UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

The Joint-Effects of Conservatism and Market Inefficiency on the Returns-Earnings Relation

A dissertation

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

By

Roxanne Gooch

Norman, Oklahoma

2005

UMI Number: 3203286

UMI Microform 3203286

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

The Joint-Effects of Conservatism and Market Inefficiency
on the Returns-Earnings Relation

A Dissertation

Approved for the
Michael F. Price College of Business

BY

Dr. Robert C. Lipe

Dr. G. Lee Willinger

Dr. Frances L. Ayres

Dr. Wayne B. Thomas

Dr. Scott C. Linn

Acknowledgments

My deepest gratitude and appreciation go to my major professor, Robert Lipe for his guidance and advice during my years in the doctoral program, especially throughout the development and completion of this dissertation. I express my sincere thanks to the other members of my committee: Lee Willinger, Fran Ayres, Wayne Thomas and Scott Linn. I would also like to thank the workshop participants at The University of Oklahoma and my doctoral colleagues for their helpful comments. Finally, I am deeply grateful for the patient support and encouragement of my daughter Alison, my son Jonathan and his wife Carrie, my granddaughter Alexandra and other family members without whose support this project could not have been completed.

Table of Contents

1. Introduction

Capital markets researchers frequently use regressions of contemporaneous stock returns on accounting variables to evaluate whether or how effectively accounting information captures events that have affected the firm over the fiscal period. Accounting's ability to capture these events is measured against stock price (or change in stock price) which presumably reflects the same events completely and unbiasedly, i.e. is "efficient." Many of the studies using returns/earnings regressions draw inferences by testing the magnitude of regression coefficients (earnings response coefficients or ERCs) or differences in ERCs. However, the coefficients can be affected by factors not central to the research question such as market inefficiency. Prior research gives little recognition to the bias in coefficients resulting when the dependent variable - stock price - reacts to information with a delay.

The assumption of market efficiency is contradicted by a growing body of evidence that the market is not completely efficient with respect to publicly available accounting information.[1] In particular, associations exist between accounting information and future returns suggesting some of the impact of information captured by current period accounting variables is reflected in future rather than contemporaneous stock returns. For example, Penman and Zhang (PZ) (2002) observe a positive association between their measure of accounting conservatism and future excess returns, suggesting the market is misled by conservative accounting.

[1]Kothari's 2001 comprehensive survey of capital market research includes a review of accounting studies investigating market inefficiency.

An interesting question not addressed by PZ then arises: Would results of studies on conservatism that focus on tests of ERCs and assume market efficiency actually reflect conservatism or market inefficiency? The answer gains importance in light of conservatism's pervasive influence on accounting and the amount of attention it has attracted from accounting researchers in recent years.

In this study, I adapt a procedure developed by Aboody, Hughes and Liu (AHL) (2002) to adjust coefficients for bias caused by inefficiencies induced by accounting conservatism. The AHL adjustment reduces bias in the estimation of coefficients in a value relevance regression when inefficient stock prices measure intrinsic value with error. Exploiting the notion that market inefficiencies are resolved over time, AHL adjust the dependent variable (price or returns) for future price changes that include delayed market reactions to current accounting information. As a result, the adjusted regressions in AHL produce coefficient estimates that capture contemporaneous and delayed market reactions to accounting variables. AHL do not attempt to identify specific aspects of accounting that contribute to the inefficiency-induced bias, but their procedure provides a means to test the influence of conservatism on market efficiency and allows an examination of the consequent implications for coefficients in value relevance regressions.

In tests of return predictability, I find strong associations between market inefficiency and three measures of conservatism: two capture conservatism's temporary effects on earnings quality and the third captures conservatism's effect on the asymmetric timeliness of earnings. Results from the AHL adjusted regressions are less significant but give some support to the notion that failure to control for the joint effects of conservatism and market efficiency may bias ERCs enough to affect economic inferences. I find significant differences between the magnitude of average ERC adjustments for market inefficiencies between samples exhibiting high and medium levels of conservatism as measured by Penman and Zhang's Q-score. I also find significant differences between the ERC adjustments for samples exhibiting high and low conservatism, but the sign of the difference is opposite of predictions based on the patterns of returns observed in PZ. They claim, but do not directly test, that the market inefficiency suggested by the results in their study is due to the market responding inappropriately to current earnings because of an inability to penetrate the implications of conservatism for earnings quality. While intuitive, their claim is not supported when subjected to closer scrutiny. My results suggest the abnormal future returns to a trading strategy based on the Q-score demonstrated by PZ are not associated with current earnings.

The rest of the paper is organized as follows. The next section provides background and contains a review of related literature. Section 3 outlines my research design and describes the empirical proxies and sample I use in the study. Results are presented in Section 4, and Section 5 provides concluding comments.

2. Background and Literature Review

2.1 Conservatism

Conservatism serves as a constraint on the measurement and recognition criteria in GAAP. Application of conservatism leads to a tendency toward early recognition of unfavorable events and the minimization of net assets and income, and this introduces a bias into financial reporting (FASB 1984). Conservative policies include the immediate expensing of costs whose future economic benefit cannot be measured with sufficient reliability (e.g. advertising and R&D), recording depreciation in excess of economic depreciation, LIFO accounting for inventories in periods of rising prices, and higher verifiability requirements for the recognition of good news relative to bad news.

Staubus (1985) views conservatism as a major component of approximately one-third of accounting standards. Sterling (1970) deems conservatism the most influential valuation principle in accounting. Numerous recent studies confirm accounting is conservative and has become more conservative over time (Givoly and Hayn 2000, Basu 1997, and others). Despite conservatism's prominence in accounting standards, a "generally accepted definition of conservatism" does not exist (Givoly and Hayn 2000, p. 291). Thus, researchers examine the effects of conservatism from different points of view. For instance, a set of studies views conservatism as conditional, or news dependent; the lower of cost or market method of valuing inventory and impairment rules for long-lived assets are examples of conditional conservatism. Basu (1997) was the first to empirically examine accounting conservatism from this perspective. He posits that the lag in accounting recognition is less pronounced for the recognition of "bad news" relative to "good news" due to conservatism's higher verifiability requirements for the recording of gains. To test his prediction, Basu (1997) uses negative (positive) returns to proxy for bad (good) news in a regression in which earnings is the dependent variable. He finds earnings are more sensitive to bad news, suggesting that the timeliness of earnings varies with the type of news.

Conservatism can also be unconditional, or news independent, meaning that the accounting rules yield unrecorded goodwill at the inception of an economic event (Beaver and Ryan 2004).[2] For instance, the immediate expensing of costs with future benefits remaining at year end (e.g. R&D and advertising) is conservatism in the unconditional sense. Feltham and Ohlson (1995) characterize this type of conservatism as "biased accounting," meaning that reported net assets are expected to be less than market value in the long run. A large literature, beginning with

---

[2] "Conditional" and "unconditional" conservatism are relatively recent terminology used to distinguish the two types of conservatism (see Beaver and Ryan 2004). Prior studies such as Basu (1997) and others refer to "conservatism in the income statement" and "conservatism in the balance sheet." Beaver and Ryan note that all conservatism affects the balance sheet and income statement consistently if clean surplus holds and if transactions in firm's equity are accounted for at fair value. The literatures on conditional and unconditional conservatism are generally separate; however Beaver and Ryan (2004) develop a model that captures the interactions between the two types. For example they state, "…unconditional conservatism yields unrecorded goodwill that preempts the application of conditional conservatism unless news is sufficiently bad to use up that goodwill (p. 46)."

Greenball (1969), models and empirically investigates the implications of unconditional conservatism on accounting numbers in the presence of growth (see also Penman and Zhang 2002).

2.2 Accounting and Market Inefficiency

Overall, evidence from a significant body of research suggests the market reacts quickly to accounting information. However, a growing body of evidence points to instances where the market reaction to accounting data is incomplete or biased. One notable example is the market's delayed reaction to earnings, i.e. the post-earnings announcement drift, first observed in Ball and Brown (1968). Numerous subsequent studies examine possible explanations for the drift including Bernard and Thomas (1989, 1990). Bernard and Thomas rule out risk-based and transaction-cost based explanations for the drift and conclude investors appear to underestimate the implications of current earnings for future earnings (i.e ignore the serial correlation in quarterly earnings). Ou and Penman (1989) use multiple financial ratios that incorporate information other than earnings from the current period financial statements to predict earnings changes; they provide evidence that the predicted earnings changes can also predict returns. Results in Sloan (1996) suggest that although cash flows are more persistent than accruals, the market acts as if they have the same persistence implying the market under-reacts to cash flows relative to accruals (the "accrual anomaly"). Thus, a trading strategy of going long in high cash flow firms and short in high accrual firms earns abnormal profits. Frankel and Lee (1998)'s residual income intrinsic value measure constructed from analyst's forecasts also demonstrates predictive ability for future abnormal returns.

AHL (2002) appear to be the first to develop a formal analysis of the impact of market inefficiency on coefficients in value relevance studies and derive an econometric solution to correct bias that results from delayed market reactions to accounting variables. AHL posit that when the market measures stock price with error, any correlation between the error and the accounting variables of interest biases the coefficient due to the omitted correlated variable problem. AHL mitigate the bias by adjusting current prices to include risk-adjusted future price changes so that the dependent variable encompasses future as well as current price effects of information about intrinsic value contained in the current accounting variables. AHL apply the adjustment to three types of value relevance regressions found in the literature:

- the value relevance of earnings and book values (both price level and return regressions)
- the value relevance of residual income value estimates and
- the value relevance of accruals and cash flows.

