THE JOURNAL OF FINANCE • VOL. XLVIII, NO. 1 • MARCH 1993

Long-Term Market Overreaction or Biases in Computed Returns?

JENNIFER CONRAD* and GAUTAM KAUL

ABSTRACT

We show that the returns to the typical long-term contrarian strategy implemented in previous studies are upwardly biased because they are calculated by cumulating single-period (monthly) returns over long intervals. The cumulation process not only cumulates "true" returns but also the upward bias in single-period returns induced by measurement errors. We also show that the remaining "true" returns to loser or winner firms have no relation to overreaction. This study has important implications for event studies that use cumulative returns to assess the impact of information events.

RECENT RESEARCH HAS UNCOVERED substantial predictability in both short-term (Conrad and Kaul (1988, 1989) and Lo and MacKinlay (1988)) and long-term (Fama and French (1988) and Poterba and Summers (1988)) stock returns. An increasingly popular interpretation of return predictability emphasized by a number of researchers is that the stock market consistently overreacts to new information. The "stock market overreaction" hypothesis asserts that stock prices take temporary swings away from their fundamental values due to waves of optimism and pessimism (see, for example, DeBondt and Thaler (1985, 1987), Lehmann (1990), and Shefrin and Statman (1985)).

Compelling evidence in favor of long-term overreaction was first provided by DeBondt and Thaler (1985). They show that losers and winners, determined by their performance relative to the aggregate stock market over the past three to five years, consistently outperform and underperform the market in subsequent three- to five-year periods. For example, they find that the arbitrage (zero investment) portfolio of losers and winners earns an average "return" of 24.6% over a three-year period.[1] Similar findings have

*Conrad is from the Graduate School of Business, University of North Carolina, and Kaul is from the School of Business Administration, University of Michigan. We thank seminar participants at the Universities of Michigan, North Carolina, Oregon, Washington, and the Southern Methodist University, and especially Elazar Berkovitch, Thomas George, Charles Jones, Paul Malatesta, Dale Morse, M. Nimalendran, Jay Ritter, René Stulz, and an anonymous referee for their comments and suggestions. We also thank Patti Lamparter and Stacy Zimmerman for preparing this manuscript. Partial financial support for this research was provided by the School of Business Administration, University of Michigan, and the Graduate School of Business, University of North Carolina.

[1]The term "return" is loosely used for the zero-investment portfolio because it is actually not defined. Strictly speaking, the "returns" on these portfolios are profits on going long and short one dollar in losers and winners, respectively. Since these profits are the difference in returns on two (dollar) portfolios, they are measured in units of percent per unit time.

been reported in a number of subsequent papers that implement DeBondt and Thaler (1985)-type long-term contrarian strategies.

In this paper, we show that these long-term strategies suffer from a methodological drawback which could spuriously inflate their profitability. DeBondt and Thaler (1985) *cumulate* short-term (monthly) returns to loser and winner stocks over long periods (three to five years). For example, their measure of profitability of the arbitrage portfolio is the difference between the average cumulative raw returns (CRRs) of loser and winner securities. We show that measurement errors in observed prices due to bid-ask errors, nonsynchronous trading, and/or price discreteness, lead to substantial spurious returns to the long-term zero-investment contrarian strategies because single-period returns are upwardly biased (see Blume and Stambaugh (1983)). By *cumulating* short-term returns over long periods, these strategies cumulate not only the "true" short-term returns but also the upward bias in each of the single-period returns. And if loser firms are low-priced relative to winner firms, the return to the arbitrage portfolio (that is, long in losers and short in winners) will then have a spurious upward drift that is unrelated to market overreaction. This occurs because the upward bias in low-priced firms' single-period returns is substantially greater than the bias in the returns of high-priced firms.

Apart from this statistical bias, the long-term cumulative performance measure suffers from a conceptual drawback. Cumulating single-period returns over long intervals implicitly amounts to rebalancing the loser and winner portfolios to equal weights each month. The appropriate measure of performance should be the buy and hold return over long intervals. This measure has the additional advantage of minimizing transactions costs.

We, therefore, measure long-term performance of contrarian strategies using holding period returns of up to three years, that is, a buy and hold strategy. Apart from being conceptually consistent with the notion of long-term overreaction, this measure greatly reduces the statistical biases in previous cumulative performance measures. Specifically, for all k-period returns, our measure contains only a constant bias (the bias in a single period's return). This contrasts with *k times* the single-period return's bias in the cumulative k-period measure.

We use a sample of NYSE firms over the 1926 to 1988 period and show that previous CRR-based performance evaluation measures are substantially upwardly biased. For example, for non-January months the 36-month CRR method yields a return of 12.2% to the arbitrage portfolio. However, the appropriate holding period average return is -1.7%! In fact, contrary to the overreaction hypothesis the non-January evidence shows that losers consistently underperform winners for the one-, two-, and three-year evaluation periods. In January, losers outperform winners using both the CRR and buy and hold strategies. However, we provide evidence that the January returns to long-term contrarian strategies have no relation to overreaction.

An important contribution of this study is that it has potentially important implications for a number of empirical studies. Most event studies use cumulative abnormal returns (CARs) to evaluate the impact of information events on stock prices. In some cases, as Fama (1991) points out, "there is suspicious evidence that part of the response of prices to an information announcement occurs slowly over time" (p. 38). Two important cases of post-announcement drifts are the slow downward drift in the CARs of acquiring firms (Asquith (1983)) and the post-earnings drifts (see, for example, Bernard and Thomas (1989)). Brown and Warner (1980) caution against the use of CARs to assess performance because even if true returns are independently and identically distributed, the CAR "like any process which follows a random walk, ... can easily give the *appearance* of 'significant' positive or negative drift, when none is present" (p. 229). Our results suggest that there is another potential source of upward or downward spurious drift in CARs caused by cumulation of the bias in single-period returns.[2] Given that monthly cumulative returns used in overreaction studies contain a substantial bias, intraday and daily cumulative returns often used in event studies are likely to be even more biased. Therefore, our results suggest that cumulative raw, or abnormal, returns used in these studies should be substituted by holding period returns.

In Section I we discuss the statistical problems associated with previous long-term contrarian strategies and introduce an alternative long-term performance measure. In Section II we present and analyze the empirical evidence, and demonstrate the likely biases in CRR-based measures of returns to long-term contrarian strategies. We conclude the paper with a brief summary in Section III.

I. Long-Term Contrarian Strategies

DeBondt and Thaler (1985) implement a long-term strategy to evaluate the "abnormal" returns earned by stocks that are losers and winners, where losers and winners are defined by their abnormal performance over the past three to five years. This basic strategy, with minor modifications, has subsequently been implemented by a number of researchers (see, for example, Ball and Kothari (1989), Chan (1988), Chopra, Lakonishok, and Ritter (1992), DeBondt and Thaler (1987), and Zarowin (1990)).

Since DeBondt and Thaler (1985) and most subsequent studies on overreaction use monthly returns of NYSE stocks, we present a brief summary of their contrarian strategy in this context.[3] The specific strategy uses three basic steps. First, over an initial "portfolio formation period" of typically

[2] The drift will be upward or downward depending on whether the portfolio of securities under consideration has an upward bias greater or less than the bias in the returns to the market portfolio (see Section I.*A* for details).

[3] This paper is concerned solely with biases in the returns to long-term contrarian strategies implemented by previous researchers. We do not address the important issue of the riskiness of such strategies (see Chan (1988), Ball and Kothari (1989), Chopra, Lakonishok, and Ritter (1992), and Kothari and Shanken (1992)).

three years, the cumulative abnormal (monthly) return for every stock is calculated as

$$\mathrm{CAR}_i = \sum_{t=1}^{36} \mathrm{AR}_{it} \tag{1a}$$

where

$$\mathrm{AR}_{it} = \mathrm{R}_{it} - \mathrm{R}_{mt}. \tag{1b}$$

The abnormal return is therefore a market-adjusted return, $\mathrm{R}_{it} - \mathrm{R}_{mt}$, where R_{it} is the return on security i in period t and R_{mt} is the equally weighted market return in period t. The above step is repeated for all subsequent nonoverlapping three-year periods. At the end of each of the portfolio formation periods, all firms are ranked on the basis of their CARs, and typically the extreme high and low performers (for example, the top 35 and bottom 35) are assigned to a winner and a loser portfolio, respectively.

