign capital, a general increase in the willingness of market participants to speculate, and a generally rising stock market fueled the growth of trading in futures and options on stocks, indices, and futures.

## Statement of the Problem

In March, 1983 options on the Standard and Poor's 100 Stock Index (commonly called the OEX, its ticker symbol) began trading. Within two or three years the OEX options became so extremely popular with investors and speculators that the open interest in them often exceeded one million contracts, representing over one million shares of each of the one hundred stocks which make up the OEX. Since the options are highly leveraged instruments, relatively small changes in the value of the OEX would have a large impact on the value of the holdings of investors in the options.

Although the ability to forecast or predict the value of stocks or the stock market in general had been needed by participants in the stock market since its inception, that need took on more importance as more and more of the highly leveraged options and futures contracts were traded by the funds and individual investors. Since options and futures contracts are short-lived instruments investors in them do not have the luxury of "waiting until they come back" if the investments in them are poorly timed and losses are quickly incurred.

There are serious doubts as to whether the direction that the price of a particular stock or the stock market in general will take can be forecast in any given instance. Also, doubts exists that the magnitude of any price movement can be forecast. However, it is a reasonable assumption that the general behavior of the stock market in the future is likely to be very similar to its general behavior in the past. Therefore, the ability to describe the past behavior of the market would offer valuable insights about the probable behavior of the market in the future.

## Purpose of the Study

The primary purpose of the study is to develop probability density functions which will describe the relative changes over various periods of time in the Standard and Poor's 100 Stock Index that have occurred in the past. Since it is obvious that a two-point change from a starting index value of 300 is not the same as a two-point change when the starting index value is 100, the changes are expressed as relative changes which are the signed ratios of the change over a specified interval of time to the values of the index at the beginning of the time interval.

The purpose of this study, then, is to develop the probability density functions of the relative changes in the

Standard and Poor's 100 Stock Index over various intervals of time. Such a density function would serve as a valuable tool to the market participants trading financial instruments such as options and futures which have market imposed trading-time limitations. With that view in mind, this study has at its forefront the purpose of developing a pragmatic tool rather than some possibly esoteric theoretical result. In addition to being readily adaptable to practical applications, the probability density functions should be intuitively appealing.

## Significance of the Study

The study has significance in that a probability density function which describes the relative changes in the OEX can be used as a tool to determine the probability of a relative change of a specified magnitude over a certain amount of time. Though the probability density function provides little or no information as to the likely direction of a change in the index, the probability of a relative change of a given magnitude is readily calculable.

Participants in applicable financial markets can find a ready application for this tool because in many cases it is readily determined that a certain amount of profit (or loss) will be sustained if a change of some known amount in the stock market occurs within a certain timeframe. A probability density function allows the ready calculation of the probability of such a change, and, therefore, would provide the user with what might be extremely valuable information.

# CHAPTER 2

## LITERATURE REVIEW

Though the behavior of stock prices had long been of interest to investors, it did not receive the widespread, serious attention of academic researchers until the 1960's. The first important study of the behavior of stock prices was the work of Frenchman Louis Bachelier. Bachelier (1900) in his doctoral dissertation proposed that stock price changes were independent of past price movements. His hypothesis became the cornerstone of the random-walk theory and has been the impetus for much of the analysis of stock price movements. Bachelier defined price differences as follows:

$$D = P(t + k) - P(t) \qquad (2\text{-}1)$$

where $P(t)$ is the price of some stock or financial instrument at time period $t$ and $P(t + k)$ is the price $k$ periods into the future. Bachelier proposed that the differences are distributed normally with mean zero and the variance proportional to $k$. This process became known as Brownian motion and satisfies the two conditions of the random-walk theory. The two conditions are that price changes are (1) independent, and, (2) conform to some probability distribution.

Bachelier's theory was widely held for some years enjoying support from the study by Osborne (1959) who labeled the process as "Brownian motion."

Quickly, however, challenges to the theory were brought by other researchers. Olivier (1926), Mills (1927), Larson (1960), Cootner (1962, 1964) and Alexander (1961) among others presented evidence that empirical distributions of price changes were typically too peaked in the middle and too fat in the tails to be Gaussian.

Fama (1965) and Mandelbrot (1963) conducted further research which suggested that distributions of stock price changes were not normal, but were, in fact leptokurtic and often skewed. The studies of Mandelbrot (1963) and Fama (1965) dealt with stock price differences by taking the transformation:

$$D = \ln P(t + k) - \ln P(t). \qquad (2\text{-}2)$$

These studies indicated that the empirical distributions showed serious departures from normality. They suggested that the distributions' variances appeared to behave as if they were infinite and that the distributions themselves conformed best to the non-Gaussian forms of the stable Paretian distributions of which the normal distribution is a limiting form.