AHL find mean and median coefficient estimates in the adjusted return regressions are significantly larger in magnitude than those obtained from the non-adjusted regressions.[3] For example, the coefficients on earnings and book value increase by

[3]AHL do not predict overreaction or underreaction by the market and thus cannot predict the direction of the correction. An exception is the regression on residual income value estimates. To support their argument that the adjustment procedure reduces error in value relevance estimations, AHL regress stock returns on a residual income-based intrinsic value deflated by lagged price where the predicted coefficient is one. AHL's finding that the coefficient moves toward the predicted coefficient suggests the procedure is effective in removing bias in estimates induced by market inefficiencies. AHL obtain

90% and 82%, respectively, when stock price is adjusted for three-year-ahead returns. For regressions with residual income as the independent variable, the adjustment results in a 24% mean increase in regression coefficients toward the predicted value of one. After adjustment, the coefficient on cash flows is significantly higher than the coefficient on accruals consistent with their differential persistence. Overall, evidence in AHL suggests their procedure considerably reduces bias in coefficients in return regressions and also emphasizes the importance of considering market inefficiency effects when drawing inferences from coefficients in value relevance studies.

2.3 Conservatism and Market Inefficiency

Relatively few studies examine the relation between conservatism and market inefficiency. The predominant type of study in this area is related to one aspect of unconditional conservatism, accounting for R&D expenditures. While the results of many studies suggest the market capitalizes R&D expenditures although the accounting does not (e.g. Lev and Sougiannis 1996, Hirschey and Weygandt 1985, and Deng and Lev 1998), other studies suggest investors are misled by conservative accounting for R&D. Lev and Sougiannis (1996) and Chan et al. (2001) report a positive association between the level of R&D investment and subsequent excess returns. However, Chambers et al. (2002) suggest the association is more likely due

more modest adjustments (but significant) for the level regressions. They argue that since price contains information accumulated since the inception of the firm and market inefficiencies more likely relate to new information, the bias is mitigated.

to inadequately controlling for risk rather than a mispricing explanation. Similar to Lev et al. (2000) and Penman and Zhang (2002), Chambers et al. (2002) report a positive association between *changes* in R&D investment and subsequent excess returns that does not appear to be explained by failure to adjust for risk. Overall, the evidence suggests the market is inefficient with respect to information about R&D, and the inefficiency arises when R&D investment experiences a material change.

The finding that changes in (rather than levels of) R&D investment are associated with mispricings is not surprising given the fact that unconditionally conservative accounting methods affect earnings only when the level of investment in assets changes over time. As PZ explain, the potential for investors to be misled by conservative accounting methods arises if a temporary change in investment results in a temporary earnings change making current earnings a poor predictor of future earnings. Penman and Zhang illustrate the interaction between conservatism and growth using accounting for R&D as an example. In the absence of growth in investment, the immediate expensing of R&D expenditures has no affect on earnings relative to capitalizing the R&D asset and amortizing it over its economic life. However, the immediate expensing of R&D costs lowers earnings and creates "reserves" if investment in R&D increases.[4] In contrast, slowing R&D investment reduces reserves and increases earnings as "the benefits from past expenditures are recognized without matching additional new expenditures" (Penman and Zhang 2002,

[4]The term "reserve" has many connotations in accounting (an appropriation of retained earnings, for example) and accounting literature (e.g. "cookie jar" reserves). In this paper, reserve refers to the unrecorded asset resulting from the expensing of costs that have remaining future economic benefit as of the balance sheet date.

p. 241). Any accounting method that recognizes higher expenses (or lower revenues) in the early years of a transaction relative to an unbiased method lowers net assets and has a similar effect. Accelerated depreciation methods have no effect on earnings if depreciable assets are not growing, but decrease earnings through the creation of reserves if the investment in assets increases. If investments in depreciable assets decrease, earnings increase through the liquidation of the "hidden" reserves. Thus, conservative accounting methods combined with changes in investment may lead to mispricing if investors misjudge the sustainability of earnings when the change in investment is temporary.

PZ construct a measure that captures the combined effect of conservatism and growth in investment to test the market's ability to understand the implications of conservatism for earnings quality. PZ characterize earnings as high quality if current earnings are useful in predicting future earnings. The "Q-score" includes estimated reserves (ER), or unrecorded assets, that result from accounting for R&D, inventory, and advertising deflated by net operating assets (NOA):

$$Q_{it} = \frac{ER_{it}}{NOA_{it}} - \frac{ER_{it-1}}{NOA_{it-1}}$$

If a firm builds up its reserve at a rate faster (slower) than the growth in net operating assets, Q is positive (negative). Both "high" (positive) and "low" (negative) Q-scores are indicators of poor earnings quality.

Consistent with an association between market inefficiency and Q, PZ observe differential returns for portfolios formed on Q. For two years prior to and including year 0, the year of portfolio formation, returns to low-Q portfolios are higher that returns to high-Q portfolios. This is consistent with the releasing of reserves that temporarily increases earnings for low-Q firms and the building up of reserves that temporarily decreases earnings for high-Q firms. However, for years +1 to +5, the pattern reverses, suggesting the market misinterprets information in firms' earnings in years -2 to 0 but corrects this mistake in subsequent years. The ability to earn excess future returns on a trading strategy based on Q suggests an association between accounting conservatism and the market's inability to process conservatism's implications for earnings quality. In an efficient market, investors will "see through" the accounting when using financial information in assessing future prospects of the firm. However, PZ's results are consistent with investors failing to differentiate accounting effects from real economic news.

In the current study, I investigate whether the conservatism-induced inefficiencies documented by PZ affect conclusions from regressions of returns on earnings. While I replicate PZ's tests as a first step in my analyses, my study differs from PZ's in two respects. First, the procedure developed by AHL (2002) allows me to consider the effect of market inefficiency related to conservatism on ERCs. Second, I consider alternate measures of conservatism in my tests.

3. Research Design

3.1 Measuring Market Inefficiency

Accounting researchers frequently assess the association between current-period accounting variables and future returns to investigate market efficiency. Kothari (2001) and others refer to these as "return predictability" tests. One common test analyzes future abnormal returns on stock portfolios formed on a current-period variable of interest. Penman and Zhang (2002) apply this approach to investigate the market's ability to penetrate information in conservative earnings. For completeness, I replicate their test using their Q-score and also with alternate measures of conservatism. However, my primary analysis is based on the method developed by Aboody, Hughes and Liu (2002), described below. While future abnormal returns to a trading strategy can provide evidence of market inefficiencies, the results do not provide direct evidence regarding the effect of inefficiencies on the coefficients in contemporaneous value relevance regressions.

AHL (2002) analytically show how coefficients in value relevance regressions can be estimated free of bias induced by market inefficiencies by changing the dependent variable to include future returns deflated by expected returns attributable to compensation for risk. In essence, the adjustment extracts information about the current measurement error in returns from future returns and incorporates it into current returns. By adjusting returns for future risk adjusted price changes, the regressions in AHL produce coefficients that capture contemporaneous as well as delayed market reactions to current earnings and book value.

AHL assume value relevance is measured by coefficients in the following linear relationship:

$$E(V_{it} \mid X_{it}) = B_t' X_{it} \tag{1}$$

where V_{it} and the vector X_{it} denote intrinsic value and accounting variables for firm *i* at time *t*, respectively

AHL further assume stock price, P_{it}, at time *t* measures intrinsic value, V_{it}, with error, u_{it}, i.e. in an inefficient market $E(u_{it} \mid I_{it}) \neq 0$ where I_{it} is all information available for firm *i* at time *t*:

$$P_{it} = V_{it} + u_{it} \tag{2}$$

If time *t* inefficiencies are resolved by the end of the period *t*+ 1 then

$$E(u_{it+1} \mid I_{it}) = 0 \tag{3}$$

AHL also assume intrinsic value V_{it} fully reflects all information so that expected return on V_{it} is driven purely by risk:

$$\overline{R}^{\upsilon}_{it+1} = E\left(R^{\upsilon}_{it+1} \mid F_{t+1}, I_t\right) = E\left(R^{\upsilon}_{it+1} \mid F_{t+1}\right) \tag{4}$$

where R^{υ}_{it+1} is the return on intrinsic value and is equal to $\frac{V_{it+1} + D_{it+1} - V_{it}}{V_{it}}$, and F_{t+1} is the realization of time *t*+1 risk factors. The expectation of time t + 1 cum-dividend stock price conditional on X_{it}, P_{it}, V_{it}, and F_{t+1} can be written as:

$$E\left(P_{it+1} + D_{it+1} \mid F_{t+1}, X_{it}, P_{it}, V_{it}\right)$$

$$= E\left(V_{it+1} + D_{it+1} + u_{it+1} \mid F_{t+1}, X_{it}, P_{it}, V_{it}\right) \text{ substituting (2)}$$

$$= E\left(V_{it+1} + D_{it+1} \mid F_{t+1}, X_{it}, P_{it}, V_{it}\right) \text{ using (3)}$$

$$= \left(1 + \overline{R}^{\upsilon}_{it+1}\right) V_{it} \tag{5}$$

Thus, intrinsic value at time t, V_{it}, equals the expected cum-dividend stock price$_{t+1}$ discounted by the risk-adjusted return, or:

$$V_{it} = E\left(\frac{P_{it+1} + D_{it+1}}{1 + \overline{R}^{\upsilon}_{it+1}} \mid F_{t+1}, X_{it}, P_{it}, V_{it}\right) \tag{6}$$

Taking the expectation of (6) conditional only on X_{it} and applying the law of iterated expectations:

$$E\left(V_{it} \mid X_{it}\right) = E\left(\frac{P_{it+1} + D_{it+1}}{1 + \overline{R}^{\upsilon}_{it+1}} \mid X_{it}\right) = B_i' X_{it} \tag{7}$$

Since $E\left(P_{it+1} + D_{it+1}\right) = \left(1 + R_{it+1}\right) P_{it}$, the adjusted contemporaneous price that measures V_{it} free of bias becomes $\left(\frac{1 + R_{it+1}}{1 + \overline{R}^{\upsilon}_{it+1}}\right) P_{it}$. AHL show the one-period model extends to multi-periods in which the measurement errors are assumed to follow a time-series process. AHL also demonstrate the adjustment does not create unwanted bias if markets are indeed efficient.

