EFFECTS OF BID-ASK SPREADS AND PRICE DISCRETENESS ON STOCK RETURNS

by

Ajay R. Dravid

6-91

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

EFFECTS OF BID-ASK SPREADS AND PRICE DISCRETENESS ON STOCK RETURNS

Ajay R. Dravid
Assistant Professor of Finance
The Wharton School
University of Pennsylvania

This paper is based on Chapter 3 of my Ph.D. dissertation at Stanford University. I thank the members of my dissertation committee — George Foster, Allan Kleidon and especially Mike Gibbons — for their encouragement and support. Helpful comments and suggestions were received from Bruce Grundy, Bob Litzenberger, Ananth Madhavan, Paul Pfleiderer, Peter Reiss, Tong-sheng Sun and Eric Terry; seminar participants at several schools also provided useful input, especially Mike Rozeff, Hans Stoll and Jim Ohlson. Financial support from Stanford University, The Wharton School, and the University of Pennsylvania Research Foundation is gratefully acknowledged.

EFFECTS OF BID-ASK SPREADS AND PRICE DISCRETENESS ON STOCK RETURNS

Abstract

This paper shows that the effects of bid-ask spreads and price discreteness on observed stock returns are related to stock price level, properties of the bid-ask spread and the nature of the rounding process. Using a model similar to Harris (1990), we derive robust Taylor series approximations relating the moments of observed stock returns to the underlying true return moments. Previous results from the literature are shown to be simple special cases of our results. We suggest explanations for seemingly anomalous empirical results such as the average non-January size effect, changes in post-split stock return volatility and the serial correlation of stock returns.

Effects of Bid-Ask Spreads and Price Discreteness on Stock Returns

I. Introduction

Much of the early financial economics literature assumed that capital markets are perfect and that prices of securities in these markets are discovered through a trading process closely resembling the classical Walrasian auction. Frictions such as trading costs, decision and information costs, and other institutional factors that affect the trading process, were not modeled explicitly.[1] More recently, the "microstructure" of security markets has been receiving considerable attention. Cohen, Maier, Schwartz and Whitcomb (1986) point out that this is the result of major structural changes in markets, as well as recent empirical evidence that measured security returns are affected by institutional and other factors. In support of their position, they cite serial correlation in security and index returns, biases in OLS beta estimates, and several unresolved anomalies in the behavior of the returns on small firms.

The market microstructure literature recognizes explicitly that prices (and hence returns) observed in security markets differ from what can be thought of as their "true underlying values." Two broad types of studies exist. Garman (1976), Glosten and Milgrom (1985) and Easley and O'Hara (1987) are examples of theoretical models. These studies model the process by which transaction prices of securities are discovered, by making assumptions about the structure of the market, trading systems, information asymmetries, types of traders (such as liquidity trading versus informed trading) and the behavior of market makers. The second type are econometric models, which examine relations between true and observed security prices without formal behavioral assumptions. Examples of these include Blume and Stambaugh (1983), Roll (1984), Harris (1990) and this paper.

[1] Some notable exceptions include Stigler (1964), Demsetz (1968), Garman (1976) and Stoll (1978). The interested reader is referred to Cohen, Maier, Schwartz and Whitcomb (1986) for a detailed survey of the issues and literature.

These studies address the issue from the perspective of the econometrician who uses observed security prices and returns in empirical testing. If the theory or model being tested rests on assumptions of perfect and frictionless markets, then deviations between observed and true prices should be viewed as measurement errors that may influence inferences drawn from the tests.[2]

In this paper, we present a simple econometric model of the impact of two types of institutional factors — bid-ask spreads and price discreteness — on observed stock returns.[3] Our approach is similar to that of Cohen, Hawawini, Maier, Schwartz and Whitcomb (1980), who demonstrate that a series of "true returns" generated in a frictionless world behaves differently from the series of "observed" returns obtained when there is friction in the trading process. In Section II, we motivate our research by some examples of stock return histograms that illustrate the nature and magnitude of these effects. When stocks trade at low prices, their observed returns exhibit higher expected values and variances, and negative serial correlation. Departures from normality are found to be particularly striking. All these effects become much more severe at progressively lower prices as the relative importance of the spread and price discreteness increases. We discuss our model in Section III. While the basic model is very similar to that of Harris (1990), our approach is quite different. Using Taylor series expansions, we derive approximate analytical results relating the moments of observed returns and the underlying true returns. These results are quite robust, since they do not depend on distributional assumptions regarding the underlying returns. We also show that the analytical results obtained by Harris and other authors can be viewed as special cases of our results. In Section IV, we turn our attention to stock price changes, which are easier to analyze than stock returns. In Section V, by assigning reasonable numerical values to the bid-ask spread and other parameters, we

[2] An analogy from the physical sciences may be helpful: the laws of Newtonian mechanics are based on the concept of point masses moving in a perfect vacuum. If we were to test them in the real world and ignore the fact that our experimental data is affected by friction, air resistance, etc., we would reject those laws.

[3] Although we use the term "stock returns" almost exclusively, it must be emphasized that our results are applicable to any security market in which bid-ask spreads exist, and where prices are restricted to be multiples of some minimum tick size.

evaluate numerically the analytical results from previous sections. We show that some anomalous results previously reported in the literature — such as the "size" effect (Banz (1981), Keim (1983) and others) and changes in the post-split volatility of stock returns (Ohlson and Penman (1985), Dravid (1987) and others) — can largely be explained by our model. Section VI concludes the paper. Details of derivations and simulation results are presented in the Appendices.

II. Motivation: Previous Studies and an Example

In this section, we discuss some previous studies dealing with bid-ask spreads and price discreteness. We also provide motivation for our model with the help of an example that illustrates the impact of these institutional factors on observed stock returns.

A. Previous Studies

Our analysis considers the effects of bid-ask spreads and price discreteness on observed stock returns. It should be emphasized at the outset that we use a simple transaction-cost model of the observed or effective bid-ask spread. Several authors have proposed theories of market-maker inventory control and information asymmetry to model the quoted spread.[4] In our model, it is the market spread that affects observed returns, as a result of transaction prices bouncing between the highest bid and lowest ask prices. The implications of the existence of the bid-ask bounce have been the subject of a number of previous studies;

[4] See, for example, Bagehot (1971), Copeland and Galai (1983), Glosten and Milgrom (1985) and others. Most of these models assume that the specialist is the major or sole provider of liquidity to the market and sets the bid-ask spread. The difference between the bid and ask prices quoted by such a market maker can be termed as the "quoted" spread. In contrast to this quoted spread, we are concerned here with the "market" (or "effective") spread, which does not typically reflect the quotes of the specialist or any other single dealer: it is the difference between the highest bid and lowest ask prices across all limit orders and dealer quotes. For more details about institutional aspects of the trading process and the market spread, the reader is referred to Pessin (1985, pp. 218–269) and Cohen, Maier, Schwartz and Whitcomb (1986, Chapter 5.) Roll (1984) and Stoll (1989) also provide insightful discussions of the effective versus quoted spread.

Blume and Stambaugh (1983), Stoll and Whaley (1983), Roll (1984), French and Roll (1986) and Harris (1990) are of particular relevance to this paper.

A second institutional factor that affects observed stock prices and returns is price discreteness: the restriction that quotes and transaction prices be multiples of some specified minimum "tick size." On major U.S. exchanges, this is usually \$1/8.[5] This issue of price discreteness or rounding was recognized by Schwartz and Whitcomb (1977, p. 301). They note that as a result of NYSE prices being quoted in minimum units of \$1/8, "what would be a smooth price series is in effect rounded and becomes a lumpy series." Gottlieb and Kalay (1985) show that discreteness results in an upward bias in the estimation of the variance and higher moments of stock returns. Harris (1990) re-examines many of these issues and provides support for the results of Gottlieb and Kalay (1985). Hausman, Lo and MacKinlay (1991) use an ordered probit model to analyze transaction stock price changes. They conclude that price discreteness does matter, although they do not address directly its effects on the moments of observed stock price changes or returns. Closely related to price discreteness is the issue of stock price clustering discussed by Harris (1989): the empirical observation that stock prices of whole or half dollars occur much more frequently than do quarters and odd eighths of dollars. Clustering is found to increase at higher stock price levels, indicating that the "natural" tick size may be greater than \$1/8 at higher prices. Clearly, this will also affect measured means and variances of returns.