Brada, Ernst, and Van Tassel (1966) agreed that price differences across periods of time were not normally distributed but suggested that such differences were distributed normally if the differencing was taken across transactions. They argued that differencing across time allowed for observations where many, few, or zero transactions were possible which introduced inconsistencies into the data and perhaps accounted for the lack of normality in the distributions. Mandelbrot and Taylor (1967), Granger and Morgenstern (1970), and Clark (1973) also employed the concept of "transaction time" in their studies of subordinated models.

Fama and Roll (1971), Officer (1972), Barnea and Downes(1973), Brenner (1974), Fielitz and Smith (1972), and Fielitz and Rozelle (1983) investigated the stability of the stable Paretian distribution and the degree to which the empirical distributions were described by it. Generally, their studies supported the use of the stable Paretian distribution to describe the empirical distributions whereas investigations by Blattberg and Gonedes (1974) and Hse, Miller and Wickern (1974) posed concerns about its usage. These investigations were often on the transformations called "returns," that is,

$$\text{return} = R = \ln[P(t + k) / P(t)]. \qquad (2\text{-}3)$$

Mandelbrot and Taylor (1967), Praetz (1972), and Granger and Morgenstern (1970) suggested the usage of subordinated models. Subordinated models of security returns are those formed by mixing two or more normal distributions. The subordinated models have received some support from the work of Brada, Ernst and Van Tassel (1966) and Clark (1973). Brenner (1974) conducted a study on a large segment of the daily changes in the Standard and Poor's Composite Index. His findings suggested that the distributions did not follow a stable, stationary process. He also proposed that a mixture of normal distributions was the most likely hypothesis.

Blattberg and Gonedes (1974) proposed that the Student's t distributions had greater validity for describing the empirical distributions than the symmetric-stable models proposed by others.

Westerfield's (1977) "transaction-time" study of 315 NYSE stocks over a period of 412 days suggested that the subordinated model was better than the stable Paretian model.

Dowell and Grube (1978) found that return behavior was sensitive to new information; that is, the variances of the returns were stable under the absence of news and unstable in the presence of news. They suggested a model of combining normal distributions of unequal variances.

Fielitz and Rozelle (1983) proposed that stock return distributions seem to follow the mixture-of-distributions model but it was unclear whether the distributions should be normal distributions with changing variances or nonnormal stable distributions with changing scale.

Becker (1975) determined a double truncated normal distribution to be satisfactory as a model for bond prices but unsatisfactory when dealing with stock prices.

It is readily apparent from the foregoing sketch of the major research efforts into the behavior of stock price changes that no clear-cut, definitive answer is yet available. Though it is now clear from much of the research since Bachelier (as well as this present study) that price changes do not conform to the normal distribution, it is not clear what distribution best describes the changes. The implications of these research efforts (though they are important and meaningful in their context) are nonetheless confounded by the variety of hypothetical models proposed, the wide range of different data on which the models were tested, and the various transformations taken on the data.

The end result is that a truly useful tool for the market participant is not presently available from these studies.

# CHAPTER 3

# DATA DESCRIPTION

## Identifying the Index of Choice

As stated previously, the main thrust of this investigation was to develop a practical tool for market participants to use to enhance their decision making. In particular, the tool should be useful in providing information about the market in general; therefore, it immediately follows that the focus of this investigation should be one of the popular stock market indices.

Though there are several popular stock market indices which could have been selected for this study, the Standard and Poor's 100 Stock Index (also called the OEX, its ticker symbol) was chosen. The OEX is the base index for a huge market in stock index options. At the time of this writing in late May of 1988, 231,000 option contracts were traded in a single day. The open interest stood at 910,000 contracts. The OEX is also the base for a market in stock index futures, but to a much lesser extent .

One important aspect of options trading is the market imposed time constraints placed on the life of index options. Every stock index option market participant understands that the opportunity window for taking action closes at the close of business on the date of the option's expiration. At expiration profits and losses are locked in and settlement must be made. Because this study focuses on the amount of change in the stock market in a given period of time, the OEX is a good choice for an index of study since change of some period of time is of vital concern to both writers and purchasers of options.

The OEX is a capitalization-weighted index of 100 stocks traded on national stock exchanges. These stocks also have stock options traded on the underlying stock. The purpose of the OEX, as with any index, is to gauge movements in the overall stock market without giving too much or too little influence to the changes in the prices of individual stocks. Since the OEX is a capitalization-weighted index it reflects the importance of stocks by weighting their influence on the index by both their prices and their number of outstanding shares.

The index has been calculated and published since 1976 when it was originated by the Chicago Board Options Exchange. The index is calculated by the equation in 3-1.