My analysis starts by replicating the non-adjusted and adjusted cross-sectional annual regressions in AHL:

$$\mathrm{R}_{it} = \beta_0 + \beta_1 \mathrm{NI}_{it}/\mathrm{P}_{it-1} + \varepsilon_{it} \quad \text{and}$$

$$\hat{\mathrm{R}}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \mathrm{NI}_{it}/\mathrm{P}_{it-1} + \hat{\varepsilon}_{it}$$

where $\hat{R}_{it} = \left(\frac{1+R_{it+\tau}}{1+\overline{R}^{\upsilon}_{it+\tau}} \right) R_{it}$, and $R_{it+\tau}$ (τ = 1, 2, 3) is firm *i*'s stock return for 36 months beginning three months after the fiscal year end t, and $\overline{R}^{\upsilon}_{it+\tau}$ (τ =1, 2, 3) is the corresponding size-decile return to control for price changes related to risk.[5] Percentage changes in regression coefficients from unadjusted regressions to adjusted regressions are calculated as $\frac{\hat{\beta}_t - \beta_t}{\beta_t}$.[6] Each year's percentage change is then treated as a separate observation in statistical tests of significance.

To assess whether conservatism causes bias in coefficients, I repeat the above analysis separately on samples exhibiting high and low levels of the conservatism measures (described below) and substitute operating income for net income. Operating income excludes temporary components of income which are less affected by conservatism in both the conditional and unconditional sense as measured in my study. The statistical tests assess the statistical significance of differences in the two samples' mean and median annual cross-sectional percentage changes. If the percentage changes for high- and low-Q samples, then conservatism is associated with the market inefficiency bias in the slope coefficients. In the next section, I discuss measures that allow comparison of relative levels of conservatism.

[5]Like AHL, I set τ equal to 1, 2, or 3 years. AHL arbitrarily limit τ to 3 years and observe the largest difference when τ = 3 which suggests the market takes up to three years to reverse some inefficiencies. Finding significantly different coefficients for longer periods may imply AHL's adjustment inadequately controls for risk. Unlike AHL, I do not include book value as an independent variable because my primary focus is on the coefficient on earnings.

[6] AHL define the percentage change in regression coefficients as the unadjusted coefficient minus the adjusted coefficient divided by the unadjusted coefficient. Thus, an increase in coefficients in their paper leads to a negative percentage change. To simplify the interpretation of results, I subtract the unadjusted coefficient from the adjusted so that an increase in coefficients leads to a positive percentage change.

3.2 Identifying the Earnings Impact of Conservatism

In addition to PZ's Q-score, I adapt proxies developed in prior studies to measure the relative influence of conservative accounting on firms' accounting numbers. I am limited to measures constructed from accounting data; the use of market data implicitly assumes the market prices information in an unbiased manner, and such an assumption contradicts the premise of my paper. A description of each measure follows.

PZ (2002) develop two indices to analyze the quality of earnings. The first is a C-score. It equals the sum of estimated reserves derived from inventory, R&D, and advertising, and measures the effect of conservative accounting on the balance sheet. The following unrecorded assets are summed and scaled by net operating assets to form C_{it}:

- Unrecorded Inventory Asset equals the LIFO reserve reported in the footnotes.
- Unrecorded Research and Development Asset is the estimated R&D asset, net of amortization, that would appear on the balance sheet if outlays for R&D were not immediately expensed. Lev and Sougiannis' (1996) industry amortization rates and estimated useful lives can be used to estimate the unrecorded asset resulting from R&D expenditures as follows:

  $$\sum_{k=0}^{T} \delta_k \left(\text{R \& D}_{t-k}\right)$$

  where t is the current period, $R\&D_{t-k}$ is the R&D expenditure for the k^{th} lag, δ is the unamortized percentage of the k^{th} lag's R&D expenditure, and T is the useful life of R&D investments.[7]

[7] Requiring the past k lags of R&D to be nonmissing potentially results in a significant loss of observations. So I adopt a procedure from Mason (2004) where past R&D expenditures can be estimated from current period R&D expenditures and estimated industry growth using the following :

- Unrecorded Advertising Asset is advertising expenditures assumed to be capitalized and amortized over a two year useful life, using the sum-of-the-year's digits method of depreciation.[8]

To measure conservatism's income statement effects, PZ subtracts C_{it-1} from C_{it} to form a Q-score.[9] By measuring the change in estimated reserves relative to net operating assets, the Q-score has the ability to capture conservatism's temporary effects on current earnings due to the interaction of conservatism and changes in investment. For instance, if estimated reserves grow faster (slower) than net operating assets, then Q is positive (negative); either a high (i.e. positive) or a low (i.e. negative) Q-score implies poor earnings quality. If estimated reserves grow at the same rate as net operating assets, then conservatism has no effect on earnings and the resulting Q-score of zero indicates high earnings quality.

I use PZ's Q-score and the AHL regressions to investigate whether conservatism leads to bias in ERCs. If inefficiencies related to conservatism are significant enough to affect inferences from returns-earnings regressions, then I should detect significant differences in unadjusted and adjusted coefficients across

---

$$\text{Lag}_{i,n} = \left(\frac{\text{Expenditures } i,t}{(1 + \text{Growthj})^n} \right)$$

where $\text{lag}_{i,n}$ is the estimated expenditure for firm i for the n^{th} prior period. The subscript t denotes the current period and j represents firm i's two-digit SIC code.

[8] Evidence from Bublitz and Etteredge (1989) suggests advertising has a short useful life, typically from one to two years. Penman and Zhang assume a two-year life when estimating the advertising reserve.

[9] PZ equally weight two measures of "unexpected" C_{it} to form Q: the change in C_{it}, and C_{it} minus a median industry C_{it}. Because PZ arbitrarily select an equal weighting scheme and find no significant difference in results when using Q-scores calculated by alternating weights of 1 and 0 on each measure of the unexpected C, I calculate Q as the change in C.

samples formed on Q. Unfortunately, predicting the direction and magnitude of percentage changes in ERCs for the high, medium, and low Q samples is problematic since Q does not ordinally rank firms on conservatism's earnings effects. Recall that PZ observe higher year 0 returns for the lowest Q-deciles (relative to the higher deciles); they attribute this to the market interpreting the temporary boost in earnings caused by conservatism as permanent. PZ observe a decrease in returns in future years as the market resolves the inefficiencies. They find the opposite result for the highest-Q deciles: low returns in year 0 reverse in future years. Therefore, I expect low-Q firms to have lower adjusted ERCs compared to the unadjusted ERCs, and the high-Q firms should have higher adjusted average ERCs. However, AHL find positive increases for average coefficients when adjusting returns for one, two or three year forward returns. Since I do not know the level of Q for firms in their sample, the implications of their results for my predictions are unclear. Either high-Q firms dominate their sample or all firms experience increases in adjusted ERCs due to some factor unrelated to Q. For instance, a Q-score of zero implies conservatism has no effect on earnings because relative amounts of reserves to net operating assets are not changing. Accordingly, PZ find relatively flat returns for years -2 to +5 for medium-Q firms in their return predictability tests suggesting the market efficiently processes information in high quality earnings. Thus, I expect no percentage change for coefficients from regressions using the sample consisting of the middle decile-rankings of Q (where Q is approximately zero). However, AHL find significant increases in coefficients for their regressions that do not consider conservatism suggesting I may find a positive adjustment for medium-Q firms. In light of AHL's results and my inability to control for factors other than Q, I make no predictions regarding the sign of the average percentage change in ERCs for medium and low deciles. Rather, I expect a greater increase in coefficients for the high-Q sample relative to the medium-Q sample and a greater increase in the medium-Q sample relative to the low-Q sample.