The case that both these institutional factors affect moments of observed stock returns has been made quite convincingly in previous studies. Blume and Stambaugh (1983) demonstrate that bid-ask spreads affect observed mean returns. Roll (1984) and French and Roll (1986) show that serial correlation and variance of observed price changes are

[5] Certain foreign exchanges have adopted more flexible rules. For example, on the Toronto Stock Exchange, the minimum allowable price fluctuation ranges are as follows: one half-cent for stocks selling below 50 cents, 1 cent for stocks trading between 50 cents and \$3, 5 cents for stocks in the \$3–\$5 range, and \$1/8 thereafter. Further, unlike the NYSE, where a round lot consists of 100 shares, round lots on the Toronto exchange range from 10 shares for stocks selling over \$100 to 1000 and 500 shares for those below 10 cents and \$1 respectively. Even on the NYSE, the tick size is \$1/16 for a stock price below \$1, and tick sizes as low as \$1/32 are permitted for some other types of securities.

both affected by spreads. Price discreteness is not addressed by these authors. Gottlieb and Kalay (1985) examine the effects of price discreteness (but not bid-ask spreads) on the variance and higher moments of price changes. They do not consider its influence on mean returns or serial correlation. Harris (1990) is closest to this study. Harris analyzes the effects of both spreads and discreteness on the serial correlation and variance of observed stock price changes (rather than returns).

This paper synthesizes and extends the studies of all the authors cited above. We derive robust approximations for the effects of bid-ask spreads and price discreteness on the moments of observed stock returns. The analysis is carried out in the context of a simple model, very similar to that of Harris (1990). In order to motivate this model, we first present some graphic examples of the effects of institutional factors on the distributions of observed stock returns, particularly for low-priced stocks.

B. An Illustration

We analyze an AMEX firm, National Health Enterprises, over the six year period between April 1974 and March 1980. Figure 1 shows the histogram of daily returns on this stock for the subperiod from April 1977 to March 1980. Although zero returns predominate, the distribution appears to be continuous and approximately normal or bell-shaped. In Figure 2, we present the histogram for the April 1974 to March 1977 subperiod.

Insert Figures 1 and 2 here

Differences between the two figures are immediately apparent. Overall, we find that any semblance of normality has disappeared. Figure 2 shows a higher frequency of zero returns as well as some discontinuities: for example, there are no (non-zero) returns between -2% and $+3\%$. Returns of large (absolute) magnitude are much more frequent, so that the distribution has fatter tails. In order to quantify these differences between the returns for the two subperiods, we compare some sample moments. We find that the returns in Figure 2 are characterized by a higher mean, standard deviation and negative first-order autocorrelation. After the reverse split, the annualized mean return and standard deviation are 42.5% and 41% respectively, in contrast to pre-split values of 67.5% and 123%. Pre-split returns exhibit fairly strong negative first-order autocorrelation (−0.26), while the autocorrelation coefficient for post-split returns is 0.05. Recall that we are comparing distributions of returns on the same security during contiguous three-year subperiods. Their divergence may suggest that the firm experienced some major change between the two subperiods, which significantly altered its operating and risk characteristics.

In fact, April 1, 1977 was the ex-date of a 5–for–1 reverse split.[6] Figures 1 and 2 exhibit the distributions of returns before and after the ex-date of this reverse split. According to financial theory, the ex-date for a split (or reverse split) should be a "non-event:" an increase (decrease) in the total number of shares outstanding should be offset by a proportional decrease (increase) in the price per share so that the market value of the equity and shareholders' claims remain unchanged. There are at least three possible explanations for our results. First, investors may have interpreted the reverse split as a signal by the firm's managers about some private information.[7] However, any such reaction should have been observed at the announcement date of the split rather than at the ex-date, which is known to investors when the split is declared. Nor is it clear what type of investor reaction would cause the observed changes in the return distribution. This leads to a second possible explanation: market inefficiency or investor irrationality. Clearly, this is an unacceptable alternative to most financial economists, since these factors could be used to "explain" virtually any empirical result. Finally, and perhaps most plausibly, differences between the two distributions may simply be a result of the large change in the trading price level of the stock at the ex-date of the reverse split. Institutional factors such as bid-ask spreads and price discreteness are obviously much less important at higher prices and their differential effects could well explain the findings.[8]

[6] We confirmed from the *Wall Street Journal* that no other significant announcement or event took place on or around this date.

[7] Signalling models of this type are discussed by Grinblatt, Masulis and Titman (1984), Brennan and Copeland (1988a) McNichols and Dravid (1990) and others.

[8] We repeated this experiment for other stocks that reverse-split, and obtained very similar results. Somewhat weaker results were obtained for some large stock splits, for two

Figures 1 and 2 represent stock returns for a firm that actually declared a reverse split. In Appendix A, we describe an experiment in which we simulate splits for a high-priced stock in an attempt to replicate these results. As we increase the size of the split, thereby lowering the trading price level, we find that the stock return histograms follow a pattern similar to that in Figures 1 and 2. Departures from any resemblance to normality become quite striking. The frequency of zero returns increases, as does the frequency of large absolute returns. In addition, the distributions begin to look more discontinuous. In terms of moments, we see the same pattern as before: the mean, standard deviation and negative serial correlation all increase as the price level falls.[9]

Based on the evidence cited above, we claim that distributions of returns on stocks that trade at low price levels are generally discontinuous and leptokurtic. Further, returns will exhibit negative serial correlation. When such a low-priced stock undergoes a reverse split (actual or simulated) and trades at higher price levels, its return distribution begins to resemble a normal distribution more closely. The new distribution of returns has a lower mean and standard deviation, and returns tend to be less negatively correlated than before. Conversely, when a stock that trades at high price levels undergoes a split (actual or simulated), the distribution of returns follows the reverse pattern. These differences in return distributions are clearly a function of the price level of the stock, and our simulations show that they can be attributed to the effects of the bid-ask spread and price discreteness.

reasons: even the largest stock splits have smaller (absolute) split factors (up to 3–for–1) relative to typical reverse split factors of 1–for–5 or 1–for–10. Further, post-split trading price levels are generally much higher for splits than for reverse splits, which mitigates the differential pre- and post-split effects of institutional factors.

[9] The fact that we can generate simulated plots that resemble the *actual* distributions of returns on low-priced stocks is very interesting. Of course, in the real world, it is not at all obvious that prices at which transactions actually take place are in fact related to "frictionless" prices in the simplistic way that we have modeled them. Because investors are aware of the existence of price discreteness and bid-ask spreads, it is likely that they factor these institutional frictions into their trading decisions. The examples provided and our results in subsequent sections indicate that our naive model performs well in describing observed return distributions. The marginal benefits from more sophisticated market microstructure models may be small, at least in this limited context.

We now offer more formal evidence to support these claims.[10]

III. The Model

Using a simple econometric model that takes into account both the bid-ask spread and price discreteness, we now derive relations between the moments of observed returns and true underlying returns. The approach we take is similar in many respects to Harris (1990), where discrete bid-ask prices are modeled explicitly. Harris assumes that true stock prices follow a random walk. He uses maximum likelihood techniques to estimate the variance of true stock price changes (rather than returns) and the bid-ask spread. Harris does not analyze the effects of the spread and discreteness on expected values of returns. He obtains limiting results that are very close to the approximate results that we derive directly. Since we do not make any distributional assumptions about prices or returns, our results are more robust. Unlike Harris, we work with stock returns rather than price changes, since returns are usually of greater interest to most econometricians for reasons of stationarity. Our model also accommodates serially correlated true returns. Further, we model the bid-ask spread and price discreteness in a very general way that allows serial correlation in the spread term, as well as correlation between the spread and rounding terms. Finally, it is possible to extend our model heuristically to account for stock price clustering.[11]

A. Assumptions

We assume that true stock returns, defined in terms of the underlying true stock prices as

---

[10] We emphasize once again that most of these results can be found in the existing literature. Our contribution here has been mainly to synthesize these studies and to illustrate graphically the dramatic effects of spreads and discreteness on the entire return distribution for low-priced stocks.

[11] Harris (1990) discusses stock price clustering. A possible implication of clustering is that rounding may not always take place to the nearest $1/8, particularly at high prices. This can be approximated in our model by modifying the distribution of the rounding error term.

on p_t and q_t, and given the exact nature of the rounding process, u_t is known with certainty. Unconditionally, it is distributed uniformly on the interval $(-\frac{1}{16}, \frac{1}{16}]$, if we assume rounding to the nearest \$1/8.[14] In this case, q_t and u_t are independent, and each is independent of p_t. Further, our assumption that returns r_t are drawn exogenously implies that e_t and r_t are also independent and this fact is used in the subsequent analysis. Notationally, it is convenient to rewrite (2) as

$$\hat{p}_t = p_t + e_t, \tag{3}$$

where e_t is the "error" term representing the combined effects of the bid-ask spread and the rounding. Let $\rho(j) \equiv \text{cov}(e_t/p_{t-1}, e_{t-j}/p_{t-1})$ represent the serial covariance of e_t/p_{t-1} at lag $j \geq 0$. Using this notation, $\rho(0)$ represents the variance of the (percent) errors.