The initial base value was set at 100 on January 2, 1976 when the index was originated. The base value is adjusted when changes in the composition of index take place through mergers and acquisitions, substitutions of one stock for another, and the exercising of stock rights.

$$\text{OEX STOCK INDEX} = \frac{\sum_{j=1}^{100} P_j S_j}{\sum_{j=1}^{100} Pb_j Sb_j} \qquad \text{3-1}$$

$P_j$= current market price of stock j
$S_j$ current number of stock j shares
$P_{bj}$= price of stock j on the base date
$S_{bj}$= stock j shares on the base date

The adjusted base value is derived by the formula:

ABV = OBV ( AMV / OMV ); where,

ABV = adjusted base value

OBV = old base value

AMV = adjusted market value

OMV = old market value

Though the OEX has been in existence since January 2, 1976, option contracts (calls and puts) and futures contracts have been traded on the index only since March, 1983.

## Specifying the Data

Beginning in July, 1983 closing prices of the OEX began to be reported in hundredths of a point; prior to that time the index was reported in tenths of a point. Because of the additional precision the hundredth-point data affords, the data for this research consists of the closing values of the OEX for the period from July 22, 1983 through mid-February of 1988. The data set consists of more than eleven hundred closing values and encompasses a time period of about four and one-half years, slightly longer than what is generally considered the length of a business cycle.

This time period is characterized by both bear markets and bull markets and, further, has the distinction of having encompassed the dramatic bull market of early 1987 and the devastating plunge in stock prices now referred to as the "October Meltdown" of mid-October, 1987. It is a reasonable assumption that the market during this period of time is reasonably representative of the market generally.

The original data were purchased from Hale Computing Corporation of Palo Alto, California and were downloaded by computer. They are believed to be essentially error-free. The data is the value (price) of the index at the close of business on each trading day during the four and one-half year period.

In addition to the fact that closing prices are readily available, the use of the closing value of the index each day affords one significant advantage to the application of this study. Since the opportunity window for action on a option or futures contracts closes at the close of business on the day of expiration, it seems intuitively sensible to take measurements at the close of business each day. Further, it seems likely that a collection of measurements taken consistently at some other time during the trading day, say twelve-o'clock Noon, would be very similar to the collection of measurements taken at the close of business.

## Creating the Data Sets

Relative changes in the OEX were selected as the measurement to be studied in this investigation because absolute changes present a distorted view of market activity. It is readily apparent that a price movement of five points when the market starts at 200 is not the same as a price movement of five points when the market starts at 100. To express the movement as a relative change presents the better measurement.

Also, a relative change measurement is easily converted to an absolute change by simply taking the product of the relative change and the current level of the index. Not only is this process simple and applicable to any level of the index, it is intuitively appealing. A relative change is given by:

$$r = ( P(t+k) - P(t) ) / P(t) \qquad (3\text{-}2)$$

where;
$P(t+k)$ is the closing value k days in the future
$P(t)$ is the closing value at time period t.

All possible relative changes were calculated from the data set, differencing over the intervals of one to fifteen days in increments of one day and from twenty to forty days in increments of five days, yielding twenty sets of relative change data. These differencing interval magnitudes were selected because change in the index is extremely important to the trader as the expiration date draws near. The differencing intervals of from one day to fifteen days addresses this aspect. Longer times to expiration than fifteen days are of interest but not as crucial as the shorter time frames. Differencing intervals of from 20 to 40 days covers the span in calendar time of a month to two months into the future which is about as long as most traders have serious interest in trading contracts.

Each set of relative changes (there is one for each of the twenty specified differencing intervals) was generated by using formula 3-2 and stepping through the data sequentially one day at a time. For example, in the case of differencing over an interval of ten days, the first difference calculated was determined by subtracting the closing price of the index on the first day from the closing price on the eleventh day and then dividing the obtained difference by the closing price on the first day. The yielded value is the relative change in the index over that particular ten-day period.

The next calculation was obtained by subtracting the closing price of the second day from the closing price of the twelfth day and then dividing the obtained difference by the second day's closing price. This process was continued until all possible relative changes over the ten-day interval were calculated.

Note that by stepping through the data one day at a time the intervals overlap somewhat for differencing intervals greater than one day. Including overlap in the data is a departure from the procedures used by the researchers cited in the prior research. Since traders (or for that matter the same trader) have the opportunity to make multiple trades of the same duration but with different starting and ending dates such that the durations overlap, it seems reasonable that the inclusion of overlapping time periods in this study follows the stated intent to develop a practical tool readily applicable to use in the market.

Using this procedure described above, twenty distributions of relative changes were formed for differencing intervals of one to fifteen, twenty, twenty-five, thirty, thirty-five, and forty days.