3.3 Alternative Measures of Conservatism

Similar to PZ's Q-score (2002), Mason (2004) estimates unrecorded assets relative to net operating assets to form an accounting-based measure of bias in a firm's balance sheet, a conservatism "score." In addition to R&D, advertising, and inventory, the estimated unrecorded assets in Mason's score include these related to accounts receivable, depreciation, and pension assets. The unrecorded Accounts Receivable Asset is measured as the difference between the firm's actual allowance for uncollectible accounts and a firm-specific minimum allowance. The minimum allowance is calculated by multiplying the firm's gross receivables by an industry-based minimum percentage, the lowest ratio of the allowance to gross receivables within an industry. The minimum percentage is adjusted for days' sales in receivables to remove expected losses from liberal accounting policies from the unrecorded asset. Unrecorded Property, Plant, and Equipment is measured as a firm's accumulated depreciation in excess of the accumulated depreciation assuming an industry-based maximum asset life. While her score excludes some unrecorded assets for tractability (e.g. goodwill) as do PZ (2002), it is broader in scope than PZ's Q-score by including unrecorded assets resulting from conservatism arising from management's estimates.[10] Mason uses levels of her score to rank firms on the level of conservatism affecting the balance sheet for her study, but I calculate the change in her score to enable the measure to capture conservatism's interaction with change in investment (as in PZ). Mason also includes unrecognized gains and losses resulting from pension plan assets in her estimate of unrecorded assets. A preliminary review of the data reveals this data item is missing for a large number of firms; therefore I do not include it in my modified version of her score.

A measure of conservatism motivated by Givoly and Hayn (2000) is the sign and magnitude of accumulated accruals over time. Since accruals are the mechanism that enables the immediate recognition of unfavorable events and a gradual recognition of favorable events, Givoly and Hayn's measure captures conservatism in the conditional sense (as in Basu 1997), without relying on stock price movements as a signal for news. Without conservatism, both positive and negative accruals reverse over time and average net accumulated accruals should approach zero. For a sample of firms with conservatism, average accruals should be negative. Following the intuition in Givoly and Hayn (2000), Ahmed et al. (2002) construct an accrual-based measure to capture conservatism's income statement effects – net accruals excluding depreciation (net income before extraordinary items plus depreciation expense less operating cash flows) deflated by assets and averaged over five-year periods. After

[10] Penman and Zhang (2002) exclude unrecorded assets arising from discretion in management estimates (for example, conservatism in estimates of the allowance of uncollectible accounts receivable) to focus on earnings quality issues related to permanent accounting policy.

multiplying the measure by -1, the measure is increasing in negative accruals and conservatism. Because the accumulation of accruals is also affected by growth of sample firms, splitting the sample on the accrual-based measure in my analysis results in an examination of the combined effect of investment growth and conservatism on market efficiency.

Figure 1 provides an overview of the study's five analyses. The first set of analyses serve as a benchmark; for those I implement a trading strategy approach for each of the three conservatism measures similar to that implemented by Penman and Zhang (2002). However, my main focus (and innovation of the study) is the second set of analyses which uses the method in AHL to examine differences in estimated coefficients between firms exhibiting differing levels of conservatism.

3.4 Sample and Data

Because the different conservatism measures impose different restrictions on the data, I construct three separate samples. Firms must have sufficient data from the annual COMPUSTAT files to construct the conservatism measures. Monthly stock returns, end-of-year size-decile assignments, and decile returns are supplied by CRSP. The sample period begins in 1975 because accounting data necessary to construct the conservatism scores were missing for a significant number of firms prior to that year. However, for analyses using the Q-score and the modified version of Mason's score, the earliest year is 1976 because of the differencing required to determine both measures. The sample period ends in 2003, the last year of available data.[11] Sample sizes vary depending on the annual return period being analyzed for the returns predictability tests (i.e. years -2 to +5 relative to the year of portfolio formation year 0). For the AHL regressions, sample sizes vary with the return adjustment horizon (i.e. one-, two-, or three-years ahead). Items designated as having insignificant values by Compustat are set to zero. Following PZ, if one or two of the three score components (R&D, LIFO reserve, or advertising) are nonmissing, the industry median (2-digit SIC code) substitutes for the missing item(s). I apply the same procedure for Mason's score (five components) to conserve data points. Because some firms within SIC codes 6000-6999 (insurance, real estate and investment brokers) capitalize a portion of software development costs that are included in Compustat's variable for R&D expenditures (data item 46), these firm-years are eliminated. I also eliminate firm-years with negative net operating assets. For the accrual-based conservatism measure, firm-years without four lags of accruals are eliminated. To control for outliers in the AHL regressions, I delete observations ranking in the top and bottom 1% of the dependent and independent variables (annual return and operating income per share deflated by price). Finally, firms undergoing a fiscal year-end change are excluded in the year of the change. For the PZ and Mason measures, the same samples are used for the AHL regressions and return predictability tests (a separate sample for each measure because of differing data requirements) to allow for comparability –evidence of market inefficiency related to conservatism found in the return predictability tests then leads to an investigation of potential ERC bias due to conservatism. I limit analysis of the accrual-based measure

[11] AHL use a sample period from 1962 to 1995 and PZ use a sample period from 1975 to 1995.

to the return predictability test. The final samples include a maximum of 62,754 firm years for the tests using Q-score, 30,737 firm-years for the tests using Mason's score, and 27,967 firm-years for the accrual-based measure. Table 1 presents descriptive statistics for variables used in the AHL regressions with the Q-score.

3.5 Descriptive Statistics

Table 2 contains descriptive statistics of firm characteristics for the entire sample and by level of Q (high, medium and low). The characteristics for medium-Q firms compared to the other groups are graphically illustrated in Exhibit 1. Relative to high-decile firms, firms in the lowest decile are larger (as measured by the market value of equity), are more profitable (as measured by earnings per share, return on assets and return on equity) and have a higher book to market ratio relative to high-decile firms. Higher earnings (and thus higher return on assets and equity) for low-Q firms are consistent with the notion that conservatism and change in investment increase current earnings relative to the high-Q firms. Medium-Q firms have higher means (and medians) for most firm characteristics relative to high- or low-Q firms. Interestingly, medium-Q firms have approximately three (six) times the mean (median) net operating assets as high- or low-Q firms but about the same amount of estimated reserves. However, recall the Q-score depends on the changes in rather than the levels of estimated reserves and net operating assets. Medium-firms have significantly higher average (median) market values than high- or low-Q firms. The differences cannot be attributed to industry differences since the ranking of Q within

2-digit SIC code ensures that each industry has similar percentages of high, medium, and low Q firms.

An analysis of changes in Q-ranks reveals that relatively few firms change from high to low ranks or vice versa in a given year (2.8%). I find that 1,061 firm-years changed rank from high to low over the sample period compared to 628 that moved from low to high. Further investigation into whether the move from high to low ranks of Q is due to changes in the numerator (estimated reserves) or denominator (net operating assets) finds an average one-year 7% (.5% median) decrease in estimated reserves and an average 216% (51% median) one-year increase in net operating assets. The results suggest firms moving from high to low ranks of Q are growing as evidenced by the large average increase in net operating assets but are decreasing investment in R&D, inventory and advertising. Whether this behavior indicates earnings management via real investment decisions is left for future research.

4. Results

Table 3 reports the results from tests of the ability of Q-score to predict future returns. Following PZ, in each scoring year from 1976 up through 2002 and within 2-digit SIC code, I rank firms into ten equal-sized portfolios based on Q-scores. Portfolios are formed in year 0. For each Q-score portfolio, raw and size-adjusted buy-and-hold returns are calculated for each year, -2 to +5. The returns cumulation period in each year ends three-months after fiscal year-end. Size-adjusted returns are calculated by subtracting the raw, buy and hold return on a size-matched, value-weighted portfolio formed from CRSP size deciles. Ranking within industry attempts to control for operating risk, and the size adjustment controls for the Fama and French (1992) "size effect" premium for risk. Panel A reports the mean returns over the sample period and Panel B reports the mean size-adjusted returns.

The results indicate returns are positively related to Q for years +1 to +5 and negatively related for years -2 to 0. Returns on low Q portfolios in year 0 are considerably higher than returns on high Q portfolios suggesting that investors misinterpret the earnings effects of the conservatism-growth interaction. The relation reverses in the following year, presumably as investors correct their error. The raw (size-adjusted) return in year +1 to a investment strategy that goes long in high Q-stocks and short in low Q-stocks without considering transaction costs is 19.1% (9.7%), significant at p-value < .0001. Future returns (years +2 to +5) to the strategy are similar in magnitude to those obtained by PZ but are not significant.[12] The results are consistent with the notion that the market is inefficient with respect to earnings effects induced by a combination of conservatism and changes in investment.

Next I assess whether conservatism-related inefficiencies materially affect the estimation of coefficients in a returns/earnings regression using the method in AHL. I run year-by-year cross-sectional regressions of returns on operating income per share scaled by previous-year price for samples formed on high, low, and medium levels of PZ's Q-score. Groups are based on a ranking of firms on Q for each year and with-in

[12] PZ tests of significance are based on 5,000 replications of randomly assigning firms to Q portfolios. PZ determine significance levels by observing the frequency of returns equal to the positive (negative) return on the investment strategy. I use t-tests to test for significance.

each two-digit SIC code, similar to the ranking procedure in PZ, as a means to control for risk. The high (low) score group corresponds to the top (lowest) decile of firms in the ranking on Q. Recall both high and low Q-scores imply low earnings quality. Firms in the medium-Q group are the middle 20% where the Q score equals approximately zero (i.e. has high earnings quality). Following AHL, I then rerun the regressions for each sample after adjusting current returns for one-, two-, and three-year ahead risk-adjusted returns.