B. Results

We now derive approximate relations between the moments of observed returns, denoted by $\hat{r}_t = \hat{p}_t/\hat{p}_{t-1} - 1$ and the moments of true returns r_t, defined in (1). It is more convenient to work with the price relatives $1 + r_t$ and $1 + \hat{r}_t$. The following results are obtained (see Appendix B for details):

$$\mathcal{E}(1 + \hat{r}_t) \approx \mathcal{E}(1 + r_t)[1 + \rho(0)] - \rho(1), \tag{4a}$$

$$\text{var}(\hat{r}_t) \approx \text{var}(r_t) + \rho(0)\left(1 + [\mathcal{E}(1 + r_t)]^2\right) - 2\rho(1)\mathcal{E}(1 + r_t), \tag{4b}$$

[14] This result has been proved rigorously by Gottlieb and Kalay (1985) for the case of a lognormal diffusion process for stock prices. Their result is asymptotic ($t \to \infty$), but simulations show that departures from a uniform distribution are of no practical significance, even for very small values of t. For other reasonable rounding assumptions, such as discussed in the previous footnote, the rounding term is still uniform. Again, this can be verified by simple simulation experiments. The support of the distribution and the correlation between q_t and u_t will depend on the assumption made about the rounding process.

and

$$\text{cov}(\hat{r}_t, \hat{r}_{t-1}) \approx \phi(1) - \rho(0)[\mathcal{E}(1+r_t)]^2 + \rho(1)\mathcal{E}(1+r_t)\left(1 + (\text{var}(r_t) + [\mathcal{E}(1+r_t)]^2)\right) - \rho(2)\left(\text{var}(r_t) + [\mathcal{E}(1+r_t)]^2\right). \tag{4c}$$

Similar results obtained by other authors are now shown to be special cases of the above. First, some comments are in order. In (4a)–(4c), moments of observed returns are expressed as functions of the moments of true returns and the error terms. The importance of the stock price level is immediately obvious: as price increases, terms such as $\rho(j)$ become smaller, and so does the difference between true and observed moments. Next, it is possible to derive expressions for higher order moments and autocovariances at higher lags. However, the difference between these higher moments of true and observed returns is negligible, even at moderate price levels. Finally, while the Taylor series approximations are necessary when considering returns, it is possible to replicate the entire analysis much more easily for price changes. This is done in Section IV.

1. Expected Value of Returns

Let us ignore price discreteness and assume that the observed price at time t is equally likely to be a bid or an ask, independent of past transactions. Then, $\rho(1) = 0$ and relation (4a) can be written as

$$\mathcal{E}(\hat{r}_t) \approx \mathcal{E}(r_t) + \rho(0) + \rho(0)\mathcal{E}(r_t). \tag{4d}$$

Ignoring the last term, which is negligible in magnitude compared to the first two, this reduces to (6) from Blume and Stambaugh (1983):

$$\mathcal{E}(\hat{r}_t) \approx \mathcal{E}(r_t) + \text{var}\left(\frac{q_{t-1}s}{p_{t-1}}\right). \tag{4d}$$

Since $\text{var}(q_{t-1}s/p_{t-1}) > 0$, (4d) implies that the expected value of observed returns will be greater than the true expected value. This is the "bid-ask bias" of Blume and Stambaugh (1983) that will be discussed at greater length in Section V.B.

2. Serial Covariance of Returns

Roll (1984) shows that observed returns will exhibit negative first-order autocorrelation if observed prices bounce between bids and asks. Under the same assumptions as in the previous subsection, and the fact that $\mathcal{E}(r_t) = 0$ and $\phi(1) = 0$ in Roll's model, our relation (4c) becomes

$$\begin{aligned} \operatorname{cov}(\hat{r}_t, \hat{r}_{t-1}) &\approx \phi(1) - \rho(0)[\mathcal{E}(1 + r_t)]^2 \\ &\approx -\operatorname{var}\left(\frac{q_{t-1}s}{p_{t-1}}\right), \end{aligned} \tag{4f}$$

which is Roll's result. Roll finds that his simple expression for imputing the spread from the first-order autocovariance of observed returns fails almost half the time since it involves the square root of a negative number whenever the observed autocovariance is positive. Our relation (4c) shows that the first-order autocovariance of observed returns can be positive, for example, when $\phi(1)$ is positive and bigger than $\rho(0)[\mathcal{E}(1 + r_t)]^2$. Further, as we show in Section IV below, our model can explain non-zero higher-order autocovariances of observed returns, which Roll's model cannot.

3. Variance of Returns

French and Roll (1986) demonstrate that the variance of observed returns is affected by bid-ask spreads. Under the same assumptions as in the previous subsection, relation (4b) reduces to

$$\begin{aligned} \operatorname{var}(\hat{r}_t) &\approx \operatorname{var}(r_t) + 2\rho(0) \\ &\approx \operatorname{var}(r_t) - 2\operatorname{cov}(\hat{r}_t, \hat{r}_{t-1}), \end{aligned} \tag{4e}$$

which is the French and Roll result. The variance of observed returns is greater than the variance of true returns, a natural consequence of adding noise (in the form of the bid-ask spread and price discreteness) to the price process.

Gottlieb and Kalay (1985) study the implications of price discreteness for estimating the variance and higher order moments of returns. They conclude that observed moments are biased upward relative to the moments of true returns. If we ignore bid-ask spreads, as they do, the predictions of our equation (4b) for the variance are very close to theirs.

IV. Price Changes Versus Returns

The preceding analysis established relations between the moments of true and observed returns. As shown by Roll (1984), Harris (1990) and others, it is much more tractable to work with stock price changes rather than returns. We show next that relations between true and observed price changes can be obtained quite easily, and can be viewed as minor modifications of results (4a)–(4c).

A. A General Model of Price Changes

We begin with the same model as before:

$$\hat{p}_t = p_t + e_t, \tag{3}$$

where e_t is the error term consisting of the bid-ask spread and rounding. Instead of returns or price relatives, we now examine observed price changes defined by

$$\Delta\hat{p}_t \equiv \hat{p}_t - \hat{p}_{t-1} = \Delta p_t + \Delta e_t, \tag{5}$$

where Δp_t and Δe_t are defined analogously. First, we note that the expected observed price change will be equal to the expected true price change provided $\Delta e_t = 0$. The following expression for the covariance of the observed price changes is easily obtained:

$$\begin{aligned} \text{cov}(\Delta\hat{p}_t, \Delta\hat{p}_{t-j}) &= \text{cov}(\Delta p_t + \Delta e_t, \Delta p_{t-j} + \Delta e_{t-j}) \\ &= \phi(j) + \text{cov}(e_t - e_{t-1}, e_{t-j} - e_{t-j-1}) \\ &= \phi(j) - \rho(j-1) + 2\rho(j) - \rho(j+1), \end{aligned} \tag{6}$$

where $\phi(j)$ and $\rho(j)$ are now appropriately redefined in terms of price changes. The following expressions for the variance and first-order autocovariance are special cases of equation (6):

$$\text{var}(\Delta\hat{p}_t) = \text{var}(\Delta p_t) + 2\rho(0) - 2\rho(1), \tag{7}$$

$$\text{cov}(\Delta\hat{p}_t, \Delta\hat{p}_{t-1}) = \phi(1) - \rho(0) + 2\rho(1) - \rho(2). \tag{8}$$

Note that these expressions for observed price changes are very similar to equations (4b) and (4c) for observed returns. For transactions data or for daily data over short periods, the price level of the stock generally changes very little. When returns are measured over such intervals, it is more convenient to use the simpler expressions for price changes and then simply divide through by the (approximately constant) stock price level to obtain approximate results for returns. Evidence from simulations indicates that such approximations are quite accurate for practical purposes.

B. Special Cases

In order to obtain more concrete predictions of our model, we need to make some assumptions about the transaction price process. Let us assume that the conditional probability of a bid (ask) price being observed at time t depends on the price observed at time $t-1$ as follows:

$$P(q_t = 1 | q_{t-1} = 1) = P(q_t = -1 | q_{t-1} = -1) = \pi, \tag{9}$$

that is, the probability of a continuation is π while the probability of a reversal is $1-\pi$. The unconditional probabilities of observing a bid or an ask price at any time t are assumed to be equal. In this framework, it can be shown that

$$\rho(j) \equiv \text{cov}(e_t, e_{t-j}) = (2\pi - 1)^j \rho(0) \qquad \forall j \geq 0. \tag{10}$$

Substituting (10) in (6) results in a general expression for the autocovariance at lag j:[15]

$$\text{cov}(\Delta\hat{p}_t, \Delta\hat{p}_{t-j}) = \phi(j) - 4(1-\pi)^2(2\pi - 1)^j \rho(0). \tag{11}$$

[15] Equation (11) does not hold when $\pi = 0.5$, the standard assumption that q_t is serially uncorrelated. In that case, equations (7) and (8) can be used with $\rho(j) = 0 \; \forall j \geq 1$. These are just the well-known French and Roll (1986) and Roll (1984) expressions.