Descriptive statistics, means, medians, modes, standard deviations, and measures of skewness of the distributions of changes are provided in Table 1.

Further visual inspections of the histograms of the twenty distributions revealed that the distributions typically exhibited shapes similar to the characteristic bell-shape of normal distributions. A closer inspection revealed that the half of the distributions to the left of the mean were more peaked than the right half, and further, appeared to have more outliers greatly distant from the mean in the left tail than in the right tail.

TABLE 1

STATISTICS ON DISTRIBUTIONS OF RELATIVE CHANGES

| K | SIZE | MEAN | MEDIAN | MODE | STD.DEV. | SKEWNESS |
|---|---|---|---|---|---|---|
| 1 | 1134 | 0.000470 | 0.0004 | 0.0000 | 0.01296 | -3.79747 |
| 2 | 1133 | 0.000929 | 0.0014 | 0.0047 | 0.01828 | -2.29246 |
| 3 | 1132 | 0.001369 | 0.0024 | -.0020 | 0.02162 | -2.48977 |
| 4 | 1131 | 0.001800 | 0.0029 | -.0132 | 0.02420 | -2.37263 |
| 5 | 1130 | 0.002213 | 0.0031 | -.0002 | 0.02622 | -2.30005 |
| 6 | 1129 | 0.002624 | 0.0036 | -.0067 | 0.02856 | -2.19200 |
| 7 | 1128 | 0.003042 | 0.0041 | 0.0014 | 0.03073 | -2.22598 |
| 8 | 1127 | 0.003442 | 0.0040 | -.0203 | 0.03288 | -2.17803 |
| 9 | 1126 | 0.003814 | 0.0044 | -.0201 | 0.03487 | -2.16506 |
| 10 | 1125 | 0.004195 | 0.0054 | -.0290 | 0.03677 | -2.13807 |
| 11 | 1124 | 0.004581 | 0.0065 | -.0260 | 0.03847 | -2.10902 |
| 12 | 1123 | 0.004973 | 0.0071 | -.0313 | 0.04013 | -2.00182 |
| 13 | 1122 | 0.005393 | 0.0073 | 0.0067 | 0.04187 | -1.86126 |
| 14 | 1121 | 0.005808 | 0.0076 | 0.0278 | 0.04337 | -1.80367 |
| 15 | 1120 | 0.006229 | 0.0083 | -.0058 | 0.04477 | -1.79325 |
| 20 | 1115 | 0.008195 | 0.0102 | 0.0116 | 0.05155 | -1.54070 |
| 25 | 1110 | 0.010076 | 0.0107 | 0.0167 | 0.05692 | -1.45600 |
| 30 | 1105 | 0.012021 | 0.0121 | -.0457 | 0.06223 | -1.39859 |
| 35 | 1100 | 0.014186 | 0.0129 | -.0174 | 0.06858 | -1.40292 |
| 40 | 1095 | 0.016338 | 0.0158 | -.0013 | 0.07470 | -1.41478 |

Note: Skewness is calculated as:
$\underline{Sk} = n / [(n-1)(n-2)\Sigma(X - X)^3 / s^3]$

# CHAPTER 4

# METHODOLOGY AND RESULTS

## Tests for Normality

Chi-square goodness-of-fit tests with equally probable classes were performed on the twenty distributions to determine if they were normally distributed. Each distribution was divided into ten equally-probable classes providing a goodness-of-fit test with nine degrees of freedom. The hypotheses statements for each test were:

Ho: The distribution is normally distributed.

Ha: The distribution is not normally distributed.

A computer program was written to determine the frequencies in each class and calculate the Chi-square statistics for each test.

The null hypothesis was rejected for all twenty distributions with $p$ less than 0.001 for each test. This result indicates that the distributions of relative changes are not normally distributed and is consistent with the findings of other researchers. Note, however, that the only directly comparable result here with those of previous research is the test for normality on the differencing interval of one day. The other distributions contain overlap in the differencing intervals which other researchers did not include, but, which is reasonable for this study.

The Chi-square statistics for these tests are given in Table 2.

TABLE 2

GOODNESS-OF-FIT TEST RESULTS
(SAMPLE MEANS, STD.DEVIATIONS USED)

| K | CHI-SQUARE |
|---|---|
| 1 | 188.29 |
| 2 | 144.87 |
| 3 | 107.57 |
| 4 | 79.86 |
| 5 | 72.53 |
| 6 | 82.84 |
| 7 | 80.79 |
| 8 | 101.22 |
| 9 | 114.53 |
| 10 | 111.41 |
| 11 | 105.50 |
| 12 | 99.41 |
| 13 | 91.51 |
| 14 | 89.41 |
| 15 | 82.19 |
| 20 | 86.33 |
| 25 | 85.00 |
| 30 | 78.44 |
| 35 | 129.91 |
| 40 | 116.90 |

Note: $p < 0.001$ for all tests.