Panel A in Table 4 presents the results from the unadjusted and adjusted regressions. The first set of regressions includes all observations in the sample to allow for comparison with results in AHL. Although AHL's independent variable is earnings, results for regressions with operating income as the independent variable provide similar conclusions. Consistent with the notion that inefficiencies resolve over one to three years, estimated mean coefficients increase on average as the dependent variable includes future year's returns. The mean (median) coefficient increases from .744 (.786) for the unadjusted return to 1.095 (1.102) when adjusted for three-year ahead returns. Panel B reports mean and median percentage changes; when three-year ahead returns are used for adjustments, the mean (median) increase in coefficients is 20.5% (40.2%), and the median is significant at *p*-value = .0014. AHL report mean and median increases of 77% and 66%, both significant at *p*-value < .001, when the adjustment includes three-year ahead returns. The large difference in the magnitude of the adjustments between my study and theirs may be due to using operating income as the independent variable in this study or to sample selection since observations in my study are limited to those with Q-scores.

I then apply the AHL adjustment to regressions for samples split on levels of conservatism and obtain mixed results. In Panel B, mean percentage changes in coefficients for medium-Q firms decrease as the adjustment increases from one to three years, but the changes are not significantly different from zero. Median coefficients for medium-Q firms increase on average by 33.2 % over three years, but the increase is only marginally significant at a p-value of .0986 in year three. Taken together, the insignificant mean and median percentage changes are consistent with less inefficiency effects for firms with high quality earnings. For low-Q firms, median percentage increases for years one and two are significant (19.8% and 37% with p-values of .0043 and .0055, respectively). Three–year ahead adjustments are not significant. For high-Q firms, the two and three-year adjustments result in significant increases relative to the unadjusted amounts and higher adjustments relative to low- and medium-Q firms, as expected; the median percentage increase in ERC achieved when the adjustment includes two- and three-year ahead returns is 56.2% and 65.6%, respectively, both significant at $p < .0001$. Mean results for the high-Q group are somewhat consistent with PZ's story: unadjusted current returns are lower than those of the low-Q portfolio, but reverse in the future as the market resolves the inefficiencies.

Panel C of Table 4 presents statistical tests of differences in percentage changes across groups. A significant difference between high- and low-Q groups occurs for the one-year adjusted coefficients, but opposite of predictions: the percentage change for low-Q groups is a 0.418 (p-value of .0208) larger than the percentage change for high-Q groups. The direction of the difference reverses in future years, but is not significant. I find the results somewhat puzzling -- if the adjusted coefficients followed the pattern of return reversals observed in Table 3, I would expect high-Q percentage changes to be larger than low-Q percentage changes for the one-year adjustment. I find significantly higher two- and three-year high-Q median percentages changes compared to medium-Q median percentage changes, consistent with predictions. However, I find significantly lower one- and two-year median percentage changes for medium-Q firms relative to low-Q firms, opposite of predictions and inconsistent with the pattern of returns in PZ. Increasing the power of the tests might lead to finding more significant differences between groups. Currently, test statistics are calculated using 27 annual observations; unfortunately data availability restricts adding more years to the sample.[13]

Next I repeat the above tests with Mason's conservatism score. Table 5 contains results from a trading strategy using the change in Mason's score (for size-adjusted returns only). It shows excess returns for years +1 to +3 of 10.7% (p-value $< .0001$), 11.4% (p-value $< .001$) and 10.6% (p-value $< .10$), suggesting the market takes three years to work out the inefficiencies compared to the one-year observed in a trading strategy using Q. The adjusted coefficients in Table 6, panel A do not follow the same general pattern as observed in Table 4: the three-year adjusted

[13] I also rerun the tests with operating income changes as an additional independent variable and observe similar results (unreported).

coefficients for the high-score group are higher than the low-score group and the adjusted coefficients for the low score group are higher than the medium score group, which is inconsistent with predictions. However, panel B reveals fewer significant percentage changes than the previous regressions using Q-scores. One possible explanation for the weaker results is that the inclusion of conservatism arising from management's estimates in the score adds noise to the measure. Another potential explanation is the procedure of handling missing score components (replacing of a large number of the Mason score components with industry averages) and then ranking the score within SIC codes to control for risk may lead to little variation between high and low scores.[14]

My final analysis uses a return predictability test to investigate inefficiencies related to conditional conservatism, the conservatism that arises from the tendency to include bad news earlier in earnings than good news. I measure conditional conservatism using a five-year average of accruals scaled by total assets since the asymmetric timeliness in earnings should lead to an accumulation of negative accruals. Table 7 reports average returns for portfolios formed on the accrual measure. The trading strategy produces positive future returns only in year +1 (7%, significant at p-value < .05) and year +5 (significant at p-value < .10). Sloan (1996) performs a similar test using current year accruals in an examination of the market's ability to distinguish correctly the accrual and cash flow components of earnings.

[14] To investigate the possibility that the procedure to replace missing score components with industry averages biases my results for the Q-score sample, I select a sample (n=13,239) that requires components of the Q-score to be nonmissing. Results (untabulated) show two-year percentage changes for high-Q firms are significantly greater than zero. They are also higher relative to all percentage changes for low- and medium-Q firms, but the differences are not statistically significant.

Sloan finds excess returns of 10%, 5% and 3% for years +1 to +3 (insignificant for year +3) and interprets his results to indicate the market never realizes the error because accruals reverse before three years. The premise behind the measure in my paper is that with conditional conservatism, negative accruals will accumulate. So the five-year average of accruals should capture an information set different from Sloan's measure, but it is possible the results in my tests are reflecting the same anomaly detected by Sloan.

In tests of return predictability, we need to know if the predictive ability of conservatism to predict returns is subsumed by other variables which have been shown to predict future stock returns. PZ's findings of predictable future returns due to the interaction of conservatism and a change in investment may be due to a correlation between Q and other anomaly factors. Table 8 provides correlations of Q with factors previously identified in anomalies studies: size, book-to-market, leverage, earnings-to-price, the risk proxies suggested by Fama and French (1992), and accruals identified by Sloan (1996). Because PZ's research design matches on size and industry, some of the differences in risk proxies across Q-groups are mitigated as evidenced by the weak correlations between Q, BTM, and leverage. Q is slightly correlated with size as measured by the firm's market capitalization decile-rank provided by CRSP (Pearson correlation = -.02; p-value < .0001) and earnings-to-price (Pearson correlation = -.07; p-value < .0001). However, Q is most strongly correlated with accruals (Pearson correlation -.1911; $p < .0001$).[15] PZ suggest the

[15] Accruals are calculated as the difference between net income before depreciation and cash flows from operations scaled by total assets.

correlation between Q and accruals arises because of management's propensity to manage earnings. If accruals are used to manage earnings, then management may also use real investment decisions to manage earnings, which Q captures since Q is designed to reflect changes in investment in off-balance sheet assets relative to net operating assets. PZ consider the correlation of Q and accruals as a possible explanation for Q's ability to predict future returns. In sensitivity tests, they regress future returns on the above risk factors, accruals, and Q. They conclude Q's effect on future returns is incremental to that of the other risk proxies and accruals (the mean coefficient on Q is significantly positive). Mason (2004) suggests a correlation between her measure (similar to Q) and accruals is possible due to current-period operating accruals being included in net operating assets. However, I find insignificant correlations between accruals and net operating assets (untabulated). Although most of the differences in portfolio returns are observed in year +1, significant excess size-adjusted returns for years +2 and +5 in PZ's paper (year +3 in my paper) provides possible evidence that Q is correlated with a yet-to-be identified risk factor.

5. Conclusion

The objective of this study is to test for bias in coefficients from returns-earnings regressions resulting from market inefficiency due to conservatism. Penman and Zhang (2002) demonstrate that the market is inefficient with respect to the earnings effects produced by the interaction of conservatism and change in investment and develop Q to measure conservatism's earnings effects. Since PZ provide evidence that a trading strategy based on Q earns excess returns in future years, I investigate whether the inefficiencies have implications for ERCs. In addition, I examine two other measures of conservatism.

To assess potential coefficient bias, I use a method developed by Aboody et al. (2002) that adjusts current period returns for information related to current earnings found in future returns. Adjusting coefficients for the market's delayed reactions to accounting information should then produce coefficients free of bias. I expect higher adjustments for firms displaying high levels of conservatism relative to firms with medium and low conservatism, and higher adjustments for firms with medium levels of conservatism relative to those with low levels. Overall, the results are mixed. I observe differences between adjusted and unadjusted coefficients for companies with high conservatism that are significantly greater than the differences for companies with medium levels of conservatism as expected, when the adjustment uses two-and three-year ahead future returns. Although I expect a larger adjustment for the high conservatism group relative to the low conservatism group, the one-year adjustment yields significantly larger coefficient increase for low-Q firms compared to the high-Q firms; the other differences in adjustments are mostly insignificant. The low significance levels may be due to low power of the tests which use only 27 annual differences in coefficients. Unfortunately, adding years to the sample is not possible due to lack of data.

Because I do find some significant differences between ERC adjustments between companies exhibiting high and medium-levels of conservatism, the results lend some support to the notion that coefficient bias due to conservatism can affect inferences, particularly if the magnitude of the ERC is the focus of the investigation. Thus, conclusions from my study may be of interest to future researchers who wish to avoid drawing spurious inferences from their statistical tests. However, the inefficiency documented in PZ does not appear to be a simple story of under/overreaction to earnings because I find the AHL regressions do not lead to consistent significant changes in ERCs across the conservatism samples. Unraveling the more complicated story is left to future research.

References

Aboody, D., Hughes J., Liu, L., 2002. Measuring value relevance in a (possibly) inefficient market. Journal of Accounting Research 40, 965-986.