Substituting (10) in (7) and (8), we get

$$\text{var}(\Delta\hat{p}_t) = \text{var}(\Delta p_t) + 4\rho(0)(1-\pi) \tag{12}$$

$$\text{cov}(\Delta\hat{p}_t, \Delta\hat{p}_{t-1}) = \phi(1) - 4(1-\pi)^2(2\pi - 1)\rho(0). \tag{13}$$

As discussed earlier, the assumption of Roll (1984), Harris (1990) and others that the bid-ask indicator variable q_t is serially uncorrelated implies that observed price changes or returns will exhibit negative first-order autocorrelation, but will be uncorrelated at higher lags. Empirically, this prediction is not borne out, and French and Roll (1986) and others have reported non-zero higher-order autocovariances. Let us assume that true returns are serially uncorrelated at all lags ($\phi(j) = 0\ \forall j \geq 1$). From equation (11), we see that autocovariances at all lags will be negative if continuations are more likely than reversals ($\pi > 0.5$). If reversals are more likely, autocovariances will alternate in sign.

V. Numerical Results and Implications

We have obtained approximate analytical relations for differences between the moments of observed and true returns and price changes. It is of some interest to assess the practical and economic significance of these deviations, and their dependence on the stock price level, magnitude of the bid-ask spread and the exact nature of the rounding process. By establishing these, we shall also show that some anomalous results reported in the literature can be explained at least partially by our model.

A. Numerical Results

In Panel A of Table 1, we compute the difference between true and observed moments at various stock price levels, based on reasonable numerical values assigned to the spread. Spreads range between 3 cents (3%) for a stock trading at $1 and 30 cents (0.6%) for a $50 stock. These estimates are generally close to observed spreads (for example, Stoll and

Whaley (1983)). We assume that rounding is always to the nearest \$1/8, and that bid and ask prices are always equally likely so that $\rho(j) = 0 \; \forall j \geq 1$. True returns are assumed to be serially uncorrelated. The expected value and standard deviation of true (daily) returns are set at 0.05% and 0.025% respectively.[16] The table shows the effects of the spread and rounding to be extremely large at low price levels. For example, at a stock price of \$1, the expected value of observed returns is 0.203%, which is four times higher than the true value. As stated earlier, similar results were first proved by Blume and Stambaugh (1983) and proposed as an explanation for the non-January 'size' effect. This is discussed in the next subsection. The standard deviation of observed returns is also much higher than the true standard deviation: even for a stock price as high as \$10, the observed standard deviation is 12% higher. This suggests an explanation for the Ohlson and Penman (1985) results, which will also be discussed in the next subsection. Finally, although true returns are serially uncorrelated, the first-order autocorrelation of observed returns is found to be negative, ranging from −0.41 to −0.01.[17]

In Panel B of Table 1, we examine the implications of an alternative rounding specification. We assume now that true prices are rounded down for bids and rounded up for asks. The consequence of this assumption is perfect correlation between q_t and u_t, and a change in the support of the distribution of u_t from $(-\$1/16, \$1/16]$ to $(-\$1/8, \$1/8]$. As a result, deviations between true and observed moments become much more significant.[18]

[16] These translate into annualized values of 12.5% and 40% respectively.

[17] In all these calculations, $\rho(0) \equiv \text{var}(e_t/p_{t-1}) = \text{var}[(q_t s + u_t)/p_{t-1}]$ is computed as the sum of the variance of the spread term (s^2) and the variance of u_t, which is distributed uniformly $(-\$1/16, \$1/16]$, divided by the squared price.

[18] The assumptions in Panels A and B represent two extreme cases: prices are always rounded either to the nearest tick (Panel A), or in favor of the market-maker (Panel B), respectively. In fact, what we observe empirically may lie between these two extremes. It is possible to extend our model by assigning a probability to each type of rounding specification.

of returns at the announcement date. They investigate several potential explanations for these anomalous findings, but are unable to provide a satisfactory answer.[20] More recently, several authors including Lamoureux and Poon (1987), Brennan and Copeland (1988b), Sheikh (1989) and Conroy, Harris and Benet (1990) have proposed explanations for this post-split volatility change. Conroy, Harris and Benet provide evidence that relative bid-ask spreads increase after splits, and propose this as an explanation for changes in post-split volatility.[21] The bid-ask spread and price discreteness, which are obviously more important at low prices, result in an increase in the variance of returns after the ex-date, at which there is a large change in the price level of the stock. The increase in standard deviations of returns shown in Table 1 are of about the same order of magnitude as those reported by Ohlson and Penman.

Our explanation that post-split volatility changes are at least partly spurious is consistent with financial theory, which views the ex-date of a stock split as a "non-event:" the increase in the number of shares outstanding is offset by a corresponding price decrease, resulting in no real change. Even in the context of signalling models for stock splits, we should expect a change in the variance of returns, if any, to take place at the announcement date of the split rather than at the ex-date. Ohlson and Penman do not detect any permanent increase subsequent to the announcement date. This finding is consistent with our model, since there is no permanent and large change in the stock price level at the announcement date as there is at the ex-date. Our model also suggests that reverse splits should exhibit the opposite effect — the variance of returns should fall at the ex-date — and this finding was reported by Dravid (1987).[22]

---

[20] Ohlson and Penman themselves suggest (see their Section 4.7 and footnote 13) that discreteness and changes in relative spreads may be at least partially responsible for their findings. Amihud and Mendelson (1987) make the same suggestion.

[21] Some of their results appeared in an earlier version of this paper (Dravid (1989)), which provides evidence on this issue using simulations as well as empirical tests using transaction data.

[22] However, Dravid's result for stock dividends, where a decrease in variance is observed subsequent to the ex-date, cannot be explained by our model and remains an unresolved finding that warrants further investigation.

VI. Conclusion

This paper examines the effects of institutional factors on observed stock returns. Using a framework similar to Harris (1990), we show that bid-ask spreads and price discreteness can and should be analyzed simultaneously. We synthesize and extend the results of Blume and Stambaugh (1983), Roll (1984), Gottlieb and Kalay (1985), French and Roll (1986) and Harris (1990). Our results largely explain the "anomalous" findings of Ohlson and Penman (1985) and Dravid (1987). We confirm Blume and Stambaugh's (1983) suggestion that most of the average non-January "size" effect is attributable to the effects of bid-ask spreads (and price discreteness) on returns measured at low price levels. By incorporating serial correlation in true returns and in the spread term, we can account for two empirical facts that simpler models such as Roll's cannot: positive first-order autocorrelation and non-zero higher order autocorrelations in observed returns.

Our study raises some concerns about the use of intraday stock price data. The availability of these databases has created new areas for research. For example, event studies can now be conducted using the exact time of the event during the trading day. However, problems such as the "bid-ask bounce" are clearly more severe when dealing with transaction data, and must be taken into consideration. The implications of stock price discreteness in the context of intraday data have not yet been fully explored. Hausman, Lo and MacKinlay (1991) examine some issues related to price discreteness that are not addressed here, and suggest areas for future research. The related issue of stock price clustering discussed by Harris (1989) also needs to be investigated in much more detail.

Appendix A: Simulation of Stock Splits

Here, we describe the simulation experiment discussed in Section II.B, designed to demonstrate the effects of the bid-ask spread and price discreteness on stock returns measured at low trading prices. Consider Figure 3, which shows the histogram of daily returns on an NYSE stock (Du Pont) over a three-year period during which the stock traded in the \$40–\$80 range. First, we demonstrate that the effect of the bid-ask spread on returns is negligible at such high price levels. We transform each price to a simulated bid or ask price by randomly (with equal probability) adding or subtracting \$0.125, which is half the assumed bid-ask spread.[23] Returns are calculated using these simulated prices, and their histogram is plotted in Figure 4. We find that there is no appreciable difference between the two figures: both distributions appear continuous and (approximately) normal. The moments of the two sets of returns are also found to be virtually identical.

Insert Figures 3 and 4 here

Next, we simulate a 2-for-1 split of this stock by dividing each observation in the original price series by a factor of 2. We assume that the bid-ask spread is \$0.25 both before and after the split.[24] Each post-split price is again transformed into a simulated bid or ask price as described earlier. Finally, the resulting price is rounded to the nearest \$1/8, and returns are calculated from this series of simulated post-split prices. In Figure 5, we present the histogram of these returns. Figures 6 and 7 show the corresponding distributions after 5-for-1 and 10-for-1 splits respectively.

Insert Figures 5, 6, 7 and 8 here

[23] For the purpose of this illustration, we are ignoring the fact that these are already bid or ask prices, rounded to the nearest \$1/8, and have assumed that the observed prices are in fact true prices. Since the stock price level is very high, this is a reasonable approximation.