The second step involves an evaluation of the future performance of past losers and winners. The three-year periods following each of the portfolio formation periods are the "performance evaluation periods." During each of these evaluation (or test) periods the CAR (as defined in equation (1a)) of each security in the loser or winner portfolio is calculated for up to 36 months. The cross-sectional average CARs of all losers or winners are then averaged (for each month of the evaluation period) across all the nonoverlapping evaluation periods.

Given the definition of abnormal returns in equation (1b), the CAR of a particular security i is

$$\mathrm{CAR}_i(k) = \sum_{t=1}^{k} (\mathrm{R}_{it} - \mathrm{R}_{mt}) = \sum_{t=1}^{k} \mathrm{R}_{it} - \sum_{t=1}^{k} \mathrm{R}_{mt} \tag{2}$$

where $k = 1, \ldots, 36$. The cross-sectional average cumulative abnormal return for a particular evaluation period can then be written as

$$\begin{aligned} \mathrm{ACAR}_p(k) &= \frac{1}{n} \sum_{i=1}^{n} \left[\sum_{t=1}^{k} \mathrm{R}_{ipt} - \sum_{t=1}^{k} \mathrm{R}_{mt} \right] \\ &= \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{k} \mathrm{R}_{ipt} - \sum_{t=1}^{k} \mathrm{R}_{mt} = \mathrm{ACRR}_p(k) - \mathrm{CRR}_M(k). \end{aligned} \tag{3}$$

The first expression on the right-hand side of equation (3), $\mathrm{ACRR}_p(k)$, is the average cumulative raw return of all securities ($i = 1, \ldots, n$) in portfolio p up to month k, where p = loser (L) or winner (W) portfolio. The second expression is $\mathrm{CRR}_M(k)$, which is simply the cumulative return of the equally weighted market up to month k. The difference between the average cumulative abnormal returns can then be defined as

$$\mathrm{DACAR}(k) = \mathrm{ACAR}_L(k) - \mathrm{ACAR}_W(k) = \mathrm{ACRR}_L(k) - \mathrm{ACRR}_W(k). \tag{4}$$

Note that since the equally weighted market return is subtracted from the returns of both the losers and winners for each k, it does not affect $\mathrm{DACAR}(k)$.

The final step in the overreaction strategy is to test whether: (a) $\text{ACAR}_L(k) > 0$, (b) $\text{ACAR}_W(k) < 0$, and (c) most importantly, $\text{DACAR}(k) > 0$. In carrying out the tests, the $\text{ACAR}_p(k)$'s are averaged across all nonoverlapping test periods for each k. The basic idea of these tests is that, if there is long-term overreaction, the loser (or winner) portfolio should outperform or underperform the market over the long term. Most importantly, however, an arbitrage (zero-investment) portfolio of the losers and winners should earn positive returns. And given the notion that the market exhibits long-term overreaction, the returns to the arbitrage portfolio should "typically" increase up to some (theoretically) unspecified k.

A. A Methodological Issue

DeBondt and Thaler (1985) find that, consistent with the notion of long-term overreaction, $\text{ACAR}_L(k) > 0$ and $\text{ACAR}_W(k) < 0$. Most importantly, $\text{DACAR}(k)$ is consistently greater than zero for all k up to 36 months, and its magnitude increases systematically with k to an average cumulative return of 24.6% over a 36-month interval. They also note that the overreaction effect is much larger for losers than winners. After 36 months the loser portfolio earns an average excess return of 19.6%, while the winner portfolio underperforms the market by only 5%.

Subsequent studies analyze the differential riskiness of, and seasonality in, the returns to losers and winners. However, there is a great deal of disagreement regarding the importance of these factors in explaining the different performance of losers and winners. For example, Chan (1988) and Ball and Kothari (1989) argue that differences in risk can explain the abnormal performance of the two types of firms. However, DeBondt and Thaler (1987) and Chopra, Lakonishok, and Ritter (1992) provide evidence which suggests that differential risk cannot explain the asymmetric overreaction of losers versus winners. Similarly, Zarowin (1990) argues that the evidence on overreaction is another manifestation of the size effect, and it occurs almost exclusively in January. But DeBondt and Thaler (1987) argue that the winner-loser effect is not primarily a size phenomenon, and Chopra, Lakonishok, and Ritter (1992) show that a substantial proportion of the overreaction (over 50%) occurs in non-January months.

However, there is an important drawback of the typical long-term contrarian strategy that has not been addressed, and which could potentially "explain" the dramatic evidence of overreaction. Note that, due to Jensen's inequality, any noise (in a statistical sense) in stock prices will lead to an upward bias in simple single-period returns (see Blume and Stambaugh (1983)). Consequently, by cumulating single-period returns, previous contrarian strategies also cumulate the upward bias in each period's return.[4] Follow-

[4] Ball and Kothari (1989) and Chopra, Lakonishok, and Ritter (1992) use annual holding period returns in most of their analyses. However, the non-January versus January return analysis to losers or winners in Chopra, Lakonishok, and Ritter (1992) is based on monthly returns. Also, we present evidence that the "true" returns to losers or winners are exclusively in January and that these returns are not due to overreaction.

ing Blume and Stambaugh (1983), suppose that the bid-ask spread is the only source of measurement errors in observed prices.[5] Instead of observing the true price of a security, P_{it}, we typically observe a closing price, P^o_{it}, which could be at the bid or the ask. Specifically, the bid-ask effect can be modeled as

$$P^o_{it} = [1 + \theta_{it}]P_{it}, \tag{5a}$$

or

$$P^o_{it} = P_{it} + \epsilon_{it} \tag{5b}$$

where $E(\theta_{it}) = 0$, θ_{it} is independently distributed over time, and is independent of P_{ik} $\forall k$. One can think of $\epsilon_{it} = \theta_{it}P_{it}$ as an (additive) measurement error in observed prices.

The "true" return for security i in period (month) t is defined, assuming no dividends for the period, as

$$R_{it} = \frac{P_{it}}{P_{it-1}} - 1. \tag{6a}$$

However, the measured or observed return is

$$R^o_{it} = \frac{P^o_{it}}{P^o_{it-1}} - 1 = \frac{[1 + \theta_{it}]P_{it}}{[1 + \theta_{it-1}]P_{it-1}} - 1. \tag{6b}$$

Using equation (6a), equation (6b) can be rewritten as

$$R^o_{it} = \frac{1 + \theta_{it}}{1 + \theta_{it-1}}[1 + R_{it}] - 1. \tag{6c}$$

Consequently,

$$E(R^o_{it}) = E\left\{\frac{1 + \theta_{it}}{1 + \theta_{it-1}}\right\}[1 + E(R_{it})] - 1. \tag{6d}$$

By Jensen's inequality, $E\{[1 + \theta_{it}]/[1 + \theta_{it-1}]\} > 1$, and the upward bias in single-period measured returns can be approximated (using a Taylor series) by

$$E(R^o_{it}) \cong E(R_{it}) + \sigma^2(\theta_{it-1}) \tag{7}$$

where $\sigma^2(\cdot)$ denotes the variance.