Ahmed, A., Billings, B., Morton, R., Stanford-Harris, M., 2002. The role of accounting conservatism in mitigating bondholder-shareholder conflicts over dividend policy and reducing debt costs. The Accounting Review 77, 867-890.

Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. Journal of Accounting Research 6, 159-177.

Basu, S., 1997. The conservatism principle and the asymmetric timeliness of earnings. Journal of Accounting and Economics 24, 3-7.

Beaver, W., Ryan, S., 2004. Conditional and unconditional conservatism: Concepts and modeling. Working Paper. Stanford University.

Bernard, V., Thomas, J., 1989. Post-earnings-announcement drift: Delayed price response or risk premium? Journal of Accounting Research 27, 1-48.

_______, _______, 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics 13, 305-340.

Bubblitz, B., Etteredge, M., 1989. The information in discretionary outlays: Advertising, research and development. The Accounting Review 64, 108-124.

Chan L., Lakonishock, J., Sougiannis, T., 2001. The stock market valuation of research and development expenditures. Journal of Finance 56, 2431-2458.

Chambers, D., Jennings, R., Thompson, R., 2002. Excess returns to R&D-intensive firms. Review of Accounting Studies 7, 133-158.

Deng, Z., Lev, B., 1998. The valuation of acquired R&D, Working Paper, New York University.

Financial Accounting Standards Board (FASB). 1984. Recognition and measurement in financial statements of business enterprises. Statement of Financial Accounting Concepts No. 5. Norwalk, CT: FASB.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 46, 427-465.

Feltham, G., Ohlson, J., 1995. Valuation and clean surplus accounting for operating and financial activities. Contemporary Accounting Research 11, 689-731.

Frankel, L., Lee, C., 1998. Accounting valuation, market expectation, and cross-sectional stock returns. Journal of Accounting and Economics 25, 283-319.

Givoly, D., Hayn, C., 2000. The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? Journal of Accounting and Economics 29, 287-320.

Greenball, M., 1969. Appraising alternative methods of accounting for accelerated tax depreciation: A relative-accuracy approach. Journal of Accounting Research 7:2, 262-289.

Hayn, C., 1995. The information content of losses. Journal of Accounting and Economics 20, 125-153.

Hirshey, M., Weygandt, J., 1985. Amortization policy for advertising and research and development expenditures. Journal of Accounting Research 23, 326-335.

Kothari, S., 2001. Capital markets research in accounting. Journal of Accounting and Economics 31, 105-231.

Kothari, S., Sloan, R., 1992. Information in prices about future earnings: Implications for earnings response coefficients. Journal of Accounting and Economics 15, 143-171.

Lev, B., Sougiannis, T., 1996. The capitalization, amortization, and value-relevance of R&D. Journal of Accounting and Economics 21, 107-108.

Lev, B., Sarath, B., Sougiannis, T., 2000. R&D-related reporting biases and their consequences. Working Paper, New York University.

Liu, J., Thomas, J., 2000. Stock returns and accounting earnings. Journal of Accounting Research 38, 71-101.

Mason, L., 2004. The impact of accounting conservatism on the magnitude of the differential information content of cash flows and accruals. Journal of Accounting, Auditing, and Finance, 19, 249-282.

Ou, J., Penman S., 1989. Financial statement analysis and the prediction of stock returns. Journal of Accounting and Economics 11, 295-329.

Penman, S., Zhang, X., 2002. Accounting conservatism, the quality of earnings, and stock returns. Accounting Review 77, 237-264.

Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71: 3, 289-315.

Staubus, G., 1985. An induced theory of accounting measurement. The Accounting Review 60, 53-75.

Sterling, R., 1970. Theory of the measurement of enterprise income. Lawrence, Kansas: University of Kansas Press, p. 256.

Figure 1. Overview of the Five Analyses

<table>
<tr><td colspan="2" rowspan="3"></td><td colspan="3">Identifying the Earnings Impact of Accounting Conservatism</td></tr>
<tr><td colspan="3">Conservatism Measure</td></tr>
<tr><td>Penman and Zhang's (2002) Q-Score</td><td>Mason's (2004) Conservatism Score (Modified)</td><td>Givoly and Hayn's Accrual-Based Measure of Conservatism (2000)</td></tr>
<tr><td rowspan="2">Measuring Market Inefficiency</td><td>Return Predictability</td><td>Table 3 (Replication)</td><td>Table 5</td><td>Table 7</td></tr>
<tr><td>Aboody, Hughes and Liu Regressions (2002)</td><td>Table 4</td><td>Table 6</td><td></td></tr>
</table>

TABLE 1
Regression Descriptive Statistics

Variable	n	Mean	Median	STD	Q1	Q3
OI_{it}	61,124	1.703	0.978	2.967	0.038	2.548
P_{it}	60,084	15.158	9.875	18.289	4.125	20.250
OI_{it}/P_{it-1}	60,038	0.087	0.091	0.203	0.009	0.178
R_{it}	60,482	0.131	0.031	0.580	-0.250	0.376
R_{it+1}	60,218	0.174	0.039	0.830	-0.243	0.389
R_{it+2}	60,220	0.350	0.108	1.823	-0.300	0.625
R_{it+3}	60,222	0.516	0.161	2.398	-0.333	0.816
AR_{it+1}	60,213	0.150	-0.042	0.978	-0.404	0.423
AR_{it+2}	57,396	0.164	-0.114	1.986	-0.506	0.410
AR_{it+3}	54,749	0.148	.0.171	1.843	-0.578	0.392
Q	61,176	-0.000	-0.000	0.246	-0.012	0.012

OI_t is operating income per share; P_{it} is stock price three months after the fiscal year end *t*. R_{it} is firm *i*'s current period return ending three months after fiscal year *t*; $R_{it+\tau}$ (τ = 1, 2, 3) are firm *i*'s returns for 12, 24, 36 months beginning three months after the fiscal year end *t*. $AR_{it+\tau}$ (τ = 1, 2, 3) are firm *i*'s size-adjusted returns for 12, 24, 36 months beginning three months after the fiscal year end *t*. Q is Penman Zhang's (2002) measure of earnings quality which captures the interactive effect of accounting conservatism and change in investment.

TABLE 2
Firm Characteristics by Level of Q-Score

Panel A: All Observations

	N	Mean	Std Dev	Q1	Median	Q3
Market Value of Equity*	60,990	815.35	7231.19	14.53	52.84	226.34
Book-to-Market	60,988	1.55	135.65	0.34	0.62	1.05
Earnings per Share	61,142	0.65	2.18	-0.10	0.47	1.34
Return on Equity**	48,126	-0.01	0.32	-0.01	0.06	0.11
Return on Assets	47,483	0.02	.173	0.01	0.06	0.10
Debt to Total Assets	61,176	0.17	0.18	0.02	0.13	0.27
Estimated Reserves*	61,176	32.41	214.16	0.59	2.49	8.96
Net Operating Assets*	61,176	421.21	2947.56	10.74	39.05	149.63

Panel B: Lowest Q- Decile

	N	Mean	Std Dev	Q1	Median	Q3
Market Value of Equity*	4,932	502.38	7457.45	9.65	32.59	121.79
Book-to-Market	4,932	0.72	3.20	0.23	0.47	0.87
Earnings per Share	4,952	0.52	1.93	-0.14	0.23	1.01
Return on Equity**	3,270	-0.02	0.38	-0.03	0.06	0.12
Return on Assets	3,240	0.01	0.23	-.01	0.06	0.11
Debt to Assets	4,956	0.14	0.18	0.00	0.07	0.22
Estimated Reserves*	4,956	20.94	118.27	0.76	2.55	7.90
Net Operating Assets*	4,956	146.63	761.40	4.45	12.77	45.46

Panel C: Highest Q-Decile

	n	Mean	Std Dev	Q1	Median	Q3
Market Value of Equity*	4,854	400.91	4353.00	7.49	25.03	89.83
Book-to-Market	4,854	0.67	1.36	0.22	0.53	1.01
Earnings per Share	4,877	-0.05	2.73	-0.77	-0.07	0.75
Return on Equity**	3,458	-0.21	0.57	-0.32	-0.01	0.09
Return on Assets	3,480	-0.12	0.32	-0.23	0.00	0.08
Debt to Assets	4,880	0.13	0.21	0.00	0.05	0.19
Estimated Reserves*	4,880	28.50	179.09	1.09	3.46	10.58
Net Operating Assets*	4,880	148.98	1047.76	2.77	9.44	36.04