[24] This assumption is not entirely realistic, since the post-split spread is likely to be lower than \$0.25. However, our results remain qualitatively the same, as long as there is a change in the percent spread.

and

$$(A3) \qquad \mathcal{E}\left(\frac{X^2}{Y^2}\right) \approx \frac{[\mathcal{E}(X)]^2}{[\mathcal{E}(Y)]^2} + \frac{\text{var}(X)}{[\mathcal{E}(Y)]^2} + 3\,\text{var}(Y)\frac{[\mathcal{E}(X)]^2}{[\mathcal{E}(Y)]^4} - 4\,\text{cov}(X,Y)\frac{\mathcal{E}(X)}{[\mathcal{E}(Y)]^3}.$$

The following assumptions, which have been discussed in Section III.A, are used frequently in the subsequent derivations:

$$(A4) \qquad \mathcal{E}\left(\frac{e_t}{p_{t-j}}\right) = 0 \qquad \forall j \geq 0,$$

$$(A5) \qquad r_t \text{ and } \frac{e_s}{p_{s-j}} \text{ are independent} \qquad \forall s \leq t, \qquad \forall j \geq 1.$$

Equation (A4) reflects the fact that the rounding error is distributed uniformly on $(-\$1/16, \$1/16]$ and the assumption that the underlying value of the stock is the average of the unrounded bid and ask prices. (A5) is simply the assumption that true returns are drawn exogenously so that the bid-ask and rounding error terms at any time are independent of the true return. Recall from Section III.A that we defined:

$$\phi(j) \equiv \text{cov}\left(\frac{p_t}{p_{t-1}}, \frac{p_{t-j}}{p_{t-j-1}}\right) \qquad \forall j \geq 1,$$

$$\rho(k) \equiv \text{cov}\left(\frac{e_t}{p_{t-1}}, \frac{e_{t-k}}{p_{t-1}}\right) \qquad \forall k \geq 0.$$

We now derive results (4a)–(4c) relating the moments of $\hat{r}_t$ and r_t. It is convenient to work with the price relatives $r_t \equiv p_t/p_{t-1}$ and $\hat{r}_t \equiv \hat{p}_t/\hat{p}_{t-1}$. For the expected return, we have from (3):

$$\begin{aligned}
\mathcal{E}(1+\hat{r}_t) &= \mathcal{E}\left(\frac{p_t + e_t}{p_{t-1} + e_{t-1}}\right) \\
&= \mathcal{E}\left(\frac{p_t/p_{t-1} + e_t/p_{t-1}}{1 + e_{t-1}/p_{t-1}}\right) \\
&= \mathcal{E}\left(\frac{1+r_t}{1+e_{t-1}/p_{t-1}}\right) + \mathcal{E}\left(\frac{e_t/p_{t-1}}{1+e_{t-1}/p_{t-1}}\right) \\
&\approx \mathcal{E}(1+r_t) + \mathcal{E}(1+r_t)\text{var}\left(\frac{e_{t-1}}{p_{t-1}}\right) + \text{cov}\left(\frac{e_t}{p_{t-1}}, \frac{e_{t-1}}{p_{t-1}}\right) \\
(4a) \qquad &= \mathcal{E}(1+r_t)[1+\rho(0)] - \rho(1),
\end{aligned}$$

where we have used results (A2), (A4) and (A5).

In order to derive result (4b) for the variance of $\hat{r}_t$, we use results (A3)–(A5) to obtain:

$$\begin{aligned}
\mathcal{E}[(1+\hat{r}_t)^2] &= \mathcal{E}\left(\frac{(p_t+e_t)^2}{(p_{t-1}+e_{t-1})^2}\right) \\
&= \mathcal{E}\left(\frac{(p_t/p_{t-1}+e_t/p_{t-1})^2}{(1+e_{t-1}/p_{t-1})^2}\right) \\
&\approx [\mathcal{E}(1+r_t)]^2+\operatorname{var}(r_t)+\rho(0)+3\rho(0)[\mathcal{E}(1+r_t)]^2-4\rho(1)\mathcal{E}(1+r_t).
\end{aligned} \tag{A6}$$

Taking squares on both sides of (4a), we get:[27]

$$\begin{aligned}
[\mathcal{E}(1+\hat{r}_t)]^2 &= (\mathcal{E}(1+r_t)[1+\rho(0)])^2-2[1+\rho(0)]\rho(1)\mathcal{E}(1+r_t)+[\rho(1)]^2 \\
&\approx [1+2\rho(0)][\mathcal{E}(1+r_t)]^2-2\rho(1)\mathcal{E}(1+r_t).
\end{aligned} \tag{A7}$$

Using (A6) and (A7), we obtain the following expression for the variance:

$$\begin{aligned}
\operatorname{var}(\hat{r}_t) &= \operatorname{var}(1+\hat{r}_t) \\
&= \mathcal{E}[(1+\hat{r}_t)^2]-[\mathcal{E}(1+\hat{r}_t)]^2 \\
&\approx \operatorname{var}(r_t)+\rho(0)\left(1+[\mathcal{E}(1+r_t)]^2\right)-2\rho(1)\mathcal{E}(1+r_t).
\end{aligned} \tag{4b}$$

To derive a relation between covariances, we first need the expressions:

$$\begin{aligned}
\mathcal{E}\left(\frac{\hat{p}_t}{\hat{p}_{t-2}}\right) &\approx \phi(1)+[\mathcal{E}(1+r_t)]^2[1+\rho(0)(\operatorname{var}(r_t)+[\mathcal{E}(1+r_t)]^2)] \\
&\quad -\rho(2)(\operatorname{var}(r_t)+[\mathcal{E}(1+r_t)]^2),
\end{aligned} \tag{A8}$$

and

$$\begin{aligned}
\mathcal{E}\left(\frac{\hat{p}_{t-1}}{\hat{p}_{t-2}}\right) &\approx [\mathcal{E}(1+r_t)][1+\rho(0)(\operatorname{var}(r_t)+[\mathcal{E}(1+r_t)]^2)] \\
&\quad -\rho(1)(\operatorname{var}(r_t)+[\mathcal{E}(1+r_t)]^2).
\end{aligned} \tag{A9}$$

[27] In obtaining (A7), we ignore second and higher order terms in $\rho(0)$ and $\rho(1)$. For a stock with a price level of \$2.00 and bid-ask spread \$0.125, $\rho(0)$ is equal to 1.3×10^{-3}, while $\rho(0)^2=1.7\times10^{-6}$, which is small enough to ignore. For the same reason, we ignore terms involving higher powers of $\rho(1)$ as well as products of $\rho(0)$ and $\rho(1)$. If $\operatorname{var}(r_t)=0.001$, the value of $\operatorname{var}(\hat{r}_t)$ from the final approximation (4b) below is 0.003604, while the more precise expression, which accounts for $\rho(0)^2$, gives a value of 0.003603.

The derivation of the above expressions is similar to that of (4a). Using (A8) and (A9), the expression we obtain for the covariance is:

$$\begin{aligned}
\operatorname{cov}(\hat{r}_t, \hat{r}_{t-1}) &= \operatorname{cov}\left(\frac{\hat{p}_t}{\hat{p}_{t-1}}, \frac{\hat{p}_{t-1}}{\hat{p}_{t-2}}\right) \\
&= \mathcal{E}\left(\frac{\hat{p}_t}{\hat{p}_{t-2}}\right) - \mathcal{E}\left(\frac{\hat{p}_t}{\hat{p}_{t-1}}\right) \mathcal{E}\left(\frac{\hat{p}_{t-1}}{\hat{p}_{t-2}}\right) \\
&\approx \phi(1) - \rho(0)[\mathcal{E}(1+r_t)]^2 + \rho(1)[\mathcal{E}(1+r_t)]\left(1 + \left(\operatorname{var}(r_t) + [\mathcal{E}(1+r_t)]^2\right)\right) \\
&\quad - \rho(2)\left(\operatorname{var}(r_t) + [\mathcal{E}(1+r_t)]^2\right). \qquad \text{(4c)}
\end{aligned}$$

Garman, M. "Market Microstructure." *Journal of Financial Economics*, 3 (June 1976), 257–275.

Glosten, L. and P. Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics*, 14 (March 1985), 71–100.

Gottlieb, G. and A. Kalay. "Implications of the Discreteness of Observed Stock Prices." *Journal of Finance*, 40 (March 1985), 135–153.

Grinblatt, M.; R. Masulis; and S. Titman. "The Valuation Effects of Stock Splits and Stock Dividends." *Journal of Financial Economics*, 13 (Dec. 1984), 461–490.

Harris, L. "Stock Price Clustering, Discreteness Regulation, and Bid-Ask Spreads," Working paper, NYSE (1989).

Harris, L. "Estimation of Stock Price Variances and Serial Covariances from Discrete Observations." *Journal of Financial and Quantitative Analysis*, 25 (Sept. 1990), 291–306.

Hausman, J.; A. Lo; and A. C. MacKinlay. "An Ordered Probit Analysis of Transaction Stock Prices," Working paper, Alfred P. Sloan School of Management, M.I.T. (1991).

Keim, D. "Size Related Anomalies and Stock Return Seasonality: Further Empirical Evidence." *Journal of Financial Economics*, 12 (June 1983), 13–32.

Lamoureux, C. and P. Poon. "The Market Reaction to Stock Splits." *Journal of Finance*, 42 (Dec. 1987), 1347–1370.

McNichols, M. and A. Dravid. "Stock Dividends, Stock Splits, and Signalling." *Journal of Finance*, 45 (July 1990), 857–879.

Ohlson, J. and S. Penman. "Volatility Increases Subsequent to Stock Splits: An Empirical Aberration." *Journal of Financial Economics*, 14 (June 1985), 251–266.

Pessin, A. *Fundamentals of the Security Industry*. New York Institute of Finance, New York, NY (1985).

Roll, R. "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance*, 39 (Sept. 1984), 1127–1139.

Schwartz, R. and D. Whitcomb. "Evidence on the Presence and Causes of Serial Correlation in Market Model Residuals." *Journal of Financial and Quantitative Analysis*, 11 (June 1977), 291–313.

Sheikh, A. "Stock Splits, Volatility Increases and Implied Volatilities." *Journal of Finance*, 44 (Dec. 1989), 1361–1372.

Simmons, D. "Common-Stock Transaction Sequences and the Random-Walk Model." *Operations Research*, 19 (July-Aug. 1971), 845–861.

Stigler, G. "Public Regulation of the Securities Markets." *Journal of Business*, 37 (April 1964), 117–142.

Stoll, H. "The Supply of Dealer Services in Security Markets." *Journal of Finance*, 33 (Sept. 1978), 1133–1151.

Stoll, H. and R. Whaley. "Transaction Costs and the Small Firm Effect." *Journal of Financial Economics*, 12 (June 1983), 57–79.

Stoll, H. "Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests." *Journal of Finance*, 44 (March 1989), 115–134.

TABLE 1

Comparison between moments of true returns and moments of observed returns

Panel A: Rounding to nearest \$1/8

Stock Price	Bid-Ask Spread	$\rho(0)$	Expected Value (%)	Standard Dev. (%)	Auto-correlation
\$ 1.00	\$0.03	0.001527	0.203	0.061	−0.41
\$ 2.00	\$0.05	0.000481	0.098	0.040	−0.30
\$ 5.00	\$0.10	0.000152	0.065	0.031	−0.16
\$10.00	\$0.15	0.000069	0.057	0.028	−0.09
\$20.00	\$0.20	0.000028	0.053	0.027	−0.04
\$50.00	\$0.30	0.000009	0.051	0.026	−0.01
True moments			0.050	0.025	0.00

Panel B: Rounding bid prices down, ask prices up

Stock Price	Bid-Ask Spread	$\rho(0)$	Expected Value (%)	Standard Dev. (%)	Auto-correlation
\$ 1.00	\$0.03	0.007598	0.810	0.126	−0.48
\$ 2.00	\$0.05	0.002360	0.286	0.073	−0.44
\$ 5.00	\$0.10	0.000597	0.110	0.043	−0.33
\$10.00	\$0.15	0.000216	0.072	0.033	−0.20
\$20.00	\$0.20	0.000074	0.057	0.028	−0.09
\$50.00	\$0.30	0.000019	0.052	0.026	−0.03
True moments			0.050	0.025	0.00

Notes: Expected value and standard deviation of true (daily) returns are set at 0.05% and 0.025% respectively, and true returns are assumed to be serially uncorrelated ($\phi(j) = 0\ \forall j$). Moments of observed returns are computed from the assumed true moments, using equations (4a)–(4c). Bid and ask prices are assumed to be equally likely at all times, so that $\rho(j) = 0\ \forall j \geq 1$.

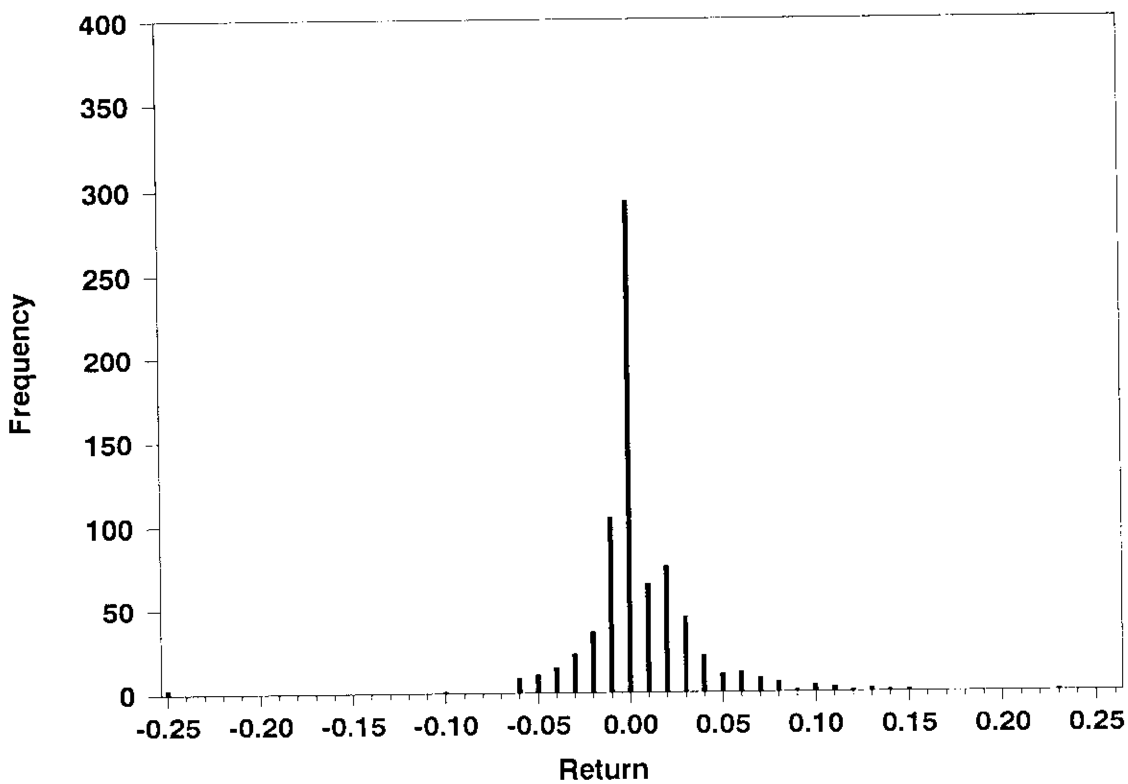

Figure 1. Frequency distribution of daily returns during three year period subsequent to 5-for-1 reverse split. (Firm: National Health Enterprises)

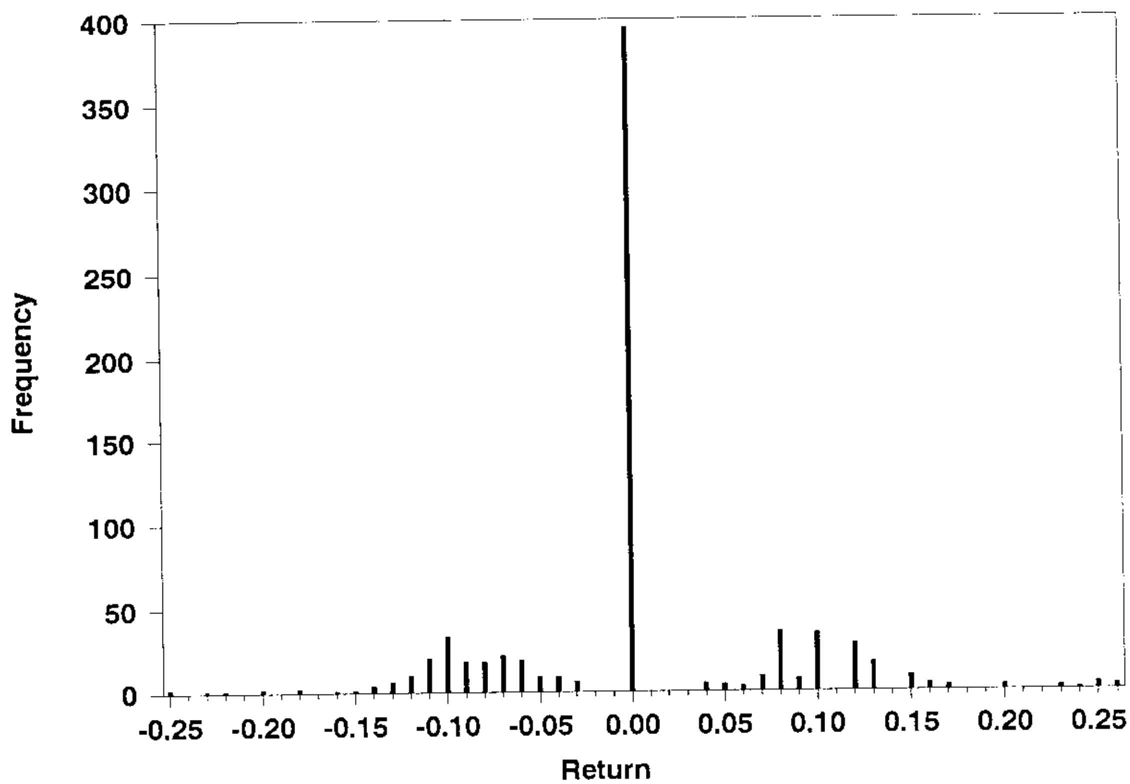

Figure 2. Frequency distribution of daily returns during three year period prior to 5-for-1 reverse split. (Firm: National Health Enterprises)