## Tests for Normality After Adjusting Variances

Noting that the distributions had a few, but extreme, outliers in each tail which would have the effect of inflating the variances of the distributions, the variances of the distributions were adjusted to reduce the influence of the outliers.

Comparing the parts of the distributions to the left of the mean to the parts to the right of the mean revealed that the parts exhibited different variances. Therefore, the first step was to calculate the standard deviations for each side of each distribution.

This was done by calculating the standard deviations for those values in the distributions which were less than the total sample mean. The standard deviations for the left part of the distributions were calculated from the total sample mean using formula 4-1.

$$s_1 = [\Sigma(x-\bar{x})^2/(n-1)]^{\frac{1}{2}} \quad ; \ x \leq \bar{x} \qquad (4\text{-}1)$$

In like fashion the standard deviations for those values greater than the total sample mean were also calculated.

Next these standard deviations were adjusted by multiplying them by an adjustment factor which would minimize the Chi-square statistic for that half of the distribution when tested for fit to a normal distribution. The procedure followed here was one of sensitivity analysis where successive goodness-of-fit tests were conducted while slightly changing the adjustment factor. By observing the resulting Chi-square statistics the optimal adjustment factor and subsequent adjusted standard deviations were determined.

The resulting adjusted standard deviations therefore provided the best fit of the distributions to a theoretical distribution made up of two parts where the left half is a normal distribution with standard deviation $\sigma_l$ and the right half is a normal distribution with standard deviation $\sigma_r$. The function is defined as follows:

$$f(x) = \begin{cases} \dfrac{1}{(2\pi)^{\frac{1}{2}}\sigma_l}\, e^{-[(x-\mu)/\sigma_l]^2} & ;\ \text{for } x \le \mu \\[2ex] \dfrac{1}{(2\pi)^{\frac{1}{2}}\sigma_r}\, e^{-[(x-\mu)/\sigma_r]^2} & ;\ \text{for } x > \mu \end{cases}$$

where the values for $\sigma_l$ and $\sigma_r$ are given in Table 3. Values for $\mu$ are the same as the sample means in Table 1. Chi-square test statistics and $p$ for the tests are also provided. Seven degrees of freedom are used because the mean and variance parameters are estimated from the data causing a loss of two degrees of freedom.

TABLE 3

CHI-SQUARE GOODNESS-OF-FIT TESTS
USING ADJUSTED STANDARD DEVIATIONS

| K | $\sigma_l$ | $\sigma_r$ | Chi-Square | p |
|---|---|---|---|---|
| 1 | 0.00850 | 0.00928 | 10.797 | 0.148 |
| 2 | 0.01394 | 0.01375 | 15.923 | 0.026 |
| 3 | 0.01665 | 0.01671 | 5.474 | 0.602 |
| 4 | 0.01968 | 0.01852 | 5.525 | 0.596 |
| 5 | 0.02072 | 0.02166 | 6.035 | 0.536 |
| 6 | 0.02241 | 0.02356 | 10.212 | 0.177 |
| 7 | 0.02300 | 0.02472 | 4.979 | 0.663 |
| 8 | 0.02375 | 0.02681 | 2.734 | 0.908 |
| 9 | 0.02492 | 0.02821 | 3.556 | 0.829 |
| 10 | 0.02658 | 0.02893 | 4.787 | 0.686 |
| 11 | 0.02876 | 0.02953 | 7.886 | 0.343 |
| 12 | 0.03122 | 0.03128 | 12.877 | 0.075 |
| 13 | 0.03229 | 0.03284 | 10.656 | 0.154 |
| 14 | 0.03434 | 0.03433 | 12.497 | 0.085 |
| 15 | 0.03565 | 0.03477 | 9.304 | 0.232 |
| 20 | 0.03888 | 0.04295 | 6.139 | 0.524 |
| 25 | 0.04182 | 0.04950 | 4.378 | 0.735 |
| 30 | 0.04408 | 0.05601 | 6.665 | 0.465 |
| 35 | 0.04485 | 0.06207 | 15.309 | 0.032 |
| 40 | 0.05385 | 0.06921 | 17.265 | 0.016 |

Note: K is the differencing interval in trading days.