Furthermore, under the assumption that all stock trades occur with equal probability at the bid or the ask prices, P_B and P_A, and that the "true" price is given by $P = (P_A + P_B)/2$, the approximate bias in single-period returns is

[5]Although price discreteness (Gottlieb and Kalay (1985)) and nonsynchronous trading (Scholes and Williams (1977)) are also important, bid-ask errors are likely to have more serious consequences for the properties of asset returns (see, for example, Blume and Stambaugh (1983), Dravid (1988), and Kaul and Nimalendran (1990)).

given by

$$E(R^o_{it}) \cong E(R_{it}) + s_i^2/4 \tag{8}$$

where s_i is the proportional spread of security i, that is, $s_i = \dfrac{(P_A - P_B)}{(P_A + P_B)/2}$.[6]

Using equations (3) and (8), we can write the observed average cumulative abnormal returns as

$$\text{ACAR}^o_L(k) \cong \text{ACAR}_L(k) + k[B_L - B_M], \tag{9a}$$

$$\text{ACAR}^o_W(k) \cong \text{ACAR}_W(k) + k[B_W - B_M], \tag{9b}$$

and

$$\text{DACAR}^o(k) \cong \text{DACAR}(k) + k[B_L - B_W] \tag{9c}$$

where $\text{ACAR}^o_p(k)$ and $\text{ACAR}_p(k)$ denote the observed and "true" average cumulative abnormal returns, and B_L, B_W, and B_M are the upward biases due to the bid-ask effect in a *single period*'s return of the portfolio of losers, winners, and all NYSE stocks, respectively. Specifically, $B_L = \dfrac{1}{n_1}\sum_{i=1}^{n_1}\dfrac{s_i^2}{4}$, and $B_W = \dfrac{1}{n_2}\sum_{i=1}^{n_2}\dfrac{s_i^2}{4}$ and $B_M = \dfrac{1}{N}\sum_{j=1}^{N}\dfrac{s_j^2}{4}$, where $n_1 = n_2 = n$ are the number of securities in each of the loser and winner portfolios and N is the total number of securities on the NYSE. Note that the biases in the single-period returns of the three portfolios reflect the average of bid-ask biases in the returns of the securities comprising the respective portfolios.

Equations (9a) to (9c) have important implications for long-term overreaction strategies used in previous studies. First, if $B_L > B_M > B_W$, then there will be an *upward* bias in the average cumulative abnormal return of loser firms and a *downward* bias in the ACAR of winner firms. This will occur if, relative to the average NYSE firm, losers are low-priced and winners are high-priced. Furthermore, given the above inequalities, the biases could also explain the asymmetric performance of loser and winner portfolios. Second, it is clear that the returns to the arbitrage portfolio of losers and winners, $\text{DACAR}^o(k)$, will be upwardly biased if losers are low-priced relative to winners. Third, the biases in $\text{ACAR}^o_L(k)$, $\text{ACAR}^o_W(k)$, and $\text{DACAR}^o(k)$ increase linearly with the measurement interval. As k increases, DACAR^o will exhibit exactly the same upward "drift" hypothesized by proponents of long-term overreaction. Finally, the absolute magnitude of the upward bias in *single-period* returns is invariant with respect to the length of the period over

[6] In equation (8) the bias is not time-dependent because the proportional spread of a particular security is assumed to remain constant over time. Also, the upward bias in single-period returns will be greater than $s_i^2/4$ if losers (or winners) are more likely to have traded at the bid (or ask) price (see Bhardwaj and Brooks (1992) and Keim (1989)). This possibility, however, is unlikely to have a significant effect on the bias in *cumulative* returns used in contrarian strategies (see below) because only the first month's returns are likely to have a bias greater than $s_i^2/4$.

which the return is measured. Therefore, the upward bias will be a larger proportion of daily as compared to monthly returns.

For illustration purposes, in Figure 1 we show the relation between the price of a stock and the bid-ask bias in its 36-month (or 36-period) cumulative raw return. We assume that the dollar spread is 25 cents. The two most important aspects of Figure 1 are: (a) the nonlinearity in the relation between the bias and the price, and (b) the potentially substantial bias in the cumulative returns of low-priced stocks (a \$1 stock has a bias of 56.25%). The nonlinearity is of particular significance for our study. For example, the bias in a \$3 stock is only 6.25%, but the average bias in two stocks with prices of \$1 and \$5, such that their average price is \$3, is a much higher 29.25%! This occurs because $\frac{\bar{s}^2}{4} < \frac{1}{n}\left(\sum_{i=1}^{n} \frac{s_i^2}{4}\right)$, where $\bar{s}$ is the average proportional spread of n securities. This implies that due to the nonlinearity of the bias, squaring the average spread of n securities will give a downward biased estimate of the average bias in the cumulative returns of the portfolio of the securities. This will occur particularly for a portfolio that contains a few low-priced securities.

Figure 1 is used simply to demonstrate the potential biases in ACAR_p^o and DACAR^o. The actual bias in the estimates reported in previous studies can only be evaluated by measuring performance using an appropriate procedure for computing returns.

B. An Alternative Performance Evaluation Measure

Consider an alternative performance evaluation measure based on holding period returns of both losers and winners for all holding periods k, where $k = 1, \ldots, K$. For each k, we first calculate the holding period return of each security in the loser and winner portfolios and then average the security returns to obtain portfolio holding period returns. To calculate the abnormal performance of the loser and winner portfolios with respect to the market portfolio, we subtract the holding period returns of the equally weighted market portfolio from the holding period returns of the portfolios of losers and winners for each k. In calculating the holding period market return, we do not compound the monthly NYSE equally weighted market return, but average the holding period returns of each NYSE security for each k. Roll (1983a) provides a detailed theoretical treatment of this buy and hold method, and compares it to other methods for calculating multiperiod returns.

The average holding period (or buy and hold) abnormal performance measures, therefore, are

$$\begin{aligned}\text{AHPAR}_p(k) &= \frac{1}{n}\sum_{i=1}^{n} \text{HPR}_{pi}(k) - \frac{1}{N}\sum_{j=1}^{N} \text{HPR}_{mj}(k) \\ &\quad i = 1, \ldots, n \quad j = 1, \ldots, N \quad p = \text{L}, \text{W} \\ &= \text{AHPR}_p(k) - \text{AHPR}_M(k) \qquad (10)\end{aligned}$$

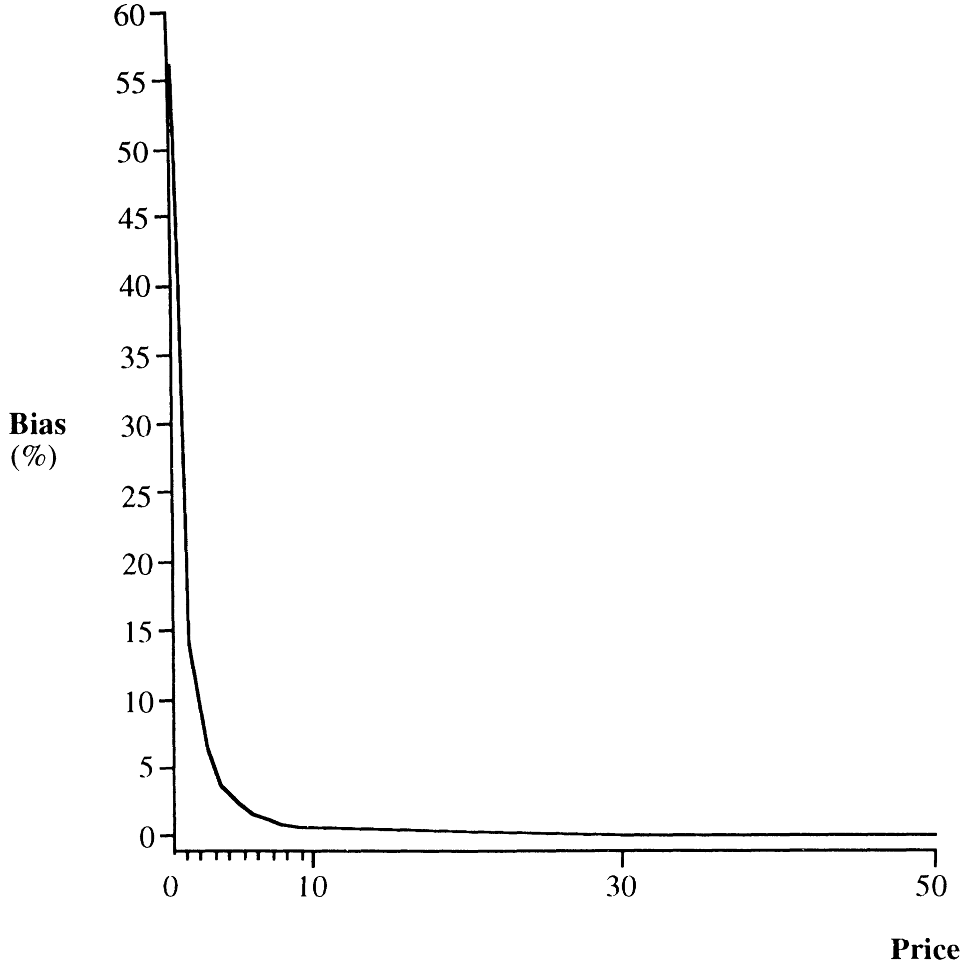

Figure 1. Approximate upward bias in 36-month cumulative raw returns of different priced stocks caused by the bid-ask spread. The single-period bias is calculated as $s_t^2/4$, where s_t is the proportional spread given by $\frac{(P_A - P_B)}{(P_A + P_B)/2}$ (see equation (5)). P_A and P_B are the ask and bid prices, and the dollar spread, $P_A - P_B$, is assumed to be 25 cents. The reported numbers are the bias in 36-month cumulative raw returns, that is, $36 * s_t^2/4$.

where $AHPR_p(k)$ is the average holding k-period return for the portfolio of losers or winners, and $AHPR_M(k)$ is the average holding k-period return for the portfolio of all NYSE stocks. The holding period return of a security over a k-period interval is calculated by compounding single period returns, that is, $HPR(k) = (1 + R_1)(1 + R_2)\ldots(1 + R_k) - 1$.

The buy and hold performance evaluation procedure implicit in equation (7) is consistent with the long-term contrarian strategy of holding securities for long periods in the expectation of long-term price reversals. In addition, it is appealing from a statistical viewpoint because it minimizes the bias in measured returns. Recall that the upward bias in single-period returns due to measurement errors is invariant with respect to the length of the measurement interval. Consequently, the bias in computed, or observed, $\text{AHPAR}_p^o(k)$ will be (approximately) constant over *all* holding periods k. Using equations (8) and (10), we can write the relation between true and observed average holding period abnormal returns as

$$\text{AHPAR}_p^o(k) \cong \text{AHPAR}_p(k) + (\text{B}_p - \text{B}_M) \quad p = \text{L}, \text{W} \tag{11}$$

and the performance measure of the zero-investment strategy is given by

$$\text{DAHPAR}^o(k) \cong \text{DAHPAR}(k) + (\text{B}_L - \text{B}_W) \tag{12}$$

where $\text{DAHPAR}^o(k)$ and $\text{DAHPAR}(k)$ are the observed and true difference between the average holding period returns of the loser and winner portfolios.[7]

Therefore, our performance evaluation measures for the loser, winner, and the arbitrage portfolios will have a bias which is constant with respect to k and, consequently, there will be no "spurious" upward drift in $\text{DAHPAR}^o(k)$. The constant bias is likely to be small even for low-priced stocks. For example, from Figure 1 the total bias even for a $1 stock (with a 25 cent spread) in a 36-month holding period return will be only 1.56%, that is, 56.25%/36.

II. Empirical Evidence

We use the DeBondt and Thaler (1985) procedure to identify 35 losers and 35 winners on the NYSE by ranking on the basis of three-year cumulative monthly market-adjusted returns over successive three-year nonoverlapping periods from 1929 to 1988 (see Section I).[8] We therefore have 20 three-year portfolio formation periods: the first starting in January 1926, and the last in January 1983. We also evaluate the performance of the losers and winners

[7]Another way to understand the reason for the lower, and constant, bias in multiperiod buy and hold returns is by considering the linking process involved in computing individual security returns. Roll (1983a) shows that the expected multiperiod buy and hold return is affected by the autocovariances of individual securities. If, for example, bid-ask errors are the only source of measurement errors in prices, then it can be shown that the negative autocovariances induced by these errors tend to offset any additional upward bias in multiperiod buy and hold returns (apart from that present in a single period's measured return). Of course, other source(s) of negative or positive autocovariances will also affect expected multiperiod buy and hold returns.

[8]There are two minor differences between our procedure and the one used by DeBondt and Thaler (1985). First, we do not require securities to have an unbroken series of 85 monthly returns, but require all securities to have all returns over each three-year portfolio formation period. Second, since we test whether long-term overreaction is a low-price (as suggested by our bias-hypothesis) or a small-firm (see Zarowin (1990)) phenomenon, we also require that the 35 losers and 35 winners have both transaction prices and share structures available at the end of every portfolio formation period. This requirement does not alter any of the conclusions of this paper.

over three-year nonoverlapping periods, the first period commencing in January 1929 and the last in January 1986. A security with a missing return in a particular evaluation period is dropped from the sample for the rest of the three-year interval. We calculate the long-term performance of the 35-firm portfolios of winners and losers using the DeBondt and Thaler (1985) cumulative-return method and the alternative holding-period method. Since the objective is to demonstrate the biases in previous performance evaluation measures, we retain the same set of losers and winners across both evaluation methods. That is, we use the DeBondt and Thaler procedure to identify winners and losers, but evaluate their subsequent performance using the two different methods.

A. The Cumulative Performance Measures

Table I reports the average cumulative abnormal returns, ACARs (see equation (3)), for one-, two-, and three-year periods. The ACARs are calculated using the equally weighted market return. Panel A contains ACARs for losers and winners, and the differences between them, that is, DACARs, using all months, while panel B contains the corresponding estimates for non-January months. The panel B numbers are calculated by restricting the January abnormal returns of all securities to be equal to zero. For comparison purposes, we also report the all-month DACARs reported by DeBondt and Thaler (1985); they do not report estimates for non-January months.

Since the relative prices of losers and winners are of critical importance to the magnitude of the upward bias in the ACARs and DACARs, we also report the average, minimum, and maximum of the *average* prices of the 35-firm loser and winner portfolios over the 20 evaluation periods. For completeness, we report similar statistics for the market values of the firms. The reported numbers characterize the distribution of the *average* prices and market values of the 35-firm loser and winner portfolios at the end of the 20 portfolio formation periods (or, alternatively, at the beginning of the 20 portfolio evaluation periods). Note that there is a substantial difference in the average prices of losers and winners: \$11.480 versus \$38.576. More importantly, the minimum *average* price of the loser firms is only \$1.623, while that of the winners is much larger (\$9.323). Since the reported minimums and maximums are also based on averages, there are a number of loser firms (over 10% of the sample) with prices less than \$1 which could lead to a substantial upward bias in both their own 36-month cumulative abnormal returns and in the 36-month portfolio ACARs (see Section *C.3*).

The ACARs of loser and winner portfolios in Table I, Panel A, reflect the same pattern of upward and downward trends documented by DeBondt and Thaler (1985) and other researchers. The loser portfolio consistently "earns" positive abnormal returns which increase systematically with the cumulation period, and eventually result in an abnormal return of 26.3% over 36 months. On the other hand, the winner portfolio consistently underperforms the market, and again at an increasing rate with the lengthening of the cumula-

Table I

Average Cumulative Abnormal Performance (With and Without January) of Loser and Winner Portfolios of NYSE Firms Using Equally Weighted Market-Adjusted Returns, 1929–1988

Average, minimum, and maximum price and market value, and average cumulative abnormal returns, ACARs, of the loser and winner portfolios of NYSE stocks at the end of one, two, and three years. Losers and winners are determined by their performance relative to the market over the past three years. We also report the differences between the ACARs of the loser and winner portfolios, i.e., DACARs. The reported ACARs are based on monthly market-adjusted returns, $R_{it} - R_{mt}$, of 35 firms within each of the loser and winner portfolios for 20 three-year nonoverlapping intervals between 1929 and 1988. We calculate abnormal returns using the equally weighted market return. The numbers in parentheses are the standard errors of the ACARs and DACARs based on the distribution of their 20 subperiod averages. For comparison purposes, we report DACARs for the 35-firm loser and winner portfolios reported in DeBondt and Thaler (1985). Panel A reports estimates based on all months, and Panel B contains estimates based on non-January months (i.e., the returns in January are assumed to be equal to zero for all securities in the loser, winner, and the market portfolios). The price and market value estimates are based on average values at the beginning of each three-year performance evaluation period.

	Price (in Dollars)			Market Value (in Millions Dollars)			Average Cumulative Abnormal Returns (ACARs)		
Portfolio	Average	Minimum	Maximum	Average	Minimum	Maximum	1-year	2-year	3-year
Panel A: All Months									
Loser	11.480	1.623	27.679	70.068	0.927	433.799	0.069	0.198	0.263
							(0.041)	(0.088)	(0.103)
Winner	38.576	9.323	123.805	160.957	3.785	496.625	0.001	−0.056	−0.112
							(0.034)	(0.068)	(0.069)
DACARs							0.068	0.254	0.375
							(0.059)	(0.139)	(0.153)
DeBondt-Thaler (1985) DACARs							0.054	0.181	0.246
							(0.070)	(0.106)	(0.112)
Panel B: Non-January Months									
Loser							−0.014	0.047	0.064
							(0.035)	(0.072)	(0.072)
Winner							0.023	−0.015	−0.058
							(0.029)	(0.060)	(0.052)
DACARs							−0.037	0.062	0.122
							(0.049)	(0.116)	(0.108)

tion period. After 36 months, the estimated $ACAR_w$ is -11.2%. The arbitrage portfolio therefore earns a 37.5% return over a 36-month interval (see DACARs), and a predominant part of this return is contributed by the loser firms. These estimates are consistently larger than the DeBondt and Thaler (1985) DACARs, reported in the last row of Panel A, mainly due to the fact that our sample period is different.

Panel B reports the ACARs for non-January months. It is important to evaluate the extent to which the "January effect" (see Keim (1983)) is responsible for the overreaction results. If biases in ACARs induced by measurement errors are at least partly responsible for the returns to long-term contrarian strategies implemented in previous studies, then the DACARs should be positive and should reflect an upward trend even for returns cumulated over non-January months. The results in Panel B show that the ACARs of losers and winners continue to exhibit the respective upward and downward drifts even in non-January months. The DACAR increases consistently to over 12% for a 36-month cumulation period. Therefore, about one-third of the returns to the arbitrage portfolio is in non-January months. This finding confirms the evidence in Chopra, Lakonishok, and Ritter (1992) that cumulative-return-based overreaction results are *not* peculiar to January.

For brevity, we do not report the ACARs and DACARs for each of the 20 portfolio evaluation periods. However, there are two interesting characteristics of the distribution of the ACARs and the corresponding prices of the underlying securities which suggests that biases in single-period returns are at least partly responsible for the differential cumulative returns to losers and winners. First, for the all-month sample the ACARs of the losers have three "extreme" positive values (out of a total of 20): 175.01%, 154.65%, and 82.70%. The corresponding *average* beginning of the period prices of the loser firms are low: \$1.75, \$2.49, and \$1.82, respectively. Similarly, the two "extreme" ACARs for the winners are -94.67% and -60.82%, and the corresponding average prices are high, \$41.05 and \$60.60.[9] This evidence is consistent with the bias hypothesis. Since the bid-ask effect is directly related to the square of the proportional spread, we would expect to see a greater upward bias in $ACAR_L$'s when prices of the losers are low, and a larger negative $ACAR_W$ when the prices of winners are high. Direct evidence in favor of the bias hypothesis is provided in the next two sections.

B. The Buy-and-Hold Performance Measures

In Table II we present the abnormal performance measures based on holding period returns. We report the average holding period abnormal returns for the losers and winners, $AHPAR_L$ and $AHPAR_W$, and for the arbitrage portfolio, DAHPAR, for one-, two-, and three-year holding periods. Panel A contains the all-month estimates, and Panel B reports the non-

[9]An extreme value is defined as a value greater in absolute magnitude than the absolute value of the grand average of all 20 ACARs plus one standard deviation.

Table II

Average Holding Period Abnormal Returns (With and Without January) of Loser and Winner Portfolios of NYSE Firms, 1929–1988

One-, two-, and three-year average holding period abnormal returns, AHPARs, of the winner and loser portfolios of NYSE stocks, and the differences between the AHPARs of loser and winner portfolios, i.e., DAHPARs. Losers and winners are determined by their performance relative to the market over the past three years, and are identical to those in Table I. The reported numbers are averages of AHPARs for 20 three-year nonoverlapping intervals between 1929 and 1988. The estimates for each subperiod are obtained by first compounding the monthly returns for each security for each holding period, and then averaging across loser or winner firms. We subtract the market holding period return from the average holding period returns of losers and winners. The numbers in parentheses are standard errors based on the distribution of the subperiod estimates. Panel A reports estimates based on all months, and Panel B contains estimates based on non-January months (i.e., the returns in January are assumed to be equal to zero for all securities in the loser, winner, and the market portfolios).

	Average Holding Period Abnormal Returns (AHPARs)		
Portfolio	1-year	2-year	3-year
Panel A: All Months			
Loser	0.026	0.139	0.204
	(0.036)	(0.080)	(0.126)
Winner	0.015	0.021	−0.067
	(0.036)	(0.074)	(0.064)
DAHPARs	0.011	0.118	0.271
	(0.053)	(0.129)	(0.157)
Panel B: Non-January Months			
Loser	−0.041	−0.020	−0.039
	(0.031)	(0.058)	(0.054)
Winner	0.027	0.038	−0.022
	(0.031)	(0.055)	(0.047)
DAHPARs	−0.068	−0.058	−0.017
	(0.046)	(0.093)	(0.081)

January estimates. For the non-January estimates we again restrict the January returns of all securities in the loser, winner, and the market portfolios to be equal to zero.

The estimates in Panel A have two interesting features. First, the holding period abnormal returns to the loser (or winner) portfolios are less (or greater) than the corresponding cumulative abnormal returns for *all k*. Second, due to the decrease in $AHPAR_L$'s and the increase in $AHPAR_W$'s, the arbitrage portfolio consistently earns lower holding period returns than the corresponding cumulative returns. For example, the 36-month DAHPAR drops to 27.1% from an estimate of 37.5% for the DACAR in Table I, Panel A. These findings are consistent with the bias hypothesis.

From our perspective, the non-January evidence in Panel B is more important because it allows us to more accurately gauge the impact of the bid-ask bias on cumulative-return-based measures of contrarian profitability. Recall that a large proportion of the returns to contrarian strategies occur in January. Consequently, the "January effect" in both cumulative and holding period returns makes it difficult to gauge the impact of the bid-ask effect. On the other hand, the non-January returns are not contaminated by the "January effect" and, therefore, should allow us to evaluate the extent of the bias in the cumulative returns to contrarian strategies. More importantly, the non-January evidence allows us to more accurately gauge the relative contributions of losers and winners to the upward bias in the cumulative returns of the arbitrage portfolio.

The non-January month estimates in Table II, Panel B, suggest that the actual returns to the arbitrage portfolio are entirely due to the January effect. The arbitrage portfolio actually earns consistently *negative* returns in non-January months, which is contrary to the predictions of the overreaction hypothesis. The importance of the upward bias in non-January months is reflected in the drop in the 36-month return to the arbitrage portfolio from 12.2% in Table I to −1.7%! Also note that most of the 13.9% drop in the 36-month non-January returns to the arbitrage portfolio is due to the decrease in the holding period returns of loser firms, as opposed to the increase in the returns of winner firms. Comparing Panels B of Tables I and II, the 36-month abnormal returns to losers drop by 10.3% (from 6.4% to −3.9%), while the abnormal returns to winners increase by only 3.6% (from −5.8% to −2.2%). This is consistent with the bias hypothesis because loser firms have considerably lower prices than winner firms.

C. Further Tests of the Bias Hypothesis

In this section, we present some additional tests to demonstrate the biases in the cumulative returns used in previous studies to measure the performance of long-term contrarian strategies.

C.1. Some Regression-Based Analysis

In Table III we present regressions of the 36-month cumulative raw returns (CRRs) of *individual* securities in each of the loser and winner portfolios for all evaluation periods on their initial (beginning of the evaluation period) prices and market values. Given the nonlinear relation between returns and price and/or market value, we use logarithm of prices and market values in the regressions. If the CRRs contain an upward bias due to measurement errors in single-period returns, they should be negatively related to prices but not necessarily to market values. Conversely, however, if CRRs contain no biases, and overreaction is primarily a small-firm phenomenon (see Chopra, Lakonishok, and Ritter (1992) and Zarowin (1990)), then the CRRs (of especially the losers) would be negatively related to market values but not necessarily to prices.

The results in Table III support the bias hypothesis, as opposed to the small-firm overreaction hypothesis. Panel A presents estimates based on the

Table III

Regressions of Long-Term Cumulative Performance Measures on Price and Market Value of Loser and Winner NYSE Firms, 1929–1988

Regressions of three-year cumulative raw returns, CRRs, of NYSE loser and winner firms, with and without January, measured over the 20 three-year nonoverlapping periods on their prices and market values at the beginning of each three-year interval over the 1929–1988 period. Losers and winners are determined by the performance relative to the market over the past three years, and are identical to those in Tables I and II. Panel A contains regressions in which CRRs are calculated for all months, while Panel B reports estimates of the same regressions with all January months assumed to earn zero returns for all securities. The estimated regression is of the form

$$CRR_{36,it} = \alpha_t + \beta_t LP_{it} + \gamma_t LM_{it} + \epsilon_{36,it}$$

where $CRR_{36,it}$ is the three-year cumulative raw return of losers or winners, LP_{it} and LM_{it} are the logarithms of beginning-of-the-period price and market value of the losers or winners, and $\epsilon_{36,it}$ is a random disturbance term. t-Statistics are reported in parentheses below the parameter estimates and are based on standard errors computed using White's (1980) heteroskedasticity-consistent method.

Portfolio		$\hat{\alpha}_t$	$\hat{\beta}_i$	$\hat{\gamma}_i$	$\bar{R}^2$
		Panel A: All Months			
Loser	(1)	1.631	−0.470		0.252
		(14.827)	(−10.628)		
	(2)	2.608		−0.204	0.119
		(8.781)		(−7.034)	
	(3)	1.875	−0.438	−0.033	0.252
		(7.324)	(−8.760)	(−1.138)	
Winner	(1)	1.229	−0.277		0.120
		(10.687)	(−7.914)		
	(2)	1.512		−0.111	0.056
		(7.376)		(−5.842)	
	(3)	1.132	−0.296	0.015	0.119
		(5.896)	(−5.920)	(0.600)	
		Panel B: Non-January Months			
Loser	(1)	0.863	−0.260		0.093
		(9.084)	(−6.842)		
	(2)	1.359		−0.108	0.039
		(5.329)		(−4.320)	
	(3)	0.940	−0.250	−0.010	0.091
		(3.715)	(−5.208)	(−0.333)	
Winner	(1)	0.854	−0.198		0.063
		(7.237)	(−5.500)		
	(2)	0.974		−0.071	0.023
		(4.870)		(−3.944)	
	(3)	0.674	−0.234	0.028	0.063
		(3.421)	(−4.333)	(1.037)	

returns of all months, and shows that the CRRs of both losers and winners are significantly negatively related to their prices (regression (1)) and market values (regression (2)). However, the R^2's of the price regressions are over twice the corresponding estimates for the market value regressions. More importantly, in the regression of CRRs on both price and market value, the latter has no marginal explanatory power and its coefficient is rendered insignificant. Hence, conditional on past prices there is *no* relation between cumulative returns to losers or winners and their market values.

Perhaps the most interesting feature of the results in Table III is the evidence in Panel B that is based on non-January months. The significant negative CRR price relation is witnessed even for non-January returns of both losers and winners. (Market value is significantly related to CRRs, but again not when the latter are conditioned on price as well.) Therefore, the negative relation between price and CRRs in Panel A is not merely a reflection of the fact that the January effect is a low-price phenomenon. There does appear to be an upward bias in CRRs of non-January months that is negatively related to prices.

Table IV contains estimates of the regressions of 36-month holding period returns on their initial prices and market values. These estimates can be compared directly with the evidence for cumulative raw returns in Table III. Panel A of Table IV contains the all-month estimates, while Panel B reports the non-January estimates. The results in Panels A and B provide strong support for the bias hypothesis. Note that the all-month coefficient estimates of the price variable, $\hat{\beta}_i$'s, are significantly negative, but are consistently smaller in magnitude than the corresponding $\hat{\beta}_i$'s in Panel A of Table III. Also the R^2's of the regressions drop dramatically. More importantly, however, in the non-January sample, the $\hat{\beta}_i$'s drop in magnitude and become indistinguishable from zero for both losers and winners. This evidence stands in contrast to the non-January estimates in Table III, Panel B, which show that all non-January $\hat{\beta}_i$'s remain significantly negative.

In summary, therefore, there appears to be a strong upward bias in the cumulative returns used in previous studies that is systematically related to prices and not to market values. The buy and hold methodology eliminates virtually all of this bias, and the all-month sample exhibits a negative relation between holding period returns and past prices only because the January effect is also a low-price, rather than a small-firm, phenomenon.

C.2. Price-Based Investment Strategies

To further evaluate the validity of the bias hypothesis, we implement a simple strategy which evaluates the long-term performance of 35 low- and high-priced stocks *without* considering their past performance. More specifically, at the beginning of each of the 20 portfolio evaluation periods we sort all NYSE firms based on their prices and assign the top 35 and bottom 35 firms into high- and low-price portfolios. We then evaluate their subsequent

Table IV

Regressions of Long-Term Holding Period Returns on Price and Market Value of Loser and Winner NYSE Firms, 1929–1988

Regressions of three-year holding period returns, HPRs, of NYSE loser and winner firms, with and without January, measured over 20 three-year nonoverlapping periods on their prices and market values at the beginning of each three-year interval over the 1929–1988 period. Losers and winners are determined by the performance relative to the market over the past three years, and are identical to those in Tables I and II. Panel A contains regressions in which HPRs are calculated for all months, and Panel B reports estimates of the same regressions with all January months assumed to earn zero returns for all securities. The estimated regression is of the form

$$HPR_{36,it} = \alpha_i + \beta_i LP_{it} + \gamma_i LM_{it} + \epsilon_{36,it}$$

where $HPR_{36,it}$ is the three-year holding period return of losers or winners, LP_{it} and LM_{it} are the logarithms of beginning-of-the-period price and market value of the losers or winners, and $\epsilon_{36,it}$ is a random disturbance term. t-Statistics are reported in parentheses below the parameter estimates and are based on standard errors computed using White's (1980) heteroskedasticity-consistent method.

Portfolio		$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\gamma}_i$	$\bar{R}^2$
		Panel A: All Months			
Loser	(1)	1.350	−0.346		0.087
		(8.232)	(−5.406)		
	(2)	1.981		−0.141	0.035
		(5.093)		(−3.711)	
	(3)	1.416	−0.337	−0.009	0.085
		(4.092)	(−4.956)	(−0.237)	
Winner	(1)	1.150	−0.234		0.042
		(4.957)	(−3.493)		
	(2)	1.276		−0.083	0.015
		(3.975)		(−2.862)	
	(3)	0.917	−0.280	0.036	0.042
		(3.568)	(−3.333)	(1.241)	
		Panel B: Non-January Months			
Loser	(1)	0.346	−0.048		0.002
		(3.977)	(−1.455)		
	(2)	0.296		−0.005	−0.002
		(1.396)		(−0.238)	
	(3)	0.178	−0.070	0.023	0.002
		(0.805)	(−1.667)	(0.852)	
Winner	(1)	0.633	−0.116		0.012
		(2.986)	(−1.902)		
	(2)	0.609		−0.033	0.002
		(2.207)		(−1.320)	
	(3)	0.401	−0.162	0.036	0.012
		(1.814)	(−2.051)	(1.333)	

performance for exactly the same three-year periods as the loser or winner portfolios.

This experiment is important for two reasons. First, if the bias hypothesis is valid then low-priced firms should have substantially higher cumulative returns than loser firms, and high-priced firms should have lower cumulative returns than winner firms. Consequently, the cumulative returns to the arbitrage portfolio should be higher. Second, if the January returns to the price-based portfolios are greater than the January returns to the performance-based portfolios, it would imply that the "January effect" is a low-price phenomenon that has little to do with past performance, or overreaction.

The cumulative abnormal performance measures and the buy and hold returns for the price-based portfolios are reported in Tables V and VI, respectively. For completeness we also report the average prices and market values of the sampled securities. The evidence in Table V has several interesting features. For both the all-month (Panel A) and non-January (Panel B) samples, the portfolio of low-priced firms has substantially higher ACARs than the loser firms, and high-priced firms typically underperform the market by more than the winner firms. Consequently, the arbitrage portfolio of low- and high-price firms has cumulative abnormal returns in Panel A that are *two to four times* larger than the loser-winner arbitrage portfolio. The one-, two-, and three-year DACARs are 23.7%, 54.9%, and 67% compared to only 5.4%, 18.1%, and 24.6% in Table I, Panel A. Similarly, the estimates in Panel B show that the arbitrage portfolio of low-high-price firms has larger non-January cumulative returns as well. For example, the 36-month DACAR is 19.7% compared to 12.2% in Panel B of Table I with a predominant part of the returns being contributed by low-priced firms.

Estimates of the holding period returns in Table VI show that virtually all of the long-term returns to the price-based portfolios are in January. More importantly, the non-January evidence provides strong support for the bias hypothesis. Note that the non-January 36-month bias in the cumulative return measure for the price-based arbitrage portfolio is 23.7%; the 36-month DACAR in Panel B of Table V is 19.7%, and the corresponding DAHPAR drops to −4.0% in Table VI, Panel B. On the other hand, the bias in the 36-month cumulative return to the loser-winner arbitrage portfolio is 13.9%. And again the low-priced firms are responsible for a predominant proportion of the 23.7% bias in the cumulative return of the price-based arbitrage portfolio. Specifically, the bias in the low-priced firms is 18.7% compared to only 5.0% for the high-priced firms.

Apart from supporting the bias hypothesis, the results in Tables V and VI also suggest that the January returns to the loser-winner portfolio have little to do with market overreaction. Note that all the holding period returns to both the performance-based and price-based arbitrage portfolios are realized in January (see Panels A of Tables II and VI). And since the returns to the price-based portfolio are *at least twice* as large as those earned by the performance-based portfolio, the "January effect" appears to have little rela-

Table V

Average Cumulative Abnormal Performance (With and Without January) of Low- and High-Priced Portfolios of NYSE Firms Using Equally Weighted Market-Adjusted Returns, 1929–1988

Average, minimum, and maximum price and market value, and average cumulative abnormal returns, ACARs, of the low- and high-priced portfolios of NYSE stocks at the end of one, two, and three years. The low- and high-priced stocks are the sets of 35 firms with the lowest and highest prices at the beginning of each performance evaluation period. We also report the differences between the ACARs of the low- and high-priced portfolios, i.e., DACARs. The reported ACARs are based on monthly market-adjusted returns, $R_{it} - R_{mt}$, of 35 firms within each of the low- and high-priced portfolios for 20 three-year nonoverlapping intervals between 1929 and 1988. We calculate abnormal returns using the equally weighted market return. The numbers in parentheses are the standard errors of the ACARs and DACARs based on the distribution of their 20 subperiod averages. Panel A reports estimates based on all months, and Panel B contains estimates based on non-January months (i.e., the returns in January are assumed to be equal to zero for all securities in the low-priced, high-priced, and the market portfolios). The price and market value estimates are based on average values at the beginning of each three-year performance evaluation period.

	Price (in Dollars)			Market Value (in Millions Dollars)			Average Cumulative Abnormal Returns (ACARs)		
Portfolio	Average	Minimum	Maximum	Average	Minimum	Maximum	1-year	2-year	3-year
Panel A: All Months									
Low-price	2.574	0.454	5.839	9.131	0.365	26.834	0.187	0.424	0.493
							(0.070)	(0.135)	(0.138)
High-price	132.913	74.925	251.500	2167.714	184.594	5849.428	−0.050	−0.125	−0.177
							(0.031)	(0.071)	(0.083)
DACARs							0.237	0.549	0.670
							(0.095)	(0.195)	(0.209)
Panel B: Non-January Months									
Low-price							0.036	0.169	0.161
							(0.056)	(0.119)	(0.097)
High-price							−0.006	−0.029	−0.036
							(0.028)	(0.066)	(0.066)
DACARs							0.042	0.198	0.197
							(0.078)	(0.173)	(0.153)

Table VI

Average Holding Period Abnormal Returns (With and Without January) of Low- and High-Priced Portfolios of NYSE Firms, 1929–1988

One-, two-, and three-year average holding period abnormal returns, AHPARs, of the low- and high-priced portfolios of NYSE stocks, and the differences between the AHPARs of the two portfolios, i.e., DAHPARs. The low- and high-priced stocks are the sets of 35 firms with the lowest and highest prices at the beginning of each performance evaluation period. The reported numbers are averages of AHPARs for 20 three-year nonoverlapping intervals between 1929 and 1988. The estimates for each subperiod are obtained by first compounding the monthly returns for each security for each holding period, and then averaging across low- and high-priced firms. We subtract the market holding period return from the average holding period returns of low- and high-priced firms. The numbers in parentheses are standard errors based on the distribution of the subperiod estimates. Panel A reports estimates based on all months, and Panel B contains estimates based on non-January months (i.e., the returns in January are assumed to be equal to zero for all securities in the low-priced, high-priced, and the market portfolios).

Portfolio	Average Holding Period Abnormal Returns (AHPARs) 1-year	2-year	3-year
	Panel A: All Months		
Low-price	0.103	0.376	0.426
	(0.074)	(0.157)	(0.172)
High-price	−0.034	−0.097	−0.140
	(0.030)	(0.068)	(0.082)
DAHPARs	0.137	0.473	0.566
	(0.096)	(0.210)	(0.238)
	Panel B: Non-January Months		
Low-price	−0.039	0.079	−0.026
	(0.054)	(0.124)	(0.080)
High-price	0.004	−0.008	0.014
	(0.027)	(0.057)	(0.051)
DAHPARs	−0.043	0.087	−0.040
	(0.072)	(0.156)	(0.118)

tion to past performance. We also estimate regressions of 36-month holding period returns of individual losers or winners and low- or high-price firms on their past performance and beginning of the period (logarithm) prices. For brevity we do not report the detailed results. However, we find that conditional on beginning of the period prices, there is *no* relation between long-term returns and past performance. Therefore, the January returns to losers and/or winners are not due to market overreaction.[10] The January effect appears to be a low-price phenomenon which may be due to tax loss selling

[10]Also, the correlation between (logarithm) price and past performance drops from 0.63 for the loser firm sample to only 0.20 for the low-price firms. This again suggests that the January returns to losers documented in previous studies are related to price and have little to do with past performance.

(see, for example, Roll (1983b), Reinganum (1983), Ritter (1988)), market microstructure biases (see Keim (1989)), and/or higher risk in January.

C.3. Some Simulated Estimates of the Bid-Ask Bias

The extent of bias induced by the bid-ask effect in the cumulative return measures appears surprisingly large considering the fact that the average price of the loser firms is over \$10 (see Table I). For example, from Figure 1 the bias in the 36-month cumulative raw return (CRR) of a \$10 stock with a 25 cent spread is only 0.56%. Based on the non-January cumulative return versus holding period analysis, however, the extent of bias is 13.9% (with losers contributing most (10.3%) of the bias). Similarly, the low-priced firms in Table V have an average price of approximately \$2.50, and from Figure 1 the implied 36-month bias is 9.0%. However, the low-priced firms again contribute a bias of 18.7% (out of a total bias of 23.7%) to the price-based arbitrage portfolio. Therefore, the actual biases in the cumulative returns appear to be considerably larger than those suggested by the bias for the *average* firms in the loser or winner or in the price-based samples.

The reason for the large bias in the "real" data lies in the highly nonlinear relation between the price of a stock and the bias it induces. Recall that the actual bias for a 36-month interval is given by $36 \times \left(\frac{1}{n} \sum_{i=1}^{n} \frac{s_i^2}{4} \right)$, which can be substantially larger than $36 \times \left(\frac{\bar{s}^2}{4} \right)$, where $\bar{s}$ is the average proportional spread of the n securities in the portfolio. In other words, the square of the average spread is much smaller than the average of squared spreads of the securities in a portfolio.

To evaluate the effects of the nonlinear bias in the real data, we simulate *lower bounds* for the biases in the 33-month cumulative returns of losers and low-priced firms. The losers and low-priced firms are chosen because they contribute most of the bias to the arbitrage portfolio, and 33 months are used because non-January returns provide a clearer picture of the bias in the real data. To simulate the bias we calculate proportional spreads for all 700 losers and 700 low-priced firms in our total sample (that is, 35 firms in each of the 20 evaluation periods). The proportional spreads are calculated by dividing the dollar spread for each security by its price at the beginning of an evaluation period. The dollar spreads for various price *categories* are chosen based on the distribution of *actual* dollar spreads of all NASDAQ-NMS firms listed on the CRSP tapes on a randomly chosen day.[11]

[11] Dollar spreads are not available for NYSE firms used in our analysis. Consequently, we use the spreads of NASDAQ-NMS securities. Since the actual and simulated biases are more sensitive to low-priced securities, we ensure that dollar spreads of such securities are not unreasonably high. For example, securities with prices less than \$1 (which comprise 10% of our loser firms) have dollar spreads ranging between 1/32 and 1/8.

The 33-month (non-January) simulated bias is then calculated as 33 $\times \left(\frac{1}{n} \sum_{i=1}^{n} \frac{s_i^2}{4} \right)$, where n = 700 for each of the loser firm and low-price firm samples. For the loser firm and the low-price firm samples, the simulated biases are 7.7% and 16.8%, respectively. These estimates are much larger than the bias calculated from the *average* spread of all 700 firms in each sample. More importantly, these estimates are close to the biases of 10.3% and 18.7% in the actual non-January loser firm and low-price firm samples. Both the simulated biases are slightly lower than the actual biases because we use a fixed dollar spread for securities in different price *categories* rather than for each security. Also, the biases are calculated under the assumption that the proportional spread of a security is constant over a three-year period. Of course, most of the bias will result from cross-sectional differences in spreads which have largely been accounted for in our simulation. Time-series variation in the proportional spreads are also likely to contribute a small fraction of the actual biases because the square of the average spread of a security over a three-year period will be less than the average of the squared monthly spreads during the same period. However, this bias (which is not accounted for in the simulation) will be small because cross-sectional differences in spreads are likely to be much larger than the time-series variation in the proportional spread of a particular security.[12]

III. Conclusion

In this paper, we show that the returns to long-term contrarian strategies implemented in previous studies are upwardly biased because they are calculated by cumulating single-period (monthly) returns over long periods. The cumulation process not only cumulates true returns, but also the upward bias in single-period returns induced by measurement errors (for example, due to the bid-ask effect). Using a buy and hold performance measure, we show that all non-January returns to long-term contrarian strategies are eliminated. The actual return to an arbitrage portfolio of losers and winners is solely due to January returns, and we show that this "January effect" has no relation to past performance of the securities. Hence, there is no evidence of market overreaction: the abnormal performance of previous long-term contrarian strategies is due to a combination of a biased performance measure and a "January effect" that is unrelated to prior performance.

Although our analysis has been limited to the returns to long-term contrarian strategies, our results have potentially important implications for a number of empirical studies. Specifically, CARs are used in virtually every event study to gauge the impact of information events on stock prices. Our results show that even for monthly returns the CARs of securities could exhibit spurious upward or downward drifts. With the increasing use of

[12] The analysis in this section resulted from comments made by the referee.

high-frequency (intraday) data this problem is likely to become more severe. Caution must therefore be exercised in interpreting the CAR of a security as a measure of returns that can be "earned" over the cumulation period.

REFERENCES

Asquith, Paul, 1983, Merger bids, uncertainty and stock-holder returns, *Journal of Financial Economics* 11, 51–83.

Ball, Ray and S. P. Kothari, 1989, Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns, *Journal of Financial Economics* 25, 51–74.

Bernard, Victor L. and Jacob K. Thomas, 1989, Post-earnings-announcement drift: Delayed price response or risk premium?, *Journal of Accounting Research* 27, 1–36.

Bhardwaj, Ravinder K. and Leroy D. Brooks, 1992, The January anomaly: Effects of low share price, transaction costs, and bid-ask bias, *Journal of Finance* 47, 553–575.

Blume, Marshall and Robert F. Stambaugh, 1983, Biases in computed returns: An application to the size effect, *Journal of Financial Economics* 12, 387–404.

Brown, Stephen J. and Jerald B. Warner, 1980, Measuring security price performance, *Journal of Financial Economics* 8, 205–258.

Chan, K. C., 1988, On the contrarian investment strategy, *Journal of Business* 61, 147–163.

Chopra, Navin, Josef Lakonishok, and Jay R. Ritter, 1992, Measuring abnormal returns: Do stocks overreact? *Journal of Financial Economics* 31, 235–268.

Conrad, Jennifer and Gautam Kaul, 1988, Time-variation in expected returns, *Journal of Business* 61, 409–425.

———, 1989, Mean reversion in short-horizon expected returns, *Review of Financial Studies* 2, 225–240.

DeBondt, Werner F. M. and Richard M. Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793–805.

———, 1987, Further evidence on investor overreaction and stock market seasonability, *Journal of Finance* 42, 557–581.

Dravid, Ajay, 1988, Effect of bid-ask spreads and price discreteness on distributions of stock returns, Working paper, Stanford University.

Fama, Eugene F., 1991, Efficient capital markets: II, *Journal of Finance* 46, 1575–1617.

——— and Kenneth R. French, 1988, Permanent and temporary components of stock market prices, *Journal of Political Economy* 96, 246–273.

Gottlieb, Gary and Avner Kalay, 1985, Implications of the discreteness of observed stock prices, *Journal of Finance* 40, 135–153.

Kaul, Gautam and M. Nimalendran, 1990, Price reversals: Bid-ask errors or market overreaction?, *Journal of Financial Economics* 28, 67–83.

Keim, Donald B., 1983, Size-related anomalies and stock return seasonality: Further empirical evidence, *Journal of Financial Economics* 12, 13–22.

———, 1989, Trading patterns, bid-ask spreads, and estimated security returns: The case of common stocks at calender turning points, *Journal of Financial Economics* 25, 75–97.

Kothari, S. P. and Jay Shanken, 1992, Stock return variation and expected dividends: A time series and cross-sectional analysis, *Journal of Financial Economics* 31, 177–210.

Lehmann, Bruce N., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1–28.

Lo, Andrew W. and A. Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41–66.

Poterba, James M. and Lawrence H. Summers, 1988, Mean reversion in stock prices: Evidence and implications, *Journal of Financial Economics* 22, 27–59.

Reinganum, Marc R., 1983, The anomalous stock market behavior of small firms in January: Empirical tests for tax-loss selling effects, *Journal of Financial Economics* 12, 89–104.

Ritter, Jay R., 1988, The buying and selling behavior of individual investors at the turn of the year, *Journal of Finance* 43, 701–717.

Roll, Richard, 1983a, On computing mean returns and the small firm premium, *Journal of Financial Economics* 12, 371–386.

———, 1983b, Vas ist das? The turn of the year effect and the return premia of small firms, *Journal of Portfolio Management* 9, 18–28.

Scholes, Myron and Joseph Williams, 1977, Estimating betas with nonsynchronous data, *Journal of Financial Economics* 5, 309–327.

Shefrin, Hersh M. and Meir Statman, 1985, The disposition to ride winners too long and sell losers too soon: Theory and evidence, *Journal of Finance* 41, 774–790.

White, Hal, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

Zarowin, Paul, 1990, Size, seasonality, and stock market overreaction, *Journal of Financial and Quantitative Analysis* 25, 113–125.