*in millions
** Firm-years with negative equity are deleted

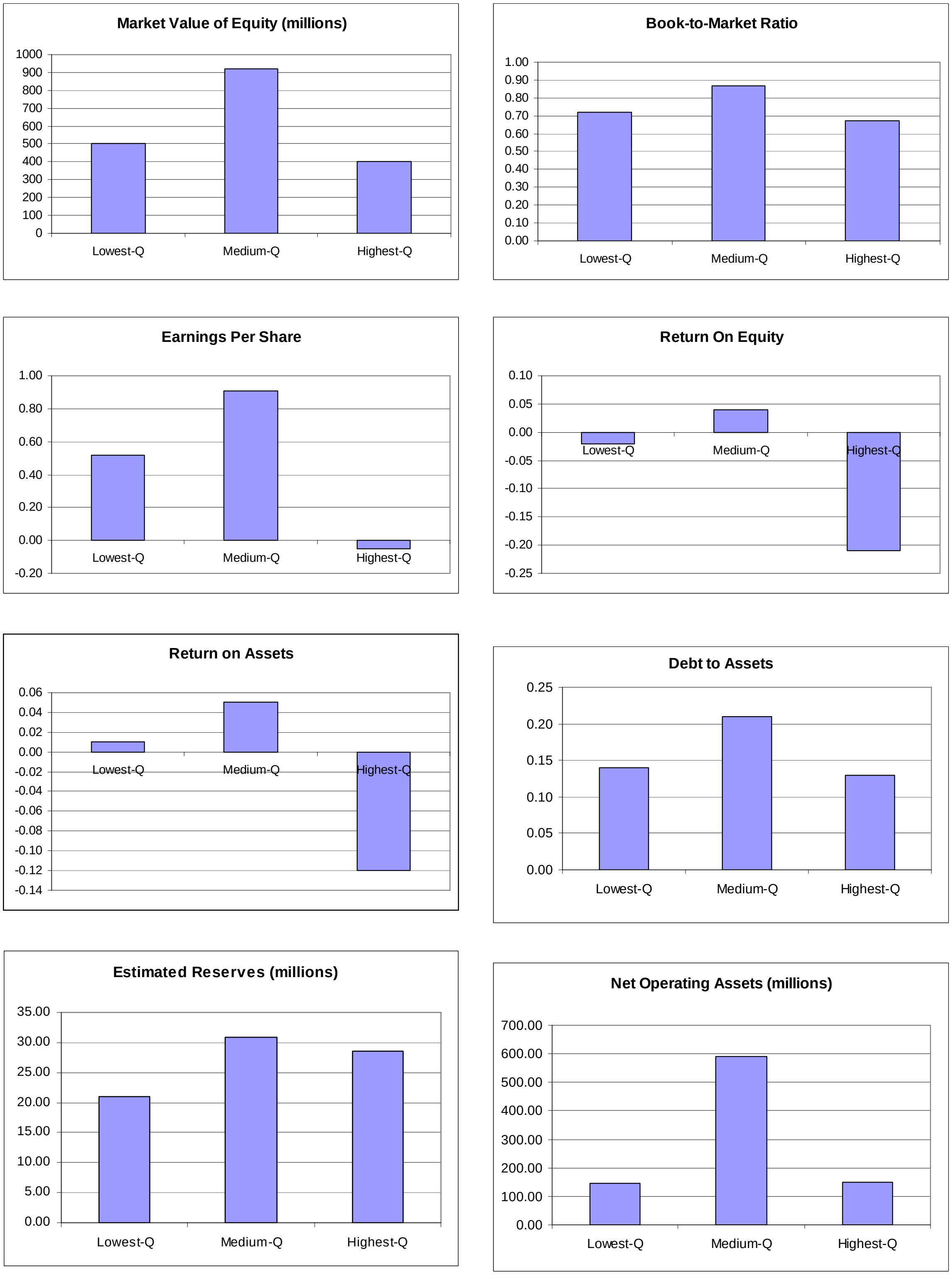
TABLE 2, continued

Panel D: Medium Q-Deciles 5 and 6 (where Q- score equals approximately zero)

	n	Mean	Std Dev	Q1	Median	Q3
Market Value of Equity*	12,957	921.78	6948.00	20.55	78.12	353.95
Book-to-Market	12,956	0.87	1.27	0.41	0.68	1.11
Earnings per Share	12,992	0.91	1.98	0.08	0.71	1.55
Return on Equity**	10,763	0.04	0.18	0.02	0.07	0.11
Return on Assets	10,677	0.05	0.09	0.03	0.06	0.10
Debt to Assets	12,999	0.21	0.18	0.05	0.18	0.30
Estimated Reserves*	12,999	30.80	218.45	0.33	1.62	6.65
Net Operating Assets*	12,999	590.96	3439.61	23.77	77.00	270.62
*in millions ** Firm-years with negative equity are deleted						

Market value of equity equals total number of common shares outstanding as of year-end multiplied by end-of-year stock price. The book-to-market ratio is shareholder's equity divided by the market value of equity. Earnings per share is basic earnings per share: Net income (including extraordinary items) divided by the weighted-average number of common shares outstanding (Compustat data item 53). Return on assets equals net income before interest expense divided by average total assets. Return on equity equals net income less dividends divided by average shareholder's equity. Firm-years with negative equity are deleted before calculating descriptive statistics for return on equity. Dividing total debt by total assets equals the debt to asset ratio. Estimated reserves comprise the numerator in the Penman and Zhang (PZ) (2002) Q-score and equal the sum of estimated unrecorded assets arising from conservative accounting for R&D, advertising and inventory; net operating assets in the Q-score denominator are defined as assets minus liabilities, excluding financial assets and liabilities. In footnote 8, p. 242, PZ (2002) describe in detail the calculation of net operating assets using Compustat item numbers.

EXHIBIT 1
Descriptive Statistics by Level of Q-Score*



*see Table 2 note

TABLE 3
Mean Percentage Annual Stock Returns for Portfolios Formed on Q-Scores

	Year Relative to Year Q-Score is Calculated (Year 0)							
Q Portfolios	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*	*5*
Panel A: Raw Returns								
Lowest Q	.222	.328	.179	.123	.216	.216	.165	.152
2	.247	.259	.198	.175	.221	.212	.171	.245
3	.232	.253	.169	.213	.178	.206	.232	.182
4	.203	.253	.136	.216	.200	.214	.159	.135
5	.240	.192	.120	.192	.193	.193	.164	.164
6	.189	.170	.134	.202	.231	.179	.178	.220
7	.183	.123	.091	.226	.222	.249	.203	.187
8	.206	.144	.116	.227	.248	.173	.222	.194
9	.195	.107	.109	.256	.257	.204	.203	.197
Highest Q	.177	.085	.067	.314	.187	.232	.180	.173
High-Low	-.044	-.243	-.112	.191	-.029	.015	.015	.021
p-value	*.2673*	*<.0001*	*.0003*	*<.0001*	*.4959*	*.7556*	*.7539*	*.6000*
Firm-years	59,224	62,437	62,754	62,432	57,058	50,101	43,934	38,508

TABLE 3, continued

	Year Relative to Year Q-Score is Calculated (Year 0)							
Q Portfolios	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*	*5*
Panel B: Size-Adjusted Returns								
Lowest Q	.086	.175	.016	-.003	.067	.046	.061	.060
2	.106	.163	.024	.014	.061	.047	.070	.042
3	.078	.100	.022	.018	.031	.060	.062	.052
4	.062	.073	-.004	.013	.055	.069	.037	.050
5	.056	.037	-.007	.022	.045	.035	.056	.043
6	.035	-.009	-.017	.036	.055	.036	.044	.065
7	.030	-.019	-.032	.065	.063	.049	.066	.048
8	.044	-.024	-.029	.063	.103	.063	.048	.054
9	.032	-.034	-.032	.070	.100	.071	.075	.084
Highest Q	-.001	-.029	-.050	.094	.085	.105	.034	.084
High-Low	-.088	-.204	-.066	.097	.019	.058	-.027	.024
p-value	*<.0001*	*<.0001*	*<.0001*	*<.0001*	*.4401*	*.0411*	*.2518*	*.3953*
Firm-years	56,850	59,772	60,062	59,748	54,666	47,989	42,102	36,930

Buy-and-hold returns for each Q-score portfolio are calculated for each year, -2 to +5, in each scoring year 1976 through 2002. Panel A reports the mean returns over the sample period. Panel B reports the mean size-adjusted returns, computed by subtracting the raw (buy and hold) return on a size-matched, value-weighted portfolio formed from CRSP size-deciles.

TABLE 4
Return Regressions on Operating Income
(Q-Score)

Stock return is cross-sectionally regressed on operating income per share deflated by stock price of the previous year. Summary statistics on the annual cross-sectional regression coefficients are reported in Panel A. To control for market inefficiencies, firm *i*'s 12-month stock return beginning 3 months after fiscal year-end *t*-1, is multiplied by $\left(\frac{1+R_{it+\tau}}{1+\overline{R}^{\upsilon}_{it+\tau}}\right)$ to form $\hat{R}_{it}$, where $R_{it+\tau}$ (τ =1,2, 3) is firm *i*'s stock return for 12, 24, 36 months beginning three months after the fiscal year end *t*, and $\overline{R}^{\upsilon}_{it+\tau}$ is the corresponding size-decile return. The Q-score groups are based on a ranking of firms each year on Q-score, within two-digit industry codes. The high Q-score group is the top 10% of firms in the ranking on Q, the medium Q-score groups are the middle 20% (where Q equals approximately 0) and the low Q-score group is the bottom 10% of firms in the ranking on Q-score. For Panel B, the percentage change in coefficients is defined as the coefficient from the adjusted regression minus the corresponding coefficient from the unadjusted regression deflated by the unadjusted coefficient. Statistical tests are based on 27 annual observations of change in coefficients from 1976 to 2002 for the one-year adjustment horizon. Statistical tests for the two (three)-year adjustments are based on 26 (25) annual regressions from 1976 to 2001 (2000). Panel C reports tests of significance for differences in percentage change percentages across Q-score groups. Reported *p*-values for mean (median) tests are based on *t*-statistics (signed-rank statistics).

Panel A: Return regression coefficients estimates

All coefficients are significant at p < .0001

Level Q-Score	Adjustment Horizon	Unadjusted	One-year τ=1	Two-year τ=2	Three -year τ=3
All	Mean	0.744	0.911	1.045	1.095
Obs	Median	0.786	1.027	1.066	1.102
Lowest	Mean	0.690	0.941	1.073	1.129
	Median	0.690	0.955	1.098	0.983
Q ≈ 0	Mean	0.843	1.001	1.017	1.107
	Median	0.882	1.163	1.077	1.019
Highest	Mean	0.657	0.763	1.367	1.519
	Median	0.747	0.771	1.034	1.132
Firm-years		59,396	59,148	56,337	53,699

TABLE 4, continued

Panel B: Summary statistics for the percentage change on the coefficient of Operating Income Per Share

Adjustment Horizon	One-year τ=1	Two-year τ=2	Three-year τ=3
All Q			
Mean	0.167	0.209	0.205
Median	0.252	0.394	0.402
p-value for mean test	*.0219*	*.3409*	*.4736*
p- value for median test	*.0020*	*.0012*	*.0014*
Lowest Q			
Mean	1.785	0.952	0.245
Median	0.198	0.370	0.505
p- value for mean test	*.1952*	*.0300*	*.6296*
p- value for median test	*.0043*	*.0055*	*.4244*
Q ≈ 0			
Mean	-0.058	-0.443	-0.578
Median	0.274	0.261	0.332
p- value for mean test	*.7751*	*.4806*	*.5249*
p- value for median test	*.1087*	*.1403*	*.0986*
Highest Q			
Mean	0.124	1.491	1.361
Median	0.117	0.562	0.656
p- value for mean test	*.4749*	*.0985*	*.0069*
p- value for median test	*.5423*	*<.0001*	*<.0001*

TABLE 4, continued

Panel C: Summary statistics for tests of differences in percentage changes across Q-groups

	Adjustment Horizon		
Difference in Percentage Changes	One-year τ=1	Two-year τ=2	Three-year τ=3
High Q - Low Q			
Mean	-1.661	0.540	1.117
Median	-0.418	0.197	0.362
p-value for mean test	*.2564*	*.4011*	*.1755*
p-value for median test	*.0208*	*.1965*	*.1793*
High Q - Medium Q			
Mean	0.182	1.934	1.940
Median	-0.116	0.578	0.668
p-value for mean test	*.3132*	*.0512*	*.0248*
p-value for median test	*.9442*	*.0059*	*.0025*
Medium Q - Low Q			
Mean	-1.843	-1.395	-0.823
Median	-0.252	-0.370	-0.122
p-value for mean test	*.2197*	*.0346*	*.4097*
p-value for median test	*.0527*	*.0078*	*.3362*

TABLE 5
Mean Percentage Annual Stock Returns for Portfolios Formed on Mason's Conservatism Score (Modified)

Year Relative to Year Score is Calculated (Year 0)

Score Portfolios	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*	*5*
Size-Adjusted Returns								
Lowest	.130	.190	.073	.005	.027	.007	.033	.001
2	.133	.191	.048	.035	.064	.041	.046	.043
3	.113	.157	.034	.029	.050	.048	.050	.031
4	.102	.095	.022	.041	.040	.062	.033	.008
5	.074	.077	.018	.047	.044	.036	.033	.030
6	.050	.035	.026	.082	.054	.035	.033	.027
7	.062	-.020	-.004	.083	.087	.054	.056	.035
8	.022	-.041	-.008	.069	.124	.084	.033	.024
9	-.015	-.060	-.015	.096	.093	.065	.078	.060
Highest	-.022	-.079	.018	.112	.141	.113	.043	.047
High-Low	-.153	-.269	-.091	.107	.114	.106	.010	.046
p-value	*<.0001*	*<.0001*	*<.0001*	*<.0001*	*.0005*	*.0136*	*.7916*	*.1566*
Firm-years	29,028	30,541	30,737	30,593	28,087	25,072	22,522	20,379

Buy-and-hold returns for each accrual-measure portfolio are calculated for each year, -2 to +5, in each scoring year 1976 through 2002. Panel A reports the mean returns over the sample period. Panel B reports the mean size-adjusted returns, computed by subtracting the raw (buy and hold) return on a size-matched, value-weighted portfolio formed from CRSP size-deciles.

TABLE 6
Return Regressions on Operating Income
(Mason's Score - Modified)

Stock return is cross-sectionally regressed on operating income per share deflated by stock price of the previous year. Summary statistics on the cross-sectional regression coefficients are reported in Panel A. To control for market inefficiencies, firm *i*'s 12-month stock return beginning 3 months after fiscal year-end *t*-1, is multiplied by $\left(\frac{1+R_{it+\tau}}{1+\overline{R}^{\upsilon}_{it+\tau}}\right)$ to form $\hat{R}_{it}$, where $R_{it+\tau}$ (τ =1,2, 3) is firm *i*'s stock return for 12, 24, 36 months beginning three months after the fiscal year end *t*, and $\overline{R}^{\upsilon}_{it+\tau}$ is the corresponding size-decile return. The groups are based on a ranking of firms each year on Mason's (2004) differenced conservatism score, within two-digit industry codes. The high -score group is the top 10% of firms in the ranking on the score, the medium score groups are the middle 20% (where the score equals approximately 0) and the low-score group is the bottom 10% of firms in the ranking. For Panel B, the change in coefficients is defined as the coefficient from the adjusted regression minus the corresponding coefficient from the unadjusted regression deflated by the unadjusted coefficient. Statistical tests are based on 27 annual observations of change in coefficients from 1976 to 2002 for the one-year adjustment horizon. Statistical tests for the two (three)-year adjustments are based on 26 (25) annual regressions from 1976 to 2001 (2000).

Panel A: Return regression coefficients estimates
All coefficients are significant at *p*-value < .0001 except as noted

Level Mason's Modified Score	Adjustment Horizon	Unadjusted	One-year τ=1	Two-year τ=2	Three -year τ=3
All	Mean	0.666	0821	0.927	1.058
Obs	Median	0.701	.958	1.071	1.144
Lowest	Mean	0.765	0.999	1.324	1.194
	Median	0.775	0.828	1.168	1.089
Q ≈ 0	Mean	0.752	1.013	1.151	0.867
	Median	0.780	0.904	0.932	0.930**
Highest	Mean	0.646	0.678	1.613*	1.490*
	Median	0.623	0.628	0.788	0.981
Firm-years		29,787	29,676	28,795	28,058

* significant at *p*-value = .01
** significant at *p*-value = .001

TABLE 6, continued

Panel B: Summary statistics for the percentage change on the coefficient of Operating Income Per Share

Adjustment Horizon	One-year τ=1	Two-year τ=2	Three -year τ=3
All Obs			
Mean	0.629	0.711	1.014
Median	-0.028	-0.451	-0.419
p-value for mean test	.4581	.5860	.5487
p-value for median test	.0015	.0025	.0146
Lowest			
Mean	-0.413	-2.263	-0.052
Median	-0.094	-0.287	-0.089
p-value for mean test	.4851	.2552	.9198
p-value for median test	.7011	.5572	1.000
Score ≈ 0			
Mean	-0.210	-0.507	0.154
Median	-0.428	-0.375	-0.375
p-value for mean test	.3402	.2100	.7778
p-value for median test	.0192	.0094	.2295
Highest			
Mean	-0.294	-2.328	-1.421
Median	-0.182	-0.319	-0.481
p-value for mean test	.2822	.1267	.0111
p-value for median test	.2478	.1686	.0433

TABLE 7
Mean Percentage Annual Stock Returns for Portfolios Formed on Givoly and Hayn's (2000) Accrual Measure

	Year Relative to Year Accrual Measure is Calculated (Year 0)							
Accrual-Measure Portfolios	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*	*5*
Size-Adjusted Returns								
Lowest	.161	.152	.092	.028	.085	.095	.116	.094
2	.057	.057	.047	.053	.070	.067	.073	.083
3	.026	.057	.048	.035	.068	.109	.061	.086
4	.039	.045	.044	.045	.076	.054	.060	.073
5	.021	.034	.038	.061	.069	.055	.071	.067
6	.017	.054	.076	.052	.069	.064	.063	.077
7	.030	.057	.064	.073	.106	.092	.074	.088
8	.027	.068	.101	.081	.077	.084	.097	.111
9	.058	.041	.089	.091	.112	.112	.091	.100
Highest	.041	.057	.040	.097	.105	.076	.112	.168
High-Low	-.120	-.095	-.053	.070	.020	-.018	-.004	.075
p-value	*.0015*	*.0143*	*.1132*	*.0254*	*.5395*	*.5826*	*.9371*	*.0843*
Firm-Years	27,964	27,965	27,967	27,804	25,316	20,991	17,192	13,823

Buy-and-hold returns for each accrual-measure portfolio are calculated for each year, -2 to +5, in each scoring year 1975 through 2002. Table reports the mean size-adjusted returns, computed by subtracting the raw (buy and hold) return on a size-matched, value-weighted portfolio formed from CRSP size-deciles.

TABLE 8
Pearson Correlations of Q with Factors Associated with Market Inefficiency

	Size	Book-to-Market	Leverage	E/P	Accruals
Q	-0.0116	-.0002	-.0006	-.0730	-.1911
(*p*-value)	(.004)	(.9593)	(.8683)	(< .0001)	(< .0001)

Size is a firm's end-of-year decile-rank of market capitalization provided by CRSP, leverage is the debt to equity ratio, and E/P is earnings per share divided by price. Accruals equal net income before depreciation less cash flows from operations scaled by total assets.