## Tests for Normality Using Linear Regression Estimates

Having determined the optimal adjusted standard deviations for the left and right sides of the twenty distributions, the next step was to determine a general model. The procedure was to perform least-squares regression on the optimal adjusted standard deviations for the left side, the optimal adjusted standard deviations for the right side, and the sample means. The theoretical model is the same as given previously in 4-2 but $\sigma_l$, $\sigma_r$, and $\mu$ are replaced by the following regression models:

$$\mu = 0.00017909 + 0.00040043\ K$$

$$\sigma_l = 0.01082537 + 0.00186737\ K - 0.0000225\ K^2$$

$$\sigma_r = 0.01124505 + 0.00180870\ K - 0.0000098\ K^2$$

The results of testing the empirical distributions for fit to the theoretical distributions using the substituted values generated by the regression models above are given in Table 4. This test was the Chi-square goodness-of-fit test with seven degrees of freedom. Note that the theoretical distribution describes the empirical distributions very well except for differencing intervals of one and two days and for more than thirty days.

## Tests Against the Weibull Distribution

Tests were performed on the twenty distributions to determine if they possess the characteristics of the three-parameter Weibull distribution. The general form

TABLE 4

GOODNESS-OF-FIT TEST RESULTS
(REGRESSION ESTIMATES MODEL)

| K | Chi-square | p |
|---|---|---|
| 1 | 181.538 | 0.000 |
| 2 | 30.892 | 0.000 |
| 3 | 7.346 | 0.394 |
| 4 | 12.033 | 0.099 |
| 5 | 12.460 | 0.086 |
| 6 | 15.260 | 0.033 |
| 7 | 7.266 | 0.402 |
| 8 | 6.052 | 0.534 |
| 9 | 7.464 | 0.382 |
| 10 | 5.871 | 0.555 |
| 11 | 8.651 | 0.279 |
| 12 | 16.653 | 0.020 |
| 13 | 12.332 | 0.090 |
| 14 | 16.707 | 0.019 |
| 15 | 11.179 | 0.131 |
| 20 | 7.502 | 0.379 |
| 25 | 8.468 | 0.293 |
| 30 | 8.765 | 0.270 |
| 35 | 23.164 | 0.002 |
| 40 | 20.114 | 0.005 |

of the Weibull distribution is:

$$F(x) = 1 - \exp[ -\{(x-\alpha)/\beta\}^m], \ x>\alpha; \ \beta>0, \ m>1$$

where:
$\alpha$ = the location parameter
$\beta$ = the scale parameter
m = the shape parameter

The two-parameter Weibull has the same general form as above except that the location parameter, $\alpha$, is known to be zero. The three-parameter Weibull distribution transforms to the two-parameter Weibull if the smallest value in the distribution is subtracted from every observation in the distribution, causing $\alpha$ to take on the value zero. The tests on the relative change distributions on the OEX were conducted by first transforming the data by the transformation:

$$X' = X - \alpha_i, \ i = 1, \ldots, n$$

The set X' was then tested against the two-parameter Weibull distribution with the scale parameter, $\beta$, and the shape parameter, m, both unknown.

The testing procedure was conducted by taking 100 randomly selected samples of size 50 from each of the twenty location-transformed distributions of relative changes. For each of the 100 samples, estimates of $\beta$ and m were made using the iterative procedure recommended by Stephens (1974). Then each of the samples was tested against a Weibull distribution using the estimated parameters. The hypothesis statements were:

Ho: The sample is from a Weibull distribution ($\beta$,m)

Ha: The sample is not from the Weibull distribution

Each sample was tested at a level of significance of 0.01 using four test statistics. The four statistics used were $D^+$, the maximum positive difference between the sample distribution and the theoretical one; $D^-$, the maximum negative difference; D, the commonly known Kolmogorov-Smirnoff D; and V, the sum of $D^+$ and $D^-$, a statistic given by Kuiper (1960). The number of rejections out of each of the 100 samples tested were counted. This procedure was repeated for each of the twenty distributions.

At alpha = 0.01 the expected frequency of rejections in a collection of 100 samples would be one. The variance for this distribution of rejections is 0.99 (100 x .01 x .99, using the binomial probability law), yielding a standard deviation of 0.995. Therefore, there is small probability of getting more than five sample rejections (1 plus 4 standard deviations of 0.995 each) out of 100. In all twenty cases the number of rejections out of 100 was more than five. This result leads to the conclusion than the samples were not taken from Weibull distributions with parameters $\beta$ and m and the adjusted location parameter equal to zero.

# CHAPTER 5

## SUMMARY AND CONCLUSIONS

Around the turn of the century Bachelier first suggested the hypothesis that stock price changes conformed well to the normal distribution. This notion, however, was firmly rejected by most of the subsequent research; and, also by this research. Chi-square goodness-of-fit tests on twenty distributions of relative changes in the Standard and Poor's 100 Stock Index (OEX) rejected the hypothesis that the sample distributions came from parent distributions which were normal in functional form.

Also, extensive testing conclusively established at all reasonable levels of significance that the sample distributions are not Weibull distributions using estimated parameters.