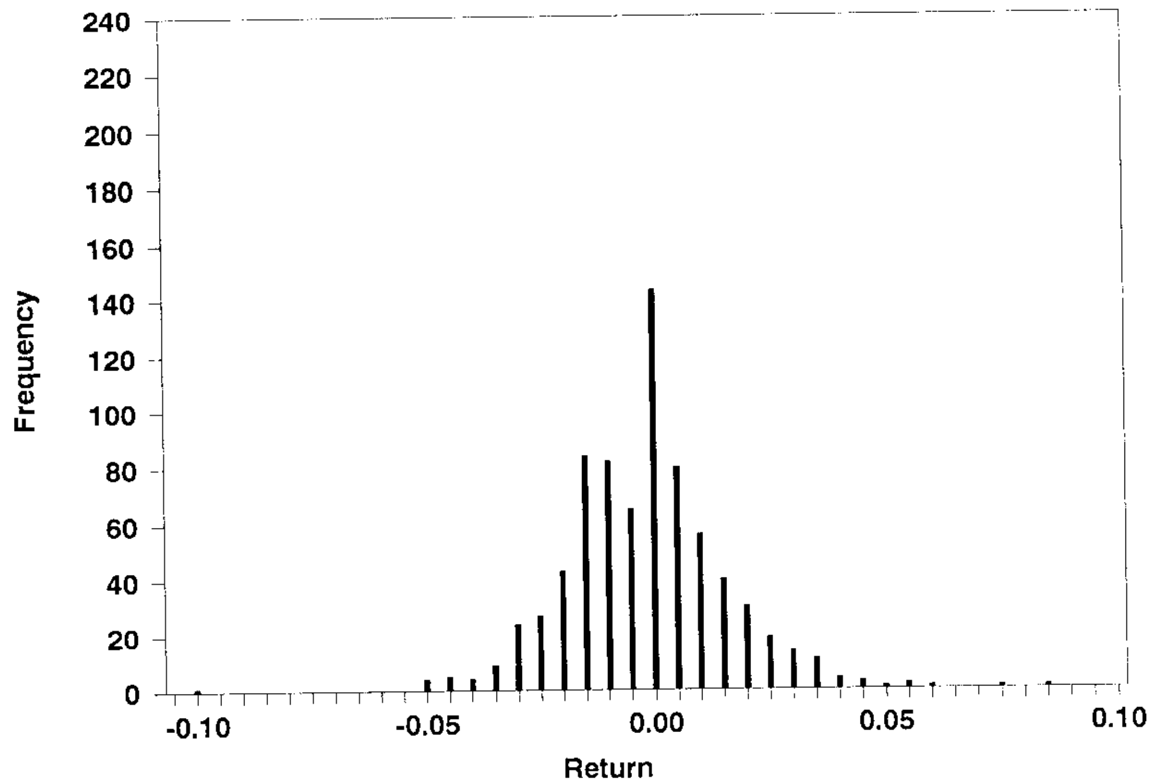

Figure 3. Frequency distribution of daily returns during three year period when stock traded in the \$40–\$80 range. (Firm: Du Pont)

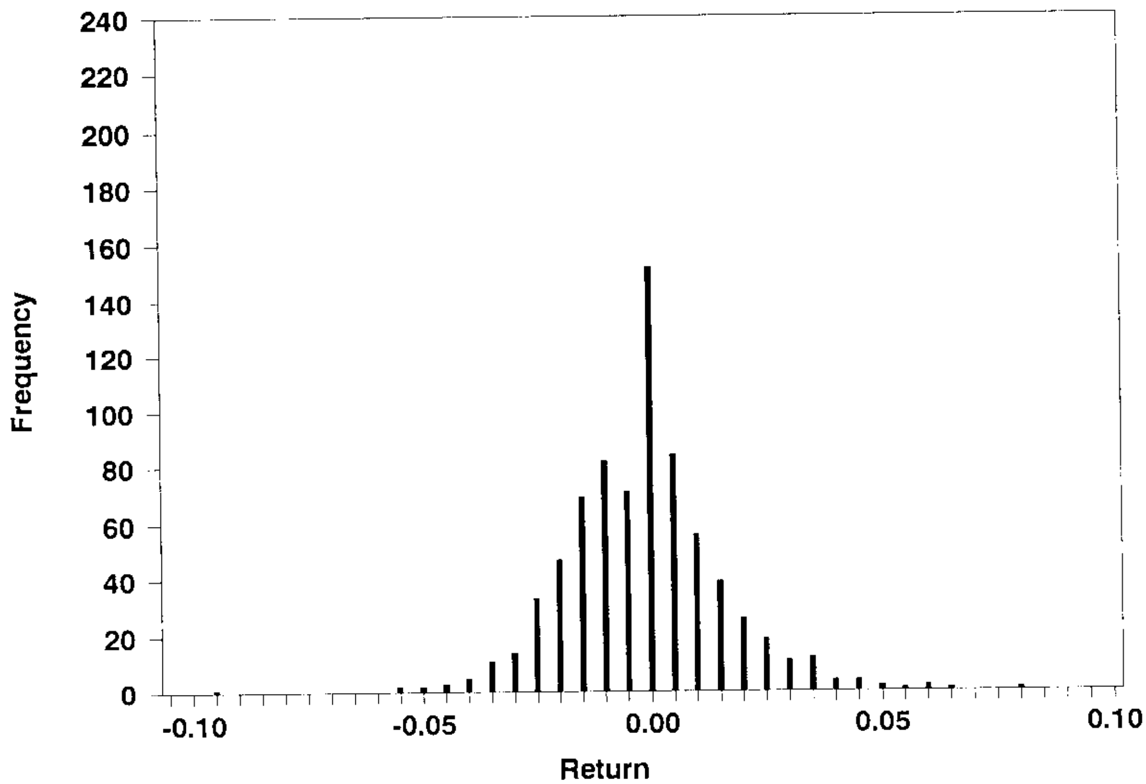

Figure 4. Frequency distribution of daily returns using simulated bid-ask prices. (Firm: Du Pont)

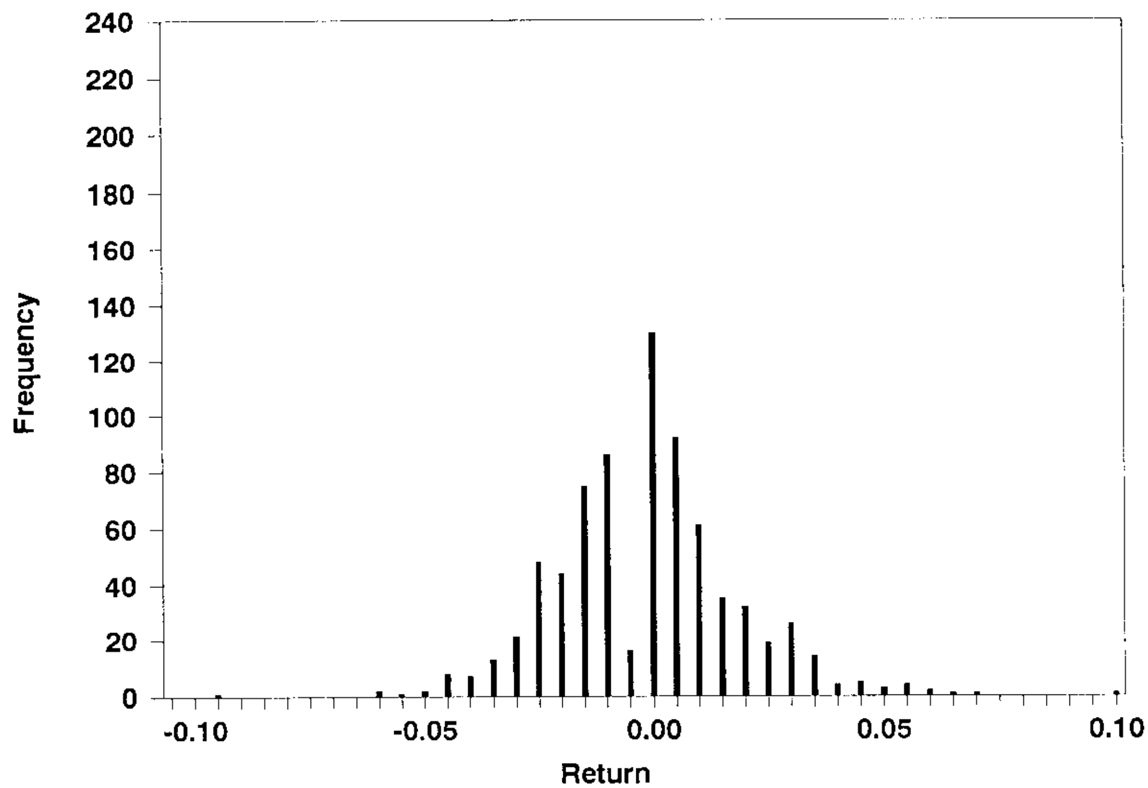

Figure 5. Frequency distribution of daily returns using simulated 2-for-1 post-split rounded bid-ask prices. (Firm: Du Pont)

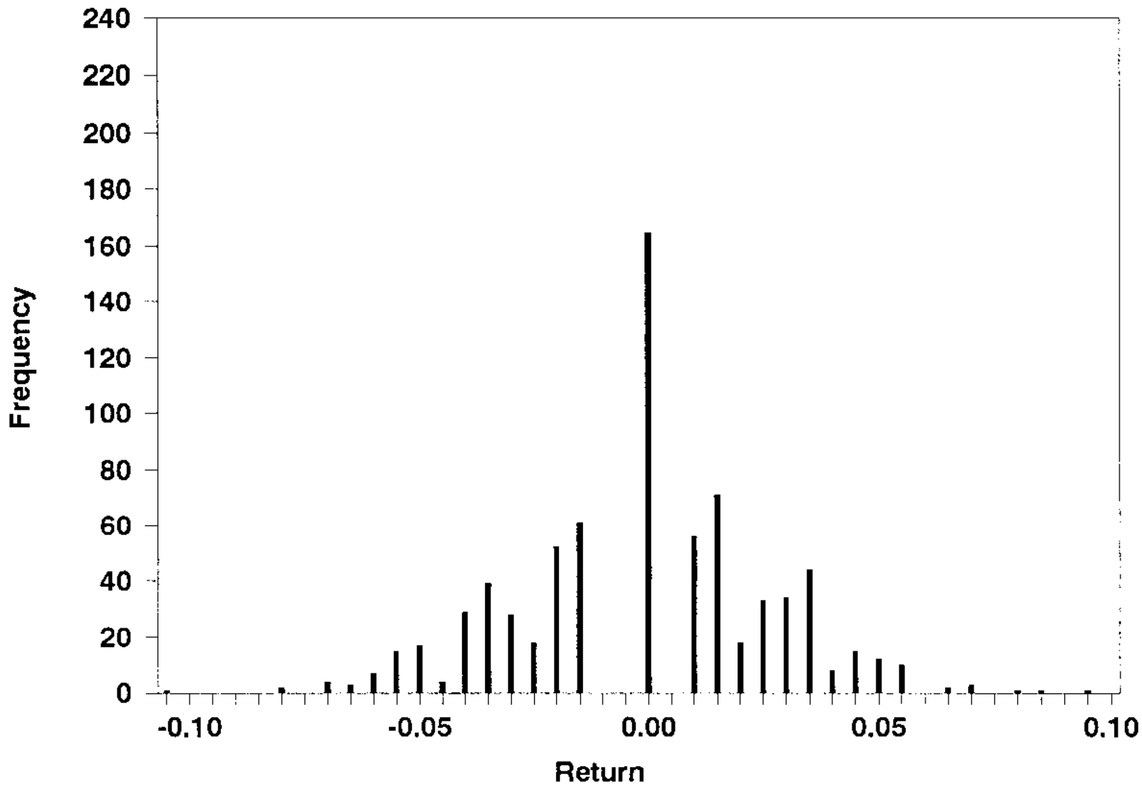

Figure 6. Frequency distribution of daily returns using simulated 5-for-1 post-split rounded bid-ask prices. (Firm: Du Pont)

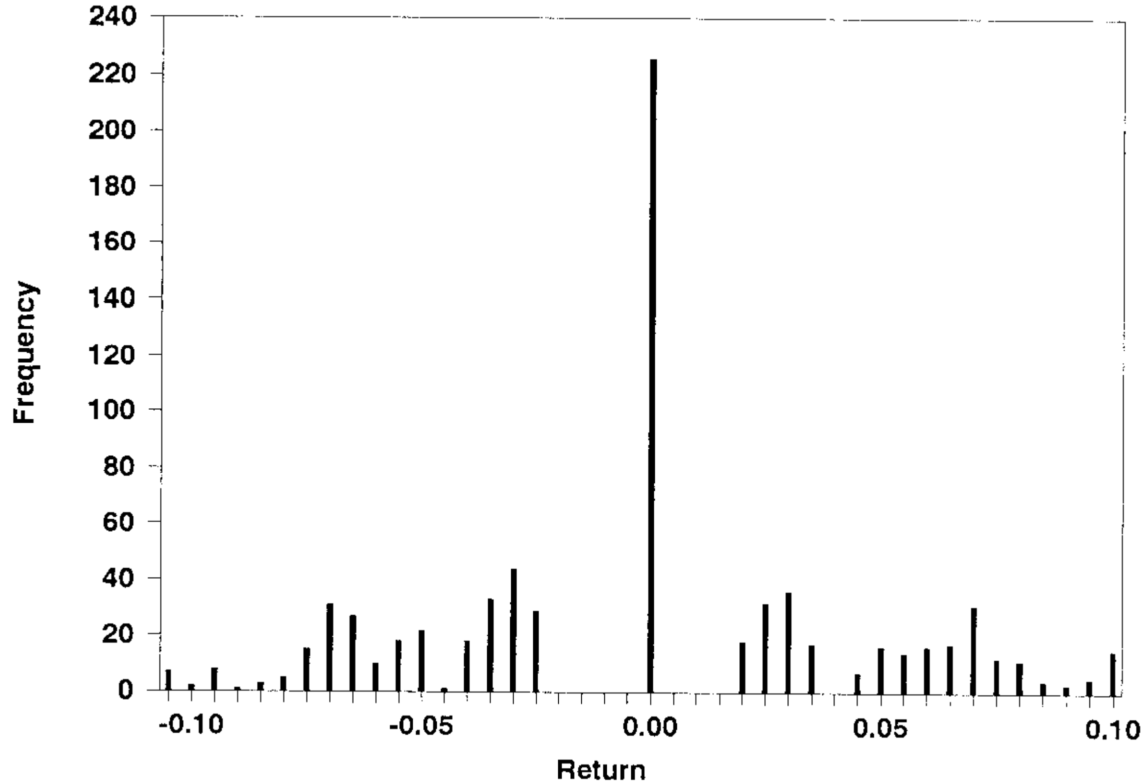

Figure 7. Frequency distribution of daily returns using simulated 10-for-1 post-split rounded bid-ask prices. (Firm: Du Pont)

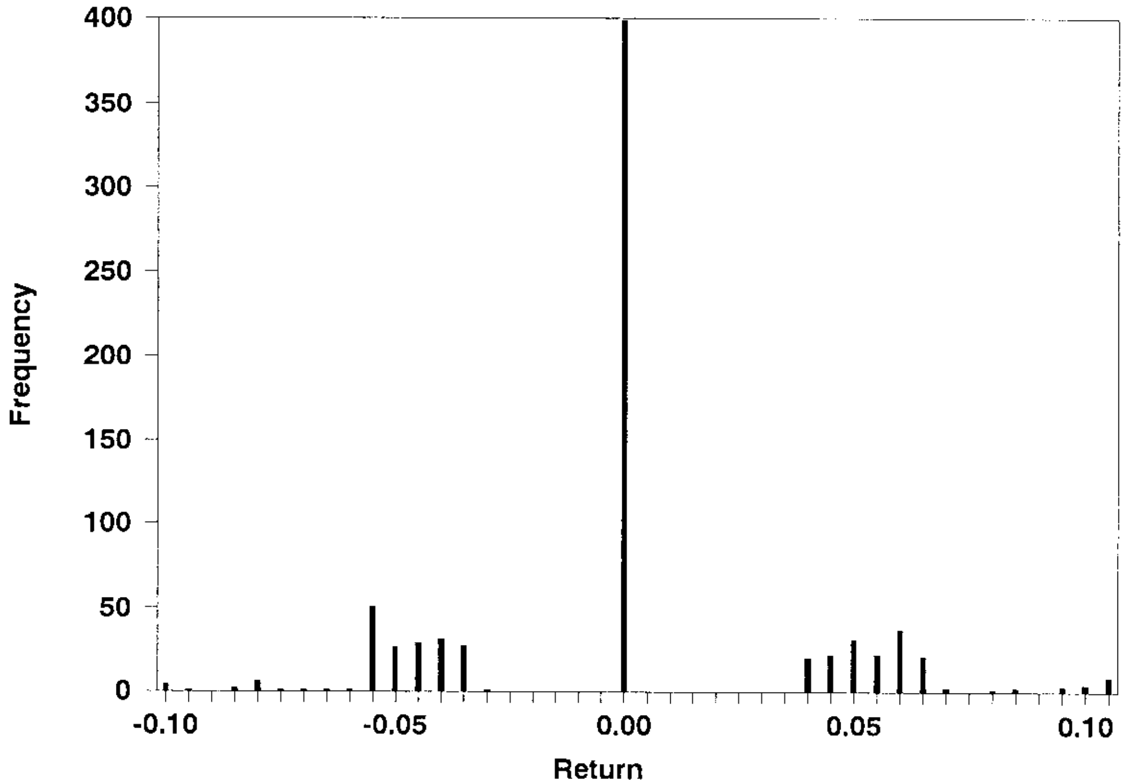

Figure 8. Frequency distribution of daily returns during three year period when stock traded below $5. (Firm: Ensource)