A Returns-Based Representation of Earnings Quality

Frank Ecker
(Duke University)

Jennifer Francis*
(Duke University)

Irene Kim
(Duke University)

Per Olsson
(Duke University)

Katherine Schipper
(Financial Accounting Standards Board)

We examine the properties of a returns-based representation of earnings quality, estimated from firm-specific asset pricing regressions augmented by an earnings quality mimicking factor. The coefficient on the earnings quality factor (the "e-loading") captures the sensitivity of the firm's returns to earnings quality in a given year or quarter, analogous to beta as a measure of the sensitivity of returns to market movements. Relative to other proxies for earnings quality, e-loadings can be calculated for larger samples of firms and can be estimated for shorter intervals at any point in time. Along all dimensions examined, we find that e-loadings perform well in capturing notions of earnings quality.

December 2005

* Corresponding author: Fuqua School of Business, Duke University, Durham, NC 27708. Email address, jfrancis@duke.edu. This research was supported by the Fuqua School of Business, Duke University. The views expressed in this paper are those of the authors, and do not represent positions of the Financial Accounting Standards Board. Positions of the Financial Accounting Standards Board are arrived at only after extensive due process and deliberation. Analyst data are from Zacks Investment Research Database. We have benefited from helpful comments from two anonymous reviewers, Moshe Bareket, Qi Chen, Dan Dhaliwal, Barry Lewis, Michael Smith, Mohan Venkatachalam, and Yun Zhang and from workshop participants at Colorado State University, Duke University, ISCTE Business School - Lisbon, Swedish Institute for Financial Research, 2005 Southeast Accounting Research Conference, Temple University and the University of Regensburg.

1. Introduction

We analyze a returns-based measure of earnings quality that can be estimated for a given firm-year (in extensions, we show the results are generalizable to a firm-quarter). Our measure is the slope coefficient from a regression of a firm's daily excess returns in year T on a factor mimicking portfolio capturing earnings quality, controlling for other factors known to affect returns (market risk premium, size and book-to-market ratio). We build on the factor mimicking portfolio approach introduced by Fama and French (1993) and applied by Francis, LaFond, Olsson and Schipper (2005) to demonstrate a risk premium for firms with poorer earnings quality, as captured by Dechow and Dichev's (2002) measure of accruals quality (*AQ*). Specifically, we follow Francis et al.'s (2005) procedures to create an *AQ* factor mimicking portfolio (*AQfactor*), except that we use daily returns rather than monthly returns to estimate the asset pricing regressions.[1] Next, for each firm with at least 100 daily returns observations in year T=1970-2003, we estimate 1-factor and 3-factor annual regressions that add *AQfactor* as an independent variable. Just as the CAPM beta captures exposure to market risk, so too will the coefficient on *AQfactor* in these regressions – the e-loading – capture investor perceptions of the firm's earnings quality exposure in year T (a larger e-loading implies greater sensitivity to poor earnings quality).

Capturing the returns consequences of a firm characteristic by applying a factor-mimicking portfolio is not new, nor are construct validity tests of those returns consequences. For example, Fama and French (1993) propose and validate the use of factor-mimicking portfolios and associated loadings as proxies for firm size and book-to-market;[2] and Lamont, Polk and Saá-Requejo (2001) probe the existence of a factor capturing the degree to which firms are financially constrained. Our paper extends this body of work by proposing and validating e-loadings as a returns-based representation of earnings quality.

[1] Using monthly returns in asset pricing tests results in less estimation error than using daily returns. However, if the true asset pricing parameters vary over time, monthly data will obscure this pattern, since a longer time period is needed to estimate an asset pricing regression with monthly data. Most studies require a minimum of 24 to 60 monthly observations to estimate an asset pricing regression; consequently, the resulting coefficient estimates are effectively assumed constant over this 2-5 year interval.

[2] The degree to which factor-mimicking portfolios proxy for underlying size and book-to-market constructs continues to be debated. In cross-sectional tests, Daniel and Titman (1997) find that additional variation in returns can be explained by firms' size and book-to-market ratios over and above 3-factor returns and loadings. However, using time-series tests, Lewellen (1999) documents no incremental information about expected returns in book-to-market ratios, beyond 3-factor returns and loadings.

We view earnings quality as a measure of information risk, and we define earnings quality in terms of precision, namely, the mapping of current accruals into current, last year and next year cash flows. Following Dechow and Dichev (2002), we term this mapping accruals quality, and denote it by *AQ*. Theoretical support for the view that information uncertainty is a nondiversifiable (i.e., priced) risk factor is provided by Easley and O'Hara (2004)[3] and Leuz and Verrecchia (2004). Empirical support for the view that earnings quality as measured by accruals quality is priced is provided by Francis et al. (2005) who show that the returns representation of the *AQ* measure, *AQfactor*, enters 1- and 3-factor asset pricing regressions with a reliably positive average coefficient estimate. Thus, conditioning on theory that shows information uncertainty is a priced risk factor, and on Francis et al.'s (2005) empirical evidence that the market prices information risk as captured by earnings quality measured by *AQ*, we view e-loadings as capturing the sensitivity of stock returns to earnings quality.

As discussed in more detail in section 2, our paper contributes to the earnings quality literature by establishing an earnings quality metric that offers several advantages relative to traditional earnings quality measures (either measures derived from accounting data or measures derived from both returns and accounting data). First, e-loadings can be measured for firms that lack the time series of accounting data that is typically required for estimating accounting-based measures of earnings quality. Therefore, e-loadings can be measured for much larger (and more representative of the population) samples of firms, increasing sampling power and generalizability. Second, e-loadings can be reliably measured over intervals as short as a quarter, so they can be used to analyze changes in earnings quality associated with specific events. In addition, the returns-based representation approach is flexible, in that other earnings attributes, such as smoothness and persistence, can also be represented by factor loadings. As we discuss

[3] In Easley and O'Hara's (2004) model, the risk premium associated with information uncertainty is a function of private information (which pertains to information asymmetry) and the precision of public and private information. Our accruals quality measures focus on the precision of current accruals with respect to cash flows; as such, we view it as capturing the precision of public information. In extensions of our main tests, we consider the overlap between e-loadings and two trading-based measures of information asymmetry: bid-ask spreads (which reflect both public and private information) and probability-of-informed trading (PIN) scores (which arguably focus more on private information).

in section 2, these advantages permit researchers to examine earnings quality in settings and for samples which are difficult (or not possible) to examine with traditional measures of earnings quality.

Our analysis of the properties of e-loadings as measures of earnings quality has four components. First, we investigate whether e-loadings vary cross-sectionally with characteristics expected to be related to earnings quality and, separately, with other proxies for earnings quality. Second, we examine whether e-loadings are associated with predictable variation in market participant behavior with respect to earnings. These tests focus on whether investors attach lower earnings response coefficients to higher e-loading firms and whether there is greater dispersion and less accuracy in analysts' earnings forecasts for higher e-loading firms. Third, we conduct over-time analyses examining whether e-loadings exhibit systematic patterns as a function of firm age, where age proxies for the amount of information available about the firm. We predict that investor perceptions of earnings quality are more uncertain (leading to higher e-loadings) and less stable (leading to lower autocorrelation in e-loadings) for young firms where less information is available. Fourth, we examine whether e-loadings are higher in three settings associated with poor earnings quality: restatements, shareholder lawsuits, and bankruptcy.

In terms of the first analysis, we find that earnings quality determinants are significant in explaining variation in e-loadings, and that e-loadings exhibit predictably positive correlations with most other proxies for earnings quality. Our second analysis shows that firms with higher e-loadings have lower earnings response coefficients and more dispersed and less accurate analysts' forecasts, consistent with market participants perceiving higher e-loading firms as having noisier earnings signals than lower e-loading firms. Results of our third analysis, of differences in the level and stability of e-loadings as a function of firm age, are also consistent with predictions: as the firm matures, we find both an over-time decline in the magnitude of its e-loading as well as an over-time increase in the autocorrelation of its e-loading. Our fourth analysis reveals that e-loadings are larger in settings characterized by poor earnings quality: e-loadings increase prior to, and are highest during, years containing restatement announcements, lawsuit filings, or bankruptcies. The level and change in e-loadings for firms affected by these events is also significantly larger than the levels and changes observed for samples of non-event firms, matched in calendar time. Results for all analyses are significant at the 0.05 level or better.

In summary, we document the reliability of e-loadings using the following measures: (i) correlation between e-loadings and both the Dechow-Dichev innate determinants of earnings quality and seven other earnings quality proxies, (ii) the relation between e-loadings and both earnings response coefficients and forecast accuracy and dispersion, (iii) over time changes in e-loadings as a function of firm age, and (iv) the behavior of e-loadings around earnings restatements, class action lawsuits, and bankruptcies. We interpret the combined results as demonstrating the reliability of e-loadings as a returns-based representation of earnings quality, controlling for other factors known to affect returns.