This present study does indicate, however, that the distributions of relative changes in the OEX do possess some of the properties of the normal distribution. It was found that a two-part normal distribution with substantially reduced variances provides a very good descriptor of the distributions of relative changes. The two-part distribution is a mixture of two normal distributions; the left half from one normal distribution and the right half from a similar but different normal distribution.

The two-part distribution is formed by taking the left side of a normal distribution with its standard deviation adjusted to a smaller value than the total sample standard deviation and the right side of a normal distribution with a similarly adjusted smaller standard deviation. It is important to note that the standard deviations of the two sides are not equal. In fact, the standard deviations for the right side are typically a little larger than the left-side standard deviations.

The empirical distributions conformed remarkably well to the two-part mixed normal distribution described above. Chi-square goodness-of-fit tests using ten equally probable classes with seven degrees of freedom resulted in tests in which only eight of the twenty p values were less than 0.20 and only eleven were less than 0.50. Though, strictly speaking, this result does not prove that the density function proposed in 4-2 describes the relative changes in the OEX, that notion can not be rejected with much confidence for most of the differencing intervals.

Using the two-part mixed normal distribution in conjunction with linear regression models to estimate the mean, the left-side standard deviation, and the right-side standard deviation also yielded a reasonably good descriptor of the distributions of relative changes. Chi-square goodness-of-fit tests were performed as described above. The tests yielded results where eleven of the p values were close to 0.10 or greater. Once again, for most of the distributions the hypothesis that they conform well to the proposed general two-part density function using estimated parameters can not be rejected with much confidence.

This study, then, supports the conclusion that the density function proposed in formula 4-2 describes the relative changes in the OEX well. Further, provided that the statistical assumption that large samples represent their respective populations well remains valid for this case (note that samples of about 1100 were used here lending strength to this assumption), the density function should represent well the relative changes in the S&P 100 Index in the future.

## Possible Applications of the Density Function

As stated previously, the thrust of this study was to develop a tool to enable stock market participants to enhance their decision-making. The density function proposed here should serve quite well as such a tool. As with any density function, 4-2 can be used to determine the probabilities. Specifically, it can be used to determine the probability of relative changes over a given range within a certain number of days. Then, knowing the current level of the index itself the range in terms of actual units of the index could be calculated having the same probability. Knowing this probability should provide extremely valuable information to participants in the OEX options market, especially writers of options.

For example, consider the case where the OEX is currently at a level of 259.930 and a writer of calls is considering writing a 265 call with nine days remaining to expiration. The writer would collect a premium of, shall we say, $2.12 per share. If the index moves upward above 265 the call would go "in the money" and have intrinsic value. Once the index goes above 267.12 the writer's collected premium is lost and, additionally, a net loss is assured if the option expires with the index above 267.12. For this scenario, then, the call writer has a keen interest in knowing the probability of the index not moving above 265, a relative change of (265-259.93) / 259.93 = +0.0195. The probability that the index will not rise above 265 is calculated to be 0.7123 using formula 4-2. Thus the writer of the call has a probability of 0.7123 of retaining all of the premium collected on the trade.

In a similar fashion, the probability of the index going above 267.12 could be calculated giving the probability that the writer would lose the $2.12 premium which he had collected as profit and some of his own money. This probability is determined to be 0.1977 which is the probability that the trader will lose some of his own money.

By considering this situation from another perspective the probabilities attendant to the buyer of the call are available. The buyer has a probability of 0.1977 of recouping his paid-out premium and making some profit; and a probability of 0.7123 of losing all of his premium.

Similar probability calculations could be applied to futures trading scenarios; however, examples will not be presented here.

## Suggestions for Further Study

The findings of this study which was limited solely to the study of the S & P 100 Stock Index suggest that the analysis techniques employed in this study would be applicable to the study of other indices and even individual stocks. Since there is a huge market in futures contracts based on the Standard and Poor's 500 Stock Index, a similar analysis of that index is suggested.

This study was conducted as if entry into the market (as based on the OEX) was a random event without consideration of any kind of information that might be available. However, it seems intuitively apparent that the variance of a distribution of relative changes would be greater during periods of high volatility than during periods of low volatility. It is immediately suggested, then, that a study would be worthwhile which would group the relative changes according to the volatility in the market at the time the change occurred. Such a study would no doubt yield interesting results.

## REFERENCES


Alexander, S. S. 1961. Price movements in speculation markets: Trends or random walks. Industrial Management Review II: 7-26.

Ansbacher, Max G. 1983. The new stock index market. New York: Walker and Company.

Bachelier, L. J. 1900. Theorie de la speculations, Paris: Gauthier-Villars, 1900. Reprinted in The random character of stock market prices, edited by P. H. Cootner. 1964. Cambridge: MIT Press.

Barnea, A., and D. H. Downes. 1973. A reexamination of the empirical distribution of stock price changes. Journal of the American Statistical Society 68: 348-350.

Becker, J. D. 1975. A truncated distribution for security returns. UMSL Working Paper Series (November).

________. 1977. The three-parameter security valuation model. Unpublished Ph.D. dissertation, Washington Univ.

________ and W. L. Eckardt. 1979. Market skewness--a leading indicator for security market performance. TIMS/ORSA Proceedings.

Blattberg, R. C., and N. J. Gonedes. 1974. A comparison of the stable and Student distribution as statistical models for stock prices. Journal of Business 47:244-280.

Brada, Josef, Harry Ernst, and John Van Tassel. 1966. The distribution of stock price differences: Gaussian after all. Operations Research 334-340.

Brenner, M. 1974. On the stability of the distribution of the market component in stock price changes. Journal of Financial and Quantitative Analysis 9: 945-961.

Clark, P. K. 1973. A subordinate stochastic process model with finite variance for speculative prices. Econometrica 41: 135-155.

Cootner, P. H. 1962. Stock prices: Random vs. systematic changes. Industrial Management Review III (Spring): 25-45.

________, ed. 1964. The random character of stock market prices. Cambridge: MIT Press.

Dowell, C. D., and R. C. Grube. 1978. Common stock return distributions during homogeneous activity periods. Journal of Financial and Quantitative Analysis 12: 79-92.

Fama, E. F. 1963. Mandelbrot and the stable paretian hypothesis. Journal of Business 36: 420-429.

________. 1965. The behavior of stock market prices. Journal of Business 38: 34-105.

________ and Richard Roll. 1971. Parameter estimates for symmetric stable distributions. Journal of the American Statistical Association (June) 66: 331-8.

Fielitz, B. D. and J. P. Rozell. 1983. Stable distributions and the mixtures of distributions hypotheses for common stock returns. Journal of the American Statistical Association 78: 28-36.

________ and E. W. Smith. 1972. Asymmetric stable distributions of stock price changes. Journal of the American Statistical Association (December) 66: 813-814.

Granger, C. W. J. and O. Morgenstern. 1970. Predictability of stock market prices. Lexington, Mass.: D. C. Heath.

Hsu, D., R. B. Miller, and D. W. Wickern. 1974. On the stable paretian behavior of stock market prices. Journal of the American Statistical Association (March) 69: 108-113.

Hines, W. W. and D. C. Montgomery. 1980. Probability and statistics in engineering and management science. New York: John Wiley and Sons.

Kendall, M. G. 1953. The analysis of economic time series, I: prices. Journal of the Royal Statistical Society Ser. A, XCVI: 11-25.

Kon, S. J. 1984. Models of stock returns--a comparison. Journal of Finance 39: 147-165.

Kuiper, N. H. 1960. Tests concerning random points on a circle. Proc. Koninkl. Neder. Akad. van. Wetenschappen A 63: 38-47.

Luskin, D. L. 1987. Index options and futures. 1987. New York: John Wiley and Sons.

Mandelbrot, B. 1963. The variation of certain speculative prices. Journal of Business 36: 394-419.

________ and H. Taylor. 1967. On the distribution of stock price differences. Operations Research (November): 1057-1062.

Moore, Arnold. 1962. A statistical analysis of common stock prices. Ph.D. diss., University of Chicago.

Niederhofer, V. 1965. Clustering of stock prices. Operations Research 13(2).

Officer, R. R. 1972. The distribution of stock returns. Journal of the American Statistical Association 67: 394-419.

Osborne, M. F. M. 1959. Brownian motion in the stock market. Operations Research VII: 145-73.

Praetz, P. D. 1972. The distribution of share price changes. Journal of Business 45: 49-55.

Press, S. J. 1967. A compound events model for security prices. Journal of Business 49: 317-335.

Roberts, H. V. 1959. Stock market 'patterns' and financial analysis: Methodological Suggestions. Journal of Finance XIV: 1-10.

Schwartz, R. A., and D. K. Whitcomb. 1977. The time variance relationship: Evidence on autocorrelation in common stock returns. Journal of Finance 32: 41-55.

Stephens, M. A. 1974. EDF statistics for goodness-of-fit and some comparisons. Journal of the American Statistical Association 69: 730-737.

Teichmoeller, J. 1971. A note on the distribution of stock price changes. Journal of the American Statistical Association 66: 282-284.

Teweles, R. J., and E. S. Bradley. 1982. The stock market, 4th ed. New York: John Wiley and Sons.

Westerfield, R. 1977. The distribution of common stock price changes: An application of transactions time and subordinated stochastic models. Journal of Financial Quantitative Analysis 12: 743-766.