To examine the robustness and generalizability of our findings, we extend our tests in several ways. First, we verify that our results are not driven by the subset of firms with the necessary data to calculate the accounting measure (*AQ*) which underlies *AQfactor*. Second, we examine whether e-loadings based on other measures of earnings quality perform as well as e-loadings based on accruals quality. Of the seven measures that we consider (persistence, predictability, smoothness, value relevance, timeliness, conservatism, and a measure of abnormal accruals), e-loadings based on mimicking factors for both persistence and smoothness appear to perform about as well as e-loadings based on accruals quality. However, none of the seven measures systematically dominates e-loadings based on accruals quality. Third, we probe the reliability of e-loadings calculated over *quarterly* estimation intervals, where we require at least 45 daily returns in quarter Q to calculate the firm-quarter e-loading. Quarterly e-loadings exhibit the same patterns found for yearly e-loadings, except that the confidence intervals are wider; however, results are significant at the 10% level or better. Fourth, we examine the overlap between e-loadings and two trading-based measures of information asymmetry – bid-ask spreads and probability of informed trading (PIN) scores. The correlations range between 0.12 and 0.16 and, while reliably different from zero, are relatively weak in economic terms. To determine the influence of spreads and PIN scores on e-loadings, we orthogonalize e-loadings with respect to these measures and repeat our tests. Results are similar, suggesting that the effects we document for e-loadings are not subsumed or driven by trading-based measures of information asymmetry, such as bid-ask spreads and PIN scores.

We believe that the strength and consistency of our results, including results of extensions, alleviate concerns that arise because *AQfactor* itself is constructed using a sample of firms with the requisite time-series data to estimate the *AQ* measure. If this sample is not sufficiently broad, the resulting returns representation of *AQ* demonstrated by *AQfactor* may not reflect the market's pricing of *AQ*, impairing our ability to identify a meaningful returns representation of earnings quality. In particular, if *AQfactor* is sufficiently noisy, we would not expect it to load positively in asset pricing regressions. Our findings that *AQfactor* exhibits significant (at the 0.001 level) average loadings for our sample and that *AQfactor* itself is associated with an economically substantial and statistically reliable (at the 0.001 level) average return indicate that whatever bias is introduced by using an accounting-based sample as the foundation for *AQfactor* does not invalidate the approach. Taken together, we believe the results support the inference that e-loadings capture the same underlying construct as reflected in other measures of earnings quality. That is, e-loadings capture what they purport to capture *and* offer several advantages, relative to accounting-based measures, in terms of flexibility and adaptability.[4]

While we believe that our results demonstrate that an earnings quality measure based on e-loadings offers significant advantages compared to existing earnings quality measures, our results do not speak to whether e-loadings are the *best* measure The choice of the best earnings quality measure will be a function of, among other things, the nature of the research question addressed, the assumptions necessary to support the chosen research design, and available data. For example, the e-loading explored in this paper is a measure of investors' perception of *total* earnings quality, not discretionary earnings quality. In a research setting that demands a measure of the portion of earnings quality that is solely

[4] Our evidence supports Chen, Shevlin and Tong's (2004) analysis of over-time changes in the information risk characteristics of firms that change dividend policies. They find that information risk (proxied by the coefficient estimate on monthly *AQfactor*) is higher in the three years following a dividend decrease relative to the three years preceding the decrease; they find mixed results for dividend-increasing firms.

attributable to managements' discretionary actions and behaviors, a different measure of earnings quality would likely be preferred.

Two other caveats about the use of e-loadings are also in order. First, unlike measures that rely exclusively on accounting data, e-loadings capture the market's *perception* of earnings quality. If market beliefs are not rational, those perceptions may differ from the reality of the firm's earnings quality. (Because our tests focus on broad samples of firms over a 34-year period, it is unlikely that our results are attributable to systematic mispricing.) Second, our analyses of the reliability of e-loadings focus on the cross-sectional distribution of empirical measures of an unobservable construct, not on the correctness of an estimated magnitude.[5] Our results have no implications for the magnitudes of the e-loadings, and theory does not predict values for these magnitudes. In this respect, e-loadings are similar to *s*- and *h*-loadings (for *SMB* and *HML*); only the CAPM beta has a value that derives from theory.

The rest of the paper is organized as follows. The next section elaborates on the characteristics of returns-based measures of earnings quality and their advantages relative to existing earnings quality measures. Section 3 lays out the construction of the *AQ*factor mimicking portfolio and the estimation of the e-loadings. Section 4 describes the main samples used in our tests, section 5 describes the empirical analyses, and section 6 presents extensions. Section 7 summarizes the findings and concludes.

2. Returns-Based Representations versus Accounting-Based Representations of Earnings Quality.

In this section we describe how a returns-based representation of earnings quality offers two distinct but related advantages over earnings quality measures derived from accounting data. (While we focus on accounting-based measures, much of the discussion also applies to measures derived from a combination of accounting and market data, for example, the explanatory power of earnings for returns.) Those advantages derive from differences in data requirements and periodicity of estimation. We also

[5] This aspect of our analysis is similar to some aspects of Botosan and Plumlee's (2005) calibration of empirical proxies for the cost of equity capital. They focus primarily on the adequacy of cross-sectional distributions by examining associations between measures of firm-specific risk (such as beta) and measures of the cost of capital.

describe the types of research settings where a returns-based measure like the one we propose and validate might be used, and where it would not be useful.

Data requirements. Measures of earnings quality based on accounting data are typically estimated using either a firm-specific time-series of annual data or industry cross-sections.[6] Either approach places significant restrictions on sample sizes and, in the case of firm-specific time-series estimation, biases the sample toward surviving firms which tend to be larger and more profitable. To illustrate this point, we refer to Table 1 of Dechow-Dichev (2002), showing the derivation of their sample used to measure accruals quality as the mapping of current accruals into this year's, last year's and next year's cash from operations (*CFO*). A key determinant of the sample size and composition is how many firms and of what type have data on current accruals and *CFO*. Table 1 in Dechow and Dichev shows that 55,850 firm-years between 1987-1999 have data on *CFO*, earnings, changes in accounts receivable and inventory and total assets (after truncation of the most extreme 1% of observations). However, only 30,317 firm years have both lead and lag values of *CFO*—reducing the sample size by nearly 46% -- and 15,234 firm years (1,725 firms) have eight or more annual observations necessary to estimate firm-specific regressions—a reduction of nearly 73% relative to the original sample.

The alternative approach of estimating measures in industry cross-sections also places severe restrictions on sample size and composition. For example, Xie (2001) follows Subramanyam's (1996) cross-sectional modified Jones model estimation and reports a sample size of 56,692 firm-years over 1971-1992 for a relatively broad industry definition: at least six firms in a 2-digit SIC code (he also excludes Nasdaq firms prior to 1982). Over this period, there are 106,645 firm-years with non-missing returns data on CRSP (also excluding Nasdaq firms prior to 1982), indicating a sample loss of about 47%. The decrease in sample size is more extreme if the industry requirements are tightened either by requiring more firms in the industry or by using more precise industry definitions. For example, Dechow-Dichev (2002, table 1) report that, starting with 55,850 firm-year observations, 27,204 firm-year observations

[6] An alternative to the use of annual data is to use quarterly accounting data; this approach introduces possible seasonal effects as well as sampling restrictions arising from availability of quarterly data.

(136 three-digit industries) remain—a reduction of about 51%--after imposing the requirements of three years of *CFO* data and at least 50 observations per industry.

To summarize, we are not aware of accounting-based measures of earnings quality that do not impose data constraints that result in substantial decreases in sample size and (in the case of firm-specific time-series estimation) increases in survivorship bias. The decreases in sample size that arise in firm-specific time-series estimations are well known. Industry-estimation can be done using short series (e.g., as short as two or three years) but even this restriction reduces sample sizes and the requirement for a minimum number of firms per industry further restricts the sample. In contrast, a returns-based representation requires only a sufficient *returns* series. We show that, despite the additional noise introduced by daily returns, reliable estimates of e-loadings can be obtained for periods as short as one quarter. Thus, e-loadings can be calculated for larger and more representative (of the population) samples than can accounting-based measures of earnings quality. This is a distinct advantage in research settings where generalizability is important or where posited effects are expected to be most likely present in smaller, younger firms that lack the data necessary to calculate accounting-based measures.

<u>Periodicity of measurement.</u> In extensions of our main findings (reported in section 6), we show that e-loadings can be reliably calculated over intervals as short as 45 trading days, or roughly one quarter. Further, these 45 days need not be aligned with reporting periods. These two features mean that e-loadings can be used to examine shifts in earnings quality over short intervals and around events that occur at any time. In contrast, accounting-based measures of earnings quality are by construction linked to annual or quarterly reporting periods, cannot be applied to short intervals, and cannot be specific to a given financial statement date because they will be based on both current and prior data.

<u>Research settings where returns-based representations of earnings quality offer advantages.</u> The two advantages of e-loadings (or other returns-based representations of earnings quality) previously discussed point to the places where these earnings quality measures can be particularly useful. Because e-loadings can be calculated for samples comprised of younger smaller firms that have (at least) one or more quarters of daily returns data but lack a time series of accounting data, research questions involving,

for example, earnings quality for firms that have gone public within the last two years are more readily addressed using a returns-based representation.

While the sampling advantage means that using e-loadings increases sampling power and generalizability, the periodicity advantage means that e-loadings can be used when the research question of interest pertains to an event, occurring at any time, that has the potential to shift investor perceptions of earnings quality. Our reliability assessments in section 5.4, for example, focus on shifts in investor perceptions of earnings quality surrounding bankruptcies, restatements and class actions. Other events where analysis of e-loadings could provide useful insights include adoptions (including early adoptions) of new accounting standards, voluntary accounting changes, and management or auditor changes that are prompted by concerns about financial reporting quality. In some settings, a shift in earnings quality might appear in conjunction with a shift in fundamental risks. Examples include changes in capital structure, mergers/acquisitions and divisive restructurings (e.g., spin-offs that change the number and nature of segments and therefore the innate factors that influence earnings quality). In these settings, an analysis of shifts in information risk would have to control for the effects of shifts in fundamental risks.

Finally, the approach used to develop e-loadings is flexible with regard to other earnings attributes. That is, the factor-mimicking approach we use could be applied to attributes called for by a specific research question, such as smoothness or persistence. Therefore, if the research question requires, for example, a short-interval assessment of possible shifts in investor perceptions of earnings persistence, it would be possible to apply the approach we use to develop persistence-based e-loadings.

<u>Research settings where returns-based representations of earnings attributes are inapplicable.</u> There are two kinds of settings where the approach we describe would be either inapplicable or perhaps of low power. With regard to the issue of power, recall that the e-loading measure captures investor perceptions of *total* earnings quality; however, in some settings it is *discretionary* earnings quality that is of interest. Previous research (Francis et al., 2005) shows that the portion of accruals quality that is due to management's reporting choices is less priced than is total or innate accruals quality. This distinction has

implications for the power of tests requiring a measure of the discretion in management's short term decisions.

With regard to the question of applicability, a returns-based representation of an earnings attribute cannot capture the sign of the underlying factor (e.g., positive or negative abnormal accruals). A returns-based representation would not, therefore, be applicable in settings which require tests of directional predictions concerning opportunistic earnings management (e.g., tests of whether earnings are managed up or down in response to some posited incentive).

3. AQ and AQFactor Mimicking Portfolios

Our main tests use Dechow and Dichev's (2002) accruals quality metric to capture earnings quality. This choice is based on Francis et al.'s (2004) finding that accruals quality has a stronger association with the cost of equity than other earnings attributes. As discussed in Francis et al. (2005, section 2.1) theories developed by Easley and O'Hara (2004) and Leuz and Verrecchia (2005) predict that information risk is priced by investors. Positing that investors price securities based on information about cash flows, Francis et al. use accruals quality (which captures the imprecision of the mapping between current accruals and cash flows) as a measure of information risk.

We operationalize accruals quality using McNichols' (2002) modification of Dechow and Dichev's model:

$$TCA_{j,T} = \phi_{0,j} + \phi_{1,j}CFO_{j,T-1} + \phi_{2,j}CFO_{j,T} + \phi_{3,j}CFO_{j,T+1} + \phi_{4,j}\Delta \text{Rev}_{j,T} + \phi_{5,j}PPE_{j,T} + \upsilon_{j,T} \qquad (1)$$

where $TCA_{j,T} = \Delta CA_{j,T} - \Delta CL_{j,T} - \Delta Cash_{j,T} + \Delta STDEBT_{j,T}$ = total current accruals in year T; $CFO_{j,T} = NIBE_{j,T} - TA_{j,T}$ = firm j's cash flow from operations in year T;[7] $NIBE_{j,T}$ = firm j's net income before extraordinary items (Compustat #18) in year T; $TA_{j,T}$ = $(\Delta CA_{j,T} - \Delta CL_{j,T} - \Delta Cash_{j,T} + \Delta STDEBT_{j,T} - DEPN_{j,T})$ = firm j's total accruals in year T; $\Delta CA_{j,T}$ = firm j's change in current assets (Compustat #4) between year T-1 and year T; $\Delta CL_{j,T}$ = firm j's change in current liabilities (Compustat #5) between year T-1 and year T; $\Delta Cash_{j,T}$ = firm j's change in cash

[7] We calculate total accruals using information from the balance sheet and income statement rather than from the statement of cash flows (as advocated by Hribar and Collins, 2002) because statement of cash flow data are not available prior to 1988 (the effective year of SFAS No. 95) and our *AQ* metric requires seven yearly observations.

(Compustat #1) between year T-1 and year T; $\Delta STDEBT_{j,T}$ = firm j's change in debt in current liabilities (Compustat #34) between year T-1 and year T; $DEPN_{j,T}$ = firm j's depreciation and amortization expense (Compustat #14) in year T; $\Delta Rev_{j,T}$ = firm j's change in revenues (Compustat #12) between year T-1 and year T; and $PPE_{j,T}$ = firm j's gross value of property, plant and equipment (Compustat #7) in year T. All variables are scaled by average assets.

We estimate (1) in annual industry cross-sections for each of the 48 Fama-French industries with at least 20 firms in that industry-year. These estimations produce firm-year residuals, $\hat{\upsilon}_{j,T}$. Our accounting-based earnings quality metric for firm j in year T is the standard deviation of firm j's residuals over the past five years, $AQ_{j,T} = \sigma(\hat{\upsilon}_{j,T})$, T= T-5,…,T-1.[8] Large (small) values of $AQ_{j,T}$ correspond to poor (good) accruals quality.

We assign firms to *AQ* deciles using a dynamic portfolio technique that allows for differences in firms' fiscal year ends as well as over-time changes in accruals quality. Specifically, we form deciles on the first day of each month *m* based on the firm's most recent value of *AQ* known prior to *m*; firms with the smallest (largest) *AQ* values are placed in the first (tenth) decile. This means that firm j's *AQ* signal for fiscal year T, where fiscal year T ends in month *n*, will influence firm j's ranking for months *n*+4 through *n*+15. We then calculate the average daily return for each decile for the period January 2, 1970 (the first trading day of 1970) to December 31, 2003, yielding a time series of 8,586 daily returns for each decile (D1,…,D10). The *AQ* factor-mimicking portfolio, *AQfactor*, equals the difference between the daily returns of the poorest *AQ* deciles (deciles 7-10) and the best *AQ* deciles (deciles 1-4). This procedure (similar to that used by Carhart (1997) to construct a factor mimicking portfolio for price momentum) yields a series of 8,586 daily *AQfactor* returns ($AQfactor_t$).[9]

[8] Calculating the *AQ* measure in year T using the residuals in years T-5 to T-1 accounts for the fact that equation (1) contains a lead term, $CFO_{j,T+1}$. In total, calculation of *AQ* requires seven years of data because we require five residuals and the model contains two lag terms.

[9] Carhart's procedure differs from Fama and French's (1993) procedures used to create size and book-to-market mimicking factors in two respects: he uses equally-weighting rather than value-weighting and does not orthogonalize with respect to firm size. As sensitivity tests, we re-create an *AQfactor* that is value-weighted within each cell and orthogonalized with respect to size. (Both value-weighting and orthogonalization serve to remove any size-shared variation from the estimation of the e-loadings.) Results based on the value-weighted, size-orthogonalized *AQfactor* (not reported) are similar in all respects to those documented.

For our sample, the average value of $AQfactor_t$ is 0.0779%, or about 22% on an annualized basis; the standard error is 0.0085%. To put these figures in perspective, over the same time period the average daily excess market return is 0.0216% (standard error of 0.0102%), the average value of the daily *SMB* is 0.0021% (standard error of 0.0056%), and the average value of the daily *HML* is 0.0224% (standard error of 0.0052%). *AQfactor* is reliably different from zero (t-statistic = 9.17, or the mean value of 0.0779% divided by its standard error of 0.0085%), and has a higher t-statistic than the other risk factors (the t-statistics for market risk, *SMB* and *HML* are 2.12, 0.37 and 4.28, respectively).

Because *AQfactor* is time-specific, not firm-specific, we can correlate *AQfactor* with the returns of any firm to determine that firm's exposure to poor earnings quality, much like we correlate a firm's returns with the market risk premium to obtain a measure of its exposure to market risk. The specific correlation measure we use is the coefficient estimate on *AQfactor* obtained from 1-factor (superscript *1f*) and 3-factor (superscript *3f*) asset pricing regressions which include *AQfactor* as an independent variable:

1-factor: $$R_{j,t} - R_{F,t} = \alpha^{1f}_{j,T} + \beta^{1f}_{j,T}(R_{M,t} - R_{F,t}) + e^{1f}_{j,T} AQfactor_t + \varepsilon^{1f}_{j,t} \quad (2)$$

3-factor: $$R_{j,t} - R_{F,t} = \alpha^{3f}_{j,T} + \beta^{3f}_{j,T}(R_{M,t} - R_{F,t}) + s^{3f}_{j,T} SMB_t + h^{3f}_{j,T} HML_t + e^{3f}_{j,T} AQfactor_t + \varepsilon^{3f}_{j,t} \quad (3)$$

where *t* = index for the number of trading days in year T; $R_{j,t} =$ firm j's return on day *t*; $R_{F,t} =$ the risk free rate on day *t*; $R_{M,t} =$ the market return on day *t*; $SMB_t =$ small-minus-big factor on day *t*; $HML_t =$ high-minus-low book-to-market factor on day *t*. $R_{M,t}$, SMB_t and HML_t are obtained from the WRDS database (where raw returns are from CRSP and *SMB* and *HML* factors are from Ken French).

For the 1-factor (3-factor) specification, $e^{1f}_{j,T}$ ($e^{3f}_{j,T}$) is the estimate of firm j's sensitivity to poor earnings quality in year T. The other slope coefficients, $\beta^{1f}_{j,T}$ (or $\beta^{3f}_{j,T}$), $s^{3f}_{j,T}$ and $h^{3f}_{j,T}$, capture the firm's exposure to returns representations of market risk, size, and book-to-market, respectively, in year T.

In summary, we create an accounting-based measure of accruals quality $AQ_{j,T}$ using a restricted sample of firms with seven years of accounting data, convert *AQ* to a time-specific returns representation ($AQfactor_t$), and use this time-specific returns representation in firm-and year-specific regressions to obtain a firm-year returns-based representation of earnings quality ($e^{\#}_{j,T}$, $\# = 1f$ or $3f$). Relative to the original accounting-based measure, e-loadings can be calculated for much larger samples because they require only enough daily returns in year T to estimate (2) or (3). An additional benefit of the returns approach is that because $AQfactor_t$ varies through time, e-loadings are not constrained to be slow to change, as is *AQ* which requires seven years of accounting data and therefore has a mechanical dependence year over year.

4. Sample and Descriptive Data

We begin by identifying all firms with the necessary data to estimate *AQ* in each year T=1970-2003 (the AQ Sample).[10] Using the *AQ* measures, we calculate *AQfactor* using the procedures described in section 3. Next, we identify all firms with at least 100 daily returns in year T (the Returns Sample). The requirement of 100 daily returns in year T to estimate (2) and (3) is ad hoc, and we assess its sensitivity in section 6. Table 1, panel A shows the number and percentage of sample firms, by year, for the two samples. As expected given the sample selection criteria, the Returns Sample is larger than the AQ Sample, both in number of firms (mean of 6,408 firms per year versus 2,147 firms per year) and as a percentage of traded firms (mean of 92.4% per year versus 30.9%). To further contrast the representativeness of the two samples, panel B reports information about the distributions of size (market capitalization, total assets, and total sales) and performance (as measured by return on assets and return on equity). Comparisons show that the AQ Sample contains larger firms (mean market capitalization is \$1,295 million versus \$942 million; mean assets are \$1,404 million versus \$1,005 million; mean sales are \$1,298 million versus \$908 million) and more successful firms (mean ROA is 3.7% versus 2.6%; mean ROE is 7.3% versus 5.5%). These data demonstrate that the Returns Sample dominates the AQ Sample on both sample size and survivorship bias. Given that the Returns Sample is minimally restricted (i.e., it requires only 100 daily returns in year T), it is more representative of the population than is the AQ Sample, which requires a firm to have seven years of accounting data for inclusion in year T.

[10] Because of the data requirements for *AQ*, as well as the additional three-month lag we require for accounting data to reach the market, we use accounting data from as early as 1963.

We next estimate firm-year e-loadings for the Returns Sample, using equations (2) and (3); we discuss but do not tabulate these results. For e-loadings based on the CAPM, the mean value of e^{1f} is 0.3021 with a standard deviation of 0.5335, and the inter-quartile range is -0.0089 (25th percentile) to 0.5249 (75th percentile). For e-loadings based on the 3-factor model, the mean value of e^{3f} is 0.1154 with a standard deviation of 0.5860, and the inter-quartile range is -0.1694 (25th percentile) to 0.3126 (75th percentile). The other loadings in these models (i.e., β for the CAPM and β, *s* and *h* for the 3-factor model) exhibit similar variation. The mean and median explained variation of the augmented 1-factor and 3-factor models is between 3% and 9%, less than the 16-20% reported by Francis et al. (2005) for estimations requiring a minimum of 18 monthly observations per firm. The lower explained variability is expected because we estimate our models using daily returns which are noisier than monthly returns.[11]

5. Analysis of e-Loadings

Our four analyses of the reliability of e-loadings are complementary, in that they use different research designs and distinct indicators of earnings quality. Section 5.1 uses a cross-sectional design to analyze the associations between e-loadings and both determinants of accruals quality and measures of earnings quality that have been used in other research settings. Section 5.2 also uses a cross-sectional design, but the focus of the analysis is on earnings response coefficients and properties of analyst forecasts. Section 5.3 shifts to a time-series design, and considers the over-time relation between e-loadings and firm age as a summary indicator of available firm-specific information. Finally, section 5.4 analyzes the relation between e-loadings and events associated with poor earnings quality.

5.1. Associations between e-loadings and determinants of and proxies for earnings quality

Our first analysis considers whether accounting-based determinants of earnings quality explain cross-sectional variation in e-loadings. We focus on the five innate determinants proposed by Dechow

---

[11] Note that these R^2s are based on firm-specific estimations; calendar-time portfolio regressions generally show much higher explanatory power (often above 40%) because firm-specific variation is largely eliminated in the portfolio design.

and Dichev: firm size (*log(assets)*, measured as the log of total assets; results are not sensitive to other measures of size, such as revenues), standard deviation of cash flow from operations ($\sigma(CFO)$), standard deviation of sales revenues ($\sigma(Sales)$), length of operating cycle ($OperCycle$, measured as the log of the sum of days accounts receivable and days inventory) and negative earnings incidence ($NegEarn$ = 1 if year T earnings are negative, 0 otherwise). We measure *log(assets)*, *log(OperCycle)* and *NegEarn* as of the end of year T and we calculate $\sigma(CFO)$ and $\sigma(Sales)$ using data from year T-6 to T. Because of the latter calculation, the sample used for these tests is the AQ Sample not the Returns Sample.

Table 2, panel A shows the results of regressing the year T e-loadings on the innate determinants. We report the average values of the coefficient estimates obtained from 34 yearly regressions; t-statistics are based on the standard errors of the 34 annual coefficients (Fama and MacBeth, 1973). The results indicate that e-loadings are highly correlated with each of the innate determinants, in the directions predicted by Dechow and Dichev: e-loadings are negatively correlated with firm size, and positively correlated with the variability of cash flows and sales, the length of the operating cycle, and the incidence of losses (t-statistics range in absolute value from 8.70 to 16.10).

Our examination of the relation between e-loadings and other measures of earnings quality focuses on the seven earnings attributes considered by Francis et al. (2004): accruals quality itself (*AQ*), earnings persistence (*Persistence*, measured as the negative of the AR1 parameter from firm-specific regressions of current earnings per share on lagged earnings per share), earnings predictability (*Predictability*, measured as the standard deviation of the error term from firm-specific AR1 models of earnings), smoothness (*Smoothness*, measured as the ratio of the standard deviation of earnings to the standard deviation of cash flows), value relevance (*Value Relevance*, measured as the negative of the explained variability of a regression of annual returns on the level and change in earnings per share), timeliness (*Timeliness*, measured as the negative of the explained variability from a Basu (1997) reverse regression of earnings on returns controlling for the sign of those returns), and conservatism (*Conservatism*, measured as the negative of the coefficient on negative returns from the aforementioned reverse regression). We follow Francis et al.'s (2004) procedures and estimate each attribute over rolling 10-year windows. Associations are calculated between these measures of earnings quality and the e-loadings averaged across the same rolling 10 year windows. Each measure is ordered consistently such that higher (lower) values indicate poorer (better) earnings quality.

Table 2, panel B shows the pairwise correlations between the e-loadings and the earnings quality proxies. To the extent that the e-loadings are correlated with the constructs captured by each of these attributes, we expect to observe positive correlations between e-loadings and each variable. For example, the correlation between e-loadings and *AQ* provides evidence on whether e-loadings capture the variable that underlies the construction of the *AQfactor*. We expect and find that this correlation is positive, with magnitudes ranging from 0.4397 to 0.5008, all significant at the 0.0001 level.[12] Francis et al.'s (2004) results suggest that associations between the e-loadings and the earnings quality attributes should be strong for the three attributes based only on accounting numbers, and weak for the three attributes that are based also on returns. Our results are for the most part consistent with their findings. Both *Persistence* and *Smoothness* exhibit significant (at the 0.0001 level) positive correlations with e-loadings ranging in magnitude from 0.1653 to 0.2575; for *Predictability*, the correlations are lower (0.0102 to 0.0760; significant at the 0.02 level or better). Results for measures of earnings quality that also rely on returns data (*Value Relevance*, *Timeliness* and *Conservatism*) are generally low, with correlations ranging from 0.0007 (insignificant) to 0.0416 ($p<0.0001$).

5.2. <u>Market participant behaviors as a function of e-loadings</u>

In this section, we examine whether e-loadings are associated with predictable variation in investor and analyst behaviors. Our first analysis builds on prior research (e.g., Imhoff and Lobo, 1992) that posits information uncertainty as a determinant of investors' response to earnings as captured by the coefficient relating returns to earnings (earnings response coefficient, or ERC). Our test of whether firms with higher e-loadings have smaller ERCs is based on the following regression:

[12] This result is similar to Fama and French's (1993) evidence that loadings on the *SMB* and *HML* mimicking factors are positively correlated with their underlying variables, market capitalization and book-to-market, respectively.

$$CAR(-1{,}0)_{j,t} = \lambda_0 + \lambda_1 UE_{j,t} + \lambda_2 UE_{j,t} \times e^{\#}_{j,T} + \zeta_{j,t} \quad (4)$$

where $CAR(-1,0)_{j,t}$ = firm j's two-day cumulative abnormal return over the quarterly earnings announcement, where abnormal return is defined as the raw return less the value-weighted market return; and $UE_{j,t}$ = unexpected earnings conveyed by firm j's quarterly earnings announcement made on day t, equal to firm j's reported earnings for quarter q less the consensus analyst forecast, scaled by firm j's share price twenty days before the earnings announcement date.

Results of estimating equation (4) are shown in Table 3, Panel A. For these tests, unexpected earnings are calculated using analyst forecast data from Zacks, 1983-2002; on average, there are 2,885 firms per quarter. We report mean values of the coefficients obtained from the 80 quarterly estimations of (4) over the period 1983-2002; the t-statistics are based on the time-series of the standard errors of the 80 coefficient estimates.[13] The results show that earnings news announced by higher e-loading firms is associated with a significantly weaker market response than is earnings news announced by lower e-loading firms: λ_2 = -0.0415 (t-statistic = -3.96) using e^{1f}, and λ_2 = -0.0370 (t-statistic = -4.32) using e^{3f}. We also estimate quarterly regressions which include *UE* interacted with other variables known to affect earnings response coefficients: whether the firm reported a loss in quarter q (*NegEarn* = 1 if reported earnings in quarter q are negative, 0 otherwise), firm size (*ln(Size)* = log of firm j's sales revenues), firm j's market-to-book ratio (*MB*), and the ratio of firm j's debt to equity (*Leverage*):

$$\begin{aligned} CAR(-1{,}0)_{j,t} &= \lambda_0 + \lambda_1 UE_{j,t} + \lambda_2 UE_{j,t} \times e^{\#}_{j,T} + \lambda_3 UE_{j,t} \times NegEarn_{j,q} + \lambda_4 UE_{j,t} \times \ln(Size)_{j,q} \\ &\quad + \lambda_5 UE_{j,t} \times MB_{j,q} + \lambda_6 UE_{j,t} \times Leverage_{j,q} + \xi_{j,t} \end{aligned} \quad (5)$$

Results of estimating (5), reported in Panel B, continue to show that firms with larger e-loadings have smaller responses to earnings news (t-statistics for λ_2 are -2.16 using e^{1f} and -3.14 using e^{3f}).

Our second analysis of the relation between e-loadings and market participant behavior examines the dispersion and accuracy of analysts' earnings forecasts. Based on prior research, we predict that firms with higher e-loadings have more difficult-to-predict earnings, resulting in both more dispersed and less accurate forecasts relative to lower e-loading firms. More specifically, prior research reports positive

[13] An alternate approach is to estimate (4) as a pooled regression, with t-statistics based on Newey-West (1987) adjusted standard errors to control for heteroscedasticity and autocorrelation. Inferences (not tabulated) are the same: in particular, the t-statistic for λ_2 is -7.17 (-7.61) for e^{1f} (e^{3f}).

relations between random walk measures of earnings surprise—which we interpret as indicators of difficult-to-predict earnings—and both forecast dispersion and forecast accuracy.[14]

Following prior literature, we measure forecast dispersion as the standard deviation of analysts' EPS forecasts for quarter q, scaled by share price at the beginning of the quarter. Tests of dispersion are conducted on all firms with at least three earnings forecasts for quarter q issued in the three months preceding the announcement of quarter q earnings; the three or more forecasts produce a standard deviation of the forecasts made in quarter q for firm j, $Dispersion_{j,q}$. We measure forecast accuracy as the absolute forecast error (the difference between reported and forecasted EPS for quarter q, scaled by share price at the beginning of the quarter). Tests of forecast accuracy are conducted on all firms with at least one quarterly earnings forecast for quarter q in the three-month period preceding the quarter q earnings announcement; our measure of forecast accuracy, $|FE|_{j,q}$, is the average of the absolute forecast errors across all forecasts made about firm j for quarter q. Larger values of $Dispersion_{j,q}$ and $|FE|_{j,q}$ indicate more dispersed and less accurate forecasts, respectively. Similar to Table 3, these tests are conducted using analyst forecast data from Zacks, 1983-2002.

Our tests of whether firms with greater forecast dispersion and larger absolute forecast errors have higher e-loadings are based on the following regressions:

$$Dispersion_{j,q} = \kappa_0 + \kappa_1 e^{\#}_{j,T} + \kappa_2 Age_{j,q} + \kappa_3 ln(Sales)_{j,q} + \varsigma_{j,q} \tag{6}$$

$$|FE|_{j,q} = \delta_0 + \delta_1 e^{\#}_{j,T} + \delta_2 Age_{j,q} + \delta_3 ln(Sales)_{j,q} + \psi_{j,q} \tag{7}$$

Equations (6) and (7) control for forecast age and firm size, both of which have been shown to affect forecast dispersion and accuracy (Bowen et al., 2002). Forecast age ($Age_{j,q}$) is the average number of

[14] Lang and Lundholm (1996) and Bowen, Davis and Matsumoto (2002) measure earnings surprise as the absolute value of the difference between reported earnings for quarter q and reported earnings for quarter q-4, scaled by share price ($\left|UE^{SRW}_{j,q}\right|$). In addition to $\left|UE^{SRW}_{j,q}\right|$, Heflin, Subramanyam and Zhang (2003) include two indicator variables to capture earnings surprise: the first equals one if the sign of the unexpected earnings news in the quarter q earnings announcement is negative ($NegUE^{SRW}_{j,q} = 1$ if $UE^{SRW}_{j,q} < 0$), the second equals one if reported earnings for quarter q are negative, and zero otherwise ($Loss_{j,q}$).

days that the forecasts issued in quarter q precede the earnings announcement date. Firm size is proxied by the log of sales in quarter q, $\ln(Sales)_{j,q}$. We estimate equations (6) and (7) using both pooled and quarterly estimations; results are similar for the two approaches, so we tabulate only the quarterly results. Table 4 reports the mean values of the coefficients from the 80 quarterly regressions, with t-statistics based on the standard errors of those 80 estimates. The results for both dispersion (Panel A) and absolute forecast errors (Panel B) indicate that e-loadings are positively associated with both forecast properties (t-statistics range between 5.39 and 12.07). We also re-estimate equations (6) and (7) including the three measures of earnings surprise considered in prior research (described in footnote 14). Even controlling for these other surprise measures, the results show that e-loadings remain significant (t-statistics, reported in Table 4, range from 2.92 to 8.40) in explaining properties of analysts' forecasts.

5.3. Patterns in e-loadings associated with firm age

This section examines patterns in e-loadings associated with firm age, which we view as an inverse measure of investor perceptions of the amount and stability of firm-specific information. Our examination is predicated on the view that the relatively sparse firm-specific information that is available for younger firms induces greater uncertainty about that firm's business, including uncertainty about its financial reporting quality. We conduct two investigations involving firm age and e-loadings. The first involves newly listed firms: for such firms, we expect information uncertainty is highest at initial listing, and declines as more information becomes available as the firm matures. This hypothesis follows from Lang's (1991) argument (in an ERC setting and in the context of IPO firms) that, as more earnings observations are revealed, investors' uncertainty about the persistence parameter underlying the earnings process decreases.[15] Similar to Pastor and Veronesi (2003) and Fama and French (2005), we measure firm age as the number of years since listing (*FirmAge*), where listing year (year 0) is the first year for which stock returns are available on CRSP. We consider age, rather than measures such as firm size,

[15] Regulation S-X requires IPO firms to provide three years of income and cash flows in their Form S-1 registration statements. Firms in existence less than three years must provide information for the period they have been in existence. Thus, for most newly listed firms, financial data exist which provide a basis for investors to develop views about the firm's financial reporting quality.

institutional holdings and analyst coverage (which have also been used to reflect the relative uncertainty or quality of information) so that we can use an over-time design to examine whether e-loadings decline after initial stock market listing.

Figure 1 shows the trend in mean e-loadings as a function of *FirmAge* for firms in the Returns Sample that listed in any year, 1926-2003. To be included in Figure 1, the firm must have data in event year T=0. Therefore, the e-loading for year T=+1 is based on all firms who survived at least one year after listing. There is a clear downward trend in e-loadings as a function of *FirmAge* for both $e^{1f}_{j,T}$ and $e^{3f}_{j,T}$. To test the significance of this trend, we regress the average e-loading of age τ (τ = 0,1,2 … 39) on *FirmAge*.[16] The results, also plotted in Figure 1, indicate a significant downward trend in e-loadings as a function of *FirmAge* (t-statistics, not reported, are -19.33 for $e^{1f}_{j,T}$ and -17.78 for $e^{3f}_{j,T}$). To control for over-time changes in the sample composition, we repeat our tests restricting the sample to firms that survived at least 10 years following listing; results (not tabulated) show a significant decline in e-loadings for this sample. Overall, we interpret declines in e-loadings as a function of firm age as evidence that, as firm-specific information increases, investor uncertainty about reporting quality decreases.

Our second analysis investigates the prediction that e-loadings exhibit greater year-over-year stability for relatively mature firms than they do for relative young firms. In examining stability, we do not suggest that a firm's e-loadings *must* be the same from one year to the next; indeed, a benefit of e-loadings is that their estimation procedures do not force them to be constant over short intervals. As a practical matter, however, we would not expect to observe significant year-to-year changes in a firm's earnings quality (or investors' perceptions of that earnings quality) for firms with relatively stable reporting and information environments because, for such firms, relatively less new information about earnings quality is likely to be revealed each year. Our measure of the stability of e-loadings is the autocorrelation (AR1) parameter estimated for each cross-section of firms of *FirmAge* = 0,…,39+ ($\omega_1^{\#}$):

[16] We truncate observations at *FirmAge* =39 because the sample size is small for values of *FirmAge* > 39; results are not sensitive to this truncation.

$$e^{\#}_{j,T} = \omega^{\#}_0 + \omega^{\#}_1 e^{\#}_{j,T-1} + \varsigma^{\#}_{j,T} \quad (8)$$

Figure 2 plots the estimates of $\omega^{\#}_1$ (representing each *FirmAge* cross-section) as a function of *FirmAge*; we also plot the results of regressing $\omega^{\#}_1$ on *FirmAge*. There is a pronounced increase in the stability of e-loadings: ω^{1f}_1 and ω^{3f}_1 increase from 0.22 and 0.08, respectively, for *FirmAge*=0, to 0.46 and 0.40 for *FirmAge*=39+ (t-statistics for these increases, not reported, are 9.86 and 7.98).

In summary, both the level of e-loadings and their stability exhibit systematic patterns related to firm age: (1) the level of e-loadings is highest at initial listing and declines as the firm matures; and (2) the year-over-year autocorrelation in e-loadings increases with firm age.

5.4. Patterns of e-loadings associated with restatements, lawsuits and bankruptcies

Our final analyses investigate whether e-loadings are higher in three situations characterized by objective evidence of poor financial reporting quality: financial statement restatements, shareholder lawsuits, and bankruptcies. These analyses are carried out in event time; so, for example, year 0 for the Restatement Sample (Lawsuit Sample) [Bankruptcy Sample] is the year the restatement was announced (the lawsuit was filed) [the firm delisted due to bankruptcy]. Using an event-time analysis allows us to control for other (unrelated) influences on over-time patterns in e-loadings by comparing e-loadings of the event firms with e-loadings of non-event firms matched on calendar time. The event-time analysis also allows us to exploit the fact that each year there are many more firms who do not restate, are not sued, and do not go bankrupt than there are firms that experience any of these events. Specifically, for each of the Event Samples, we select 100 equal-size random samples of non-event firms, with a yearly distribution that is identical to that of the Event Sample. To the extent the smaller number of Event Sample observations is associated with greater measurement error in e-loadings than are the e-loadings of the larger set of non-event firms, equating the size of the Event and Non-Event Samples minimizes measurement error differences between the estimates derived from the two samples. Separately, equating the yearly distributions controls for calendar time influences affecting e-loadings. The 100 non-event samples create an empirical distribution of *non-event* e-loadings. A comparison of *event* e-loadings to this empirical distribution provides evidence on whether the e-loadings of the event firms are statistically distinguishable from those of firms that did not experience events associated with poor earnings quality.[17]

Our sampling and test procedures are the same for each of the three events, so we illustrate them only for the Restatement Sample. The Restatement Sample (described in section 5.4.1) contains 788 firm-year restatements occurring between January 1, 1997 and June 30, 2002 with returns data. The yearly distribution of this sample is as follows: 1997 (n=82), 1998 (n=87), 1999 (n=151), 2000 (n=162), 2001 (n=199), and the first half of 2002 (n=107). We select 100 random samples of n=788 total non-restating observations with the same yearly distribution (for example, 82 non-restating firms are randomly selected from 1997). Matching by year controls for events and conditions affecting all firms in a given year and provides the pseudo-event date (year 0) for the non-event firms. For each of the 788 observations in the Restatement Sample, we estimate equations (2) and (3) for each year, T=-5,…,+1, where T=0 is the restatement year. We then calculate the mean value of the e-loadings across restatement firms, in event time, to obtain the mean value of $e^{1f}_{j,T}$ ($e^{3f}_{j,T}$) for each year T=-5,…,+1. For each of the 788 observations in each of the 100 Non-Restatement Samples, we perform the same estimation and averaging, yielding a series of 100 mean e-loadings for non-restating firms for each year, T=-5,…,+1.

Our analyses of the Event Samples are intended to provide evidence on two issues: first, whether e-loadings increase for event firms relative to the non-event benchmark samples in the years preceding events that provide objective evidence of poor earnings quality ("pre-event tests"); and second, whether e-loadings decline in post-event periods for firms in the Restatement and Lawsuit Samples ("post-event tests"). (Post-event declines in e-loadings cannot be investigated for the Bankruptcy Sample because, by definition, these firms ceased trading, so returns data are not available to construct e-loadings in post-bankruptcy periods.) The pre-event and post-event tests are predicated on the assumptions that (1)

[17] Empirical distributions of comparison samples have been used in other contexts to establish a benchmark for assessing whether a test sample is distinctive. For example, Alford, Jones, Leftwich and Zmijewski (1993) use this approach to compare the value relevance and timeliness of earnings in the US versus sixteen non-US jurisdictions. They compare measures of value relevance and timeliness for each of 16 non-US test samples with an empirical distribution of 100 year-, size-and industry-matched US samples.

earnings quality deteriorated in the period preceding each of the three events and was, to some extent, corrected in the year after the event, and (2) investors were able to discern at least some portion of this deterioration. Our pre-event tests are clearly more sensitive to the second assumption than are the post-event tests, because the events we examine were all publicly disclosed.

In terms of the first assumption, prior research provides mixed evidence on the association between earnings quality and events such as restatements, lawsuits and bankruptcies. For the most part, this research examines changes in earnings quality metrics before and after restatements. For example, Anderson and Yohn (2003) compare bid-ask spreads and earnings response coefficients before and after restatements and find no change in spreads and a decline in ERCs. Moore and Pfeiffer (2004) find no significant difference in reporting aggressiveness (measured by abnormal accruals) before versus after a restatement. Taken together, these results suggest that correcting a GAAP violation by restating has either no effect, or a worsening effect, on earnings quality. These findings are at odds with intuition: assuming that GAAP earnings are of high quality, one would expect that when a GAAP violation is cured, the quality of the firm's financial reports improves not worsens.

In terms of the second assumption, prior research suggests that accounting data have predictive ability for the three events we study. Richardson, Tuna and Wu (2002) find that restatement firms have large accruals in the years of alleged manipulation, and that restatement firms try to maintain earnings patterns (such as consecutive positive earnings surprises and earnings growth). Research on lawsuit firms shows that firms sued over allegedly defective financial disclosures have systematic patterns in accounting variables, including higher levels of accounts receivable and inventory (Stice, 1991), more income-increasing total accruals (Lys and Watts, 1994), and more income-increasing abnormal accruals (Heninger, 2001). Finally, numerous studies document systematic patterns in accounting data for bankrupt firms, including declines in profitability and increases in measures of financial distress (such as increases in book-to-market ratios and bankruptcy indices, Altman, 1968;1993). Taken as a whole, we believe previous research supports the assumption that market participants have at least some information to form expectations about restatements, lawsuits and bankruptcies before these events occur.

We analyze several predictions about the e-loadings of the Event and Non-Event Samples. First, we expect each Event Sample has larger e-loadings than its comparison Non-Event Samples in the event year, and in one or more years leading up to the event. Second, we expect that e-loadings of each Event Sample increase over time, relative to the e-loadings of comparison Non-Event Samples. For the Restatement and Litigation Samples we expect that e-loadings decline in year +1, consistent with year 0 actions correcting the financial reporting problems, e.g., restatement firms stop using incorrect GAAP and issue corrected financial statements. As noted earlier, we do not test whether e-loadings decline in year +1 for the Bankruptcy Sample because these firms cease trading after year 0.

5.4.1. *Restatement sample and tests*

Data on restatements come from the General Accounting Office's (GAO) study on Financial Statement Restatements, which covers January 1, 1997 through June 30, 2002. The GAO database identifies 919 restatement firm-years, representing 840 distinct firms. From this set, we identify all observations with usable identifier codes, resulting in a sample of 882 firm-year observations, for 794 firms. The distribution of this sample is shown in panel A, Table 5. Imposing the requirement that firms have at least 100 daily returns in year T reduces the final Restatement Sample to 788 firm-year observations, representing 713 firms. Comparison of the yearly distributions of the two samples shown in panel A indicates that the year-by-year distribution of this Restatement Sample is similar to the sample unconstrained by the requirement to have sufficient daily returns data.

Table 6, panel A shows the e-loadings for the Restatement Sample and Non-Event Samples for event years -5 through +1. The e-loading for the Non-Event sample is the mean value of $e^{\#}_{j,T}$ calculated for the 100 matched samples. The column labeled "Diff." shows the difference between the mean value of $e^{\#}_{j,T}$ between the Event and Non-Event Samples. We evaluate the statistical significance of this difference in two ways. The first measure is a conventional t-statistic, based on the standard error of the mean e-loading of the Event Sample and the mean standard error of the 100 Non-Event Samples. The second is nonparametric and is reported in the column labeled "%-tile": here we show the percentile rank of the Restatement Sample e-loading relative to the distribution of the 100 mean e-loading values of the Non-Event Samples. For example, a %-tile value of 25 implies that the Restatement Sample e-loading ranks 25th from the bottom of the 100 mean Non-Event mean e-loadings.

Panel A of Table 6 shows that, as predicted, Restatement firms' e-loadings increase over the period preceding the restatement announcement year: the mean value of e^{1f} increases from 0.2994 in year -5 to 0.3899 in year 0, and declines to 0.3368 in year +1. For e^{3f} we observe a monotonically increasing trend, from 0.1018 to 0.3153 between years -5 to 0; the e-loading in year +1 declines to 0.2694. Over the same pre-event period, the results for e^{1f} for Non-Event firms show a slight decline, while results for e^{3f} show an increase up to year 0 (from 0.1292 in year -5 to 0.2072 in year 0), followed by a decline to 0.1963 in year +1. The trend in e^{3f} for the Non-Event Sample indicates a calendar-time effect which we control for by concentrating on the *differences* between the Event and Non-Event Samples.

Differences between e-loadings of the Event and Non-Event Samples are more pronounced over the entire period between years -5 and +1 for e^{1f} than for e^{3f}. For example, the value of e^{1f} for the Restatement sample is at or above the 91st percentile of the Non-Event samples in every year considered, and parametric tests yield t-statistics well above 2.0 in years -2 through +1. In contrast, values of e^{3f} for the Restatement Sample are below the median of the Non-Event Samples in years -5 through -2 and t-statistics exceed 2.0 only in years 0 and +1. Clearly, however, differences are statistically reliable for both e^{1f} and e^{3f} in years 0 and +1, with the largest differences observed in year 0. The significant difference in year +1 suggests that market participants continue to perceive Restatement firms as having poorer earnings quality in the year after the restatement occurs.

5.4.2. *Lawsuit sample and tests*

Data on class action securities litigation are from Woodruff-Sawyer & Co. and from the Securities Class Action Clearinghouse. In total, there are 2,707 firm-year lawsuit observations over the period January 1, 1990 through April 24, 2003. Because we are interested in lawsuits over financial reporting issues, we focus on the subset of 1,067 lawsuits alleging accounting violations or concerns. We find usable identifier data for 857 lawsuits; the requirement of 100 daily returns in year T reduces the final Lawsuit Sample to 793 lawsuits against 751 distinct firms. Table 5, panel B shows the distribution of the Lawsuit observations across the sample years, before and after imposing the returns requirement; the year-by-year distributions of the two samples are similar.

Examination of the e-loadings of the Lawsuit Sample (Table 6, Panel B) suggests an increase in e^{1f} from 0.2756 to 0.4792 for lawsuit firms over years -5 to year 0, and a (slight) decline in year +1. For the matched Non-Event firms, there is a near-monotonic (but limited) decline in e^{1f} over years -5 to +1. The difference in e^{1f} between the Event and Non-Event firms shows a pronounced upward trend, from -0.0215 (year +5) to 0.2070 (year 0), further increasing to 0.2217 in year +1. The 3-factor e-loadings increase for both Lawsuit firms and Non-Event firms over the period -5 to 0; however, the increase is much more pronounced for Lawsuit firms. Comparing the two samples, the Lawsuit firms exhibit lower values of e^{3f} than Non-Event firms until year -1, and Lawsuit firms' e^{3f} are significantly larger than Non-Event firms e^{3f} in years 0 and +1. Non-parametric tests place the Lawsuit firms' 1-factor (3-factor) e-loadings at the 100$^{\text{th}}$ percentile of the Non-Event samples in years -2 through +1 (years 0 and +1). The weight of the evidence supports the view that e-loadings do not differ between Lawsuit and Non-Event firms in most pre-event years, but differ dramatically in years 0 and +1 when Lawsuit firms' e-loadings exceed those of Non-Event firms.

5.4.3. *Bankruptcy sample and tests*

We identify 618 firms that went bankrupt during 1970-2003 using delisting codes obtained from the CRSP database. Requiring returns data reduces the sample to 371 bankruptcy observations.[18] The yearly distributions of both samples are similar (panel C, Table 5). Results comparing the e-loadings of Bankrupt firms with Non-Event firms are shown in panel C of Table 6. We observe an upward trend in both e^{1f} and e^{3f} for Bankrupt firms over the event years, with the highest e-loadings found in the year

[18] The decline in sample size is due to the fact that requiring at least 100 daily returns in calendar year T effectively eliminates firms that went bankrupt in the first five months of year T.

before and the year of bankruptcy. Non-Event firms show a downward trend in e^{1f} over this same time frame, and a modest upward trend in e^{3f}. The difference in e^{1f} between Event and Non-Event firms increases monotonically from 0.1558 in year -5 to 0.3558 in year 0 while the difference in e^{3f} also increases monotonically, from 0.1293 in year -5 to 0.3888 in year 0. Both the parametric t-test and the non-parametric test indicate that the average e-loading is substantially higher for the Bankruptcy Sample than for the Non-Event Samples in every year examined.

5.4.4. *Summary of results of event tests*

The weight of the evidence for the three Event Samples supports the following inferences. First, e-loadings are significantly larger for Event firms than for Non-Event firms in the event year (for all Event Samples) and in the year following (for the Lawsuit and Restatement Samples). Second, for all Event Samples there is an increase in the mean e-loading of the Event firms over years -5 to 0. For the Restatement Sample, we also observe the predicted decline in e-loading between year 0 and year +1; however, this pattern is not systematically observed for the Lawsuit Sample. Third, the difference between the e-loadings of the Event firms and the Non-Event firms tends to increase over years -5 to year 0 for all Event Samples. This pattern, which controls for changes in e-loadings due to economy-wide factors, indicates that investors perceive declining earnings quality for Event firms over the period leading up to and culminating in the event year. This pattern is most pronounced, and begins earliest, for the Bankruptcy Sample where, arguably, investors have access to relatively more pre-event information.[19]

As a check on our treatment of the events as independent, we calculated the overlap in firms represented in the three Event Samples. The Bankruptcy Sample displays minimal overlap with either the Restatement Sample (4.3%) or the Lawsuit Sample (7.8%). For the Restatement and Lawsuit Samples, the overlap is 25-27%. We address this sample dependence by repeating the tests in Table 6 excluding all

---

[19] Note that the results for e^{3f} control for changes in other risk factors over the event periods. The importance of this is best demonstrated by the bankruptcy sample, where firms' risk profiles change as they approach bankruptcy. Consistent with Fama and French's (1995) finding that *h*-loadings proxy for relative distress, we find that *h*-loadings increase for firms approaching bankruptcy.

firms that appear in more than one Event Sample. The results (not reported) are similar to those for the full sample, in terms of both the magnitudes of the coefficients and their statistical significance.

6. Extensions

We conduct several extensions of our main analyses. First, to examine the possibility that our results are driven by the subset of firms with data on *AQ* (i.e., the AQ Sample), we repeat our tests on the subset of firms in the Returns Sample but not in the AQ Sample. Results (not tabulated) are similar in all respects to those reported. This finding demonstrates that the significant expansion of sample size made possible by not requiring data on *AQ* produces e-loadings that are as reliable as e-loadings generated for the AQ Sample itself.

Second, we consider how well e-loadings for other earnings quality proxies perform, relative to e-loadings based on *AQfactor*. Our aim is to assess whether *any* returns representation of earnings quality performs equally well along the dimensions we examine. We calculate mimicking factors for the six earnings quality proxies considered in section 5.1, yielding a returns representation for each: $Factor(k)$, *k*= *Persistence*, *Predictability*, *Smoothness*, *Value Relevance*, *Timeliness* and *Conservatism*. We substitute $Factor(k)$ for *AQfactor* in equations (2) and (3), generate firm-specific e-loadings for each of the *k* factors ($e(k)^{\#}_{j,T}$), and then repeat our tests *k* times substituting $e(k)^{\#}_{j,T}$ for $e^{\#}_{j,T}$. The results (not tabulated) show that e-loadings based on other measures of earnings quality generally show similar patterns as e-loadings based on *AQfactor*, except that e-loadings for *Timeliness* and *Conservatism* show weaker effects. Of the remaining e-loadings, those for *Smoothness* and *Persistence* are most similar, in magnitude and statistical significance, to those documented for *AQfactor*. (This result is consistent, in hindsight, with *Smoothness* and *Persistence* being closer to accruals quality in capturing the precision of earnings with respect to cash flows). There is no evidence that any of the *k* e-loadings systematically dominates the *AQ*-based e-loading ($e^{\#}_{j,T}$).

Third, we consider an abnormal accruals metric, equal to the absolute value of the performance-matched abnormal accrual ($\left|AA_{j,T}\right|$) estimated in annual cross-sections for each of the 48 Fama-French industries containing at least 20 firms. We calculate a mimicking factor based on $\left|AA_{j,T}\right|$, $Factor(|AA|)$, and repeat all tests on the resulting e-loadings based on this factor, $e(|AA|)^{\#}_{j,T}$. The results show that while $e(|AA|)^{\#}_{j,T}$ shows similar patterns as $e^{\#}_{j,T}$, the results are generally weaker for this measure. Based on this evidence, as well as that documented for $e(k)^{\#}_{j,T}$, we conclude that e-loadings based on accruals quality performs at least as well as, or better than, e-loadings based on other measures of earnings quality.

Fourth, we consider whether shorter estimation periods produce e-loadings that reliably capture earnings quality. Shorter estimation intervals are advantageous because they place fewer restrictions on the sample and allow for greater time-variation in e-loadings; they are disadvantageous because they increase the noise in the resulting e-loadings. Our extension investigates how well e-loadings perform when they are calculated at the quarterly level; here, we require a firm to have at least 45 daily returns in calendar quarter Q to estimate its quarterly e-loading, $e^{\#}_{j,Q}$. Results (not tabulated) show that $e^{\#}_{j,Q}$ exhibit similar patterns as found for $e^{\#}_{j,T}$. Further, while the confidence intervals are wider, all of the results for $e^{\#}_{j,Q}$ are reliably significant at the 10% level or better. We conclude that returns-based representations of earnings quality calculated using as few as 45 daily returns reliably capture aspects of earnings quality.

Our fifth and final extension considers the overlap between properties of e-loadings and properties of broader measures of information risk or information asymmetry. The two most prominent such measures are bid-ask spreads and probability of informed trading (PIN) scores. The properties of both measures have been explored primarily in the market microstructure literature (see Huang and Stoll [1997] for bid-ask spreads and Easley, Kiefer and O'Hara [1996] for PIN scores). Huang and Stoll show that the bid-ask spread has three components: (i) processing costs, (ii) inventory costs, and (iii) costs due to adverse selection/information asymmetry. Based on extensive data analysis (they analyze all trades during one year for their sample firms), they conclude that the average adverse information component is 9.6% of the spread. This finding is based on very large firms; it is conceivable that the information asymmetry component is relatively more important for stocks of smaller and younger firms (Easley, Kiefer, O'Hara and Paperman [1996]). Probability of informed trading (PIN) scores incorporate information on spreads as well as order flow information to generate an estimate of the probability that a given trade order originates from a privately informed trader. Firms with larger PIN scores are viewed as having more information asymmetry among investors than firms with smaller PIN scores.

We use bid ask-spread data for Nasdaq firms for 1983-2003,[20] and we use PIN scores for NYSE/AMEX firms for the period 1983-2001.[21] As an initial assessment of the overlap between e-loadings and both bid-ask spreads and PIN scores, we correlate firm-year e-loadings with firm-year average bid-ask spreads and firm-year PIN scores. To control for scale effects, we divide the bid-ask spread by the midpoint of the bid and ask. The Pearson (Spearman) correlation between e-loadings and bid-ask spreads is 0.14 (0.12); the correlation between e-loadings and PIN scores is similar: 0.13 (0.16).[22] All correlations are significant at the 0.01 level. We interpret these results as indicating that while there is statistically non-zero overlap between e-loadings and these measures of information asymmetry, its economic magnitude is limited. This result is consistent with Huang and Stoll's conclusion that a small portion of the bid-ask spread (about 10%) can be linked to information asymmetry.

We probe these results further by orthogonalizing e-loadings with respect to bid-ask spreads and PIN scores, respectively. Our goal is to assess the sensitivity of the main results to the shared variation between e-loadings and bid-ask spreads, and between e-loadings and PIN scores. Specifically, we run the following regressions by year:

---

[20] CRSP reports bids and asks for all securities listed on The Nasdaq National Market since November 1982, and all Nasdaq securities since June 15, 1992. The inside quotation (the highest bid and lowest ask) is used as the closing bid and ask. The source data for February 1986 is missing; hence, this month is also missing on CRSP.

[21] These data are from Søren Hvidkjær's web site [http://www.smith.umd.edu/faculty/hvidkjaer/]. To be included in year T, a stock must further have at least 60 days with quotes or trades in that year. The final sample has between 1,863 and 2,414 stocks per year.

[22] We cannot assess the overlap between bid-ask spreads and PIN scores because the former are available for Nasdaq firms and the latter for NYSE/AMEX firms. However, for a sample of firms where both measures are available, Hasbrouck (2004) reports correlations between closing spreads and PIN scores of 0.40 (Pearson) and 0.66 (Spearman).

$$e_{j,T}^{3f} = \lambda_{0,T}^{BA} + \lambda_{1,T}^{BA} BA_{j,T} + \varepsilon_{j,T}^{BA} \tag{9}$$

$$e_{j,T}^{3f} = \lambda_{0,T}^{PIN} + \lambda_{1,T}^{PIN} PIN_{j,T} + \varepsilon_{j,T}^{PIN} \tag{10}$$

where BA stands for the bid-ask spread and PIN denotes the PIN score. The residual from regression (9), or (10), is the portion of the e-loading that is orthogonalized with respect to bid-ask spreads or PIN scores, $\hat{\varepsilon}_{j,T}^{Orthog}$. Note that this procedure is conservative in the sense that any shared variation between e-loadings and bid-ask spreads (or between e-loadings and PIN scores) will be attributed to the information asymmetry measure and thus 'removed' from the e-loadings.[23] Note further that due to data availability, $\hat{\varepsilon}_{j,T}^{Orthog} = \hat{\varepsilon}_{j,T}^{BA}$ exists for Nasdaq firms for 1983-2003, while $\hat{\varepsilon}_{j,T}^{Orthog} = \hat{\varepsilon}_{j,T}^{PIN}$ exists for NYSE/AMEX firms for 1983-2001. For these sub-periods, we repeat our main tests setting $\hat{\varepsilon}_{j,T}^{Orthog} = \hat{\varepsilon}_{j,T}^{BA}$ for Nasdaq firms and $\hat{\varepsilon}_{j,T}^{Orthog} = \hat{\varepsilon}_{j,T}^{PIN}$ for NYSE/AMEX firms. (Results are similar if we restrict the sample to only bid-ask spread orthogonalized firms or to PIN-score orthogonalized firms.) Table 7 contains the results from tests based on $\hat{\varepsilon}_{j,T}^{Orthog}$. For brevity, we tabulate one market participant test (analyst forecast dispersion), one test of a situation associated with objectively poor earnings quality (restatements), and one over-time test (firm age); all conclusions generalize to other tests. As is evident from Table 7, the behavior of $\hat{\varepsilon}_{j,T}^{Orthog}$ is very similar to the behavior of the non-orthogonalized e-loading, $e_{j,T}^{\#}$, documented in our main tests.

It is possible that the effects of bid-ask spreads and PIN scores are better captured by loadings on factor-mimicking portfolios for each construct rather than on the constructs themselves. Using returns representations of these constructs also allows us to extend the sample to all firms, not just the sample of firms for which spreads and PIN scores are available; further, the resulting BA-loadings and PIN-loadings are directly comparable to e-loadings. To form the BA-mimicking portfolio, we begin by calculating, for each firm-month, the average bid-ask spread; using the monthly average removes some of the variation in

[23] We also performed tests which orthogonalized e-loadings with respect to loadings on factor mimicking portfolios based on bid-ask spreads and PIN scores. Results (not reported) are similar in all respects to those where the orthogonalization is performed with respect to the construct itself (BA or PIN).

daily bid-ask spreads, stabilizing the portfolio. BA-loadings ($BA^{\#}_{j,T}$) are then calculated using the same procedures as described for the e-loadings (i.e., same ranking procedures, returns measurement with a three-month lag, and taking long-short positions). Because PIN-scores are available at the firm-year level, we follow the same procedure as *AQ* for creating a mimicking factor and estimating firm-year specific PIN-loadings ($PIN^{\#}_{j,T}$). Measures of both $BA^{\#}_{j,T}$ and $PIN^{\#}_{j,T}$ are available for the same set of firms for which we have e-loadings: 1983-2003 for $BA^{\#}_{j,T}$ and 1983-2001 for $PIN^{\#}_{j,T}$. We then repeat our tests replacing $e^{\#}_{j,T}$ with $BA^{\#}_{j,T}$ (or $PIN^{\#}_{j,T}$). In all cases, the results (not tabulated) show weaker patterns, and weaker statistical significance of those patterns, relative to tests based on $e^{\#}_{j,T}$.[24] Based on these results, as well as those in Table 7, we conclude that the earnings quality effects captured by e-loadings are not driven by information asymmetry effects captured by bid-ask spreads and PIN scores.

7. Summary and Conclusion

We describe a returns-based representation of earnings quality, in the form of the coefficient estimate (the e-loading) from firm-specific regressions of daily excess returns on a factor-mimicking portfolio capturing earnings quality, controlling for other risk factors. Our analysis is predicated on Francis et al.'s (2005) analysis of accruals quality as a valid empirical measure of information risk as a priced factor. Theoretical support for information risk as a priced factor is provided by analytical models, for example, Easley and O'Hara [2004], Leuz and Verrecchia [2005], and Lambert, Leuz and Verrecchia [2005]. Each of these studies posits a different information risk pricing mechanism: Easley and O'Hara provide a trading model in which better quality reporting reduces the information risk faced by investors who have access to public signals only; Leuz and Verrecchia provide a real effects model in which higher

[24] Further, in the market participant tests where it is straightforward to include $e^{\#}_{j,T}$ as well as $BA^{\#}_{j,T}$ (or $PIN^{\#}_{j,T}$) as independent variables, we find that $e^{\#}_{j,T}$ is always more important than either of the information asymmetry measures in explaining market reactions to unexpected earnings conveyed by earnings announcements, analyst forecast dispersion, and analyst forecast accuracy.

quality reporting supports a better alignment between investors and managers with respect to investment decisions (a solution to an agency problem); and Lambert et al. posit a framework in which information risk may be priced because of the inability to fully specify a forward-looking CAPM beta. We do not distinguish among various reasons put forward by analytical research for *why* information risk is priced (although we regard this as an interesting research question). Rather, our tests build on prior empirical research demonstrating that information risk (as proxied by accruals quality) is priced; we extend this work by developing and validating a returns-based representation of information risk that has several empirical advantages relative to accounting-based representations. While our main tests focus on accruals quality, we note that accruals quality is not the *only* valid empirical measure of information risk; moreover, accruals quality may capture effects other than information risk (although we attempt to control for known effects in our tests). As a returns-based representation of earnings quality, e-loadings are subject to similar caveats.

We document that e-loadings are a reliable returns-based representation of earnings quality as measured by accruals quality. That is, e-loadings are positively associated with other measures of earnings quality; they proxy for the uncertainty in earnings as viewed by investors and by analysts; and they exhibit expected over-time patterns as a function of firm age. Further, in settings where earnings quality has arguably changed (restatements, lawsuits, and bankruptcies), e-loadings show predictable patterns both over-time and in relation to e-loadings for firms which did not experience these events. We find similar patterns for e-loadings based on both annual and quarterly estimations, and for e-loadings that are value-weighted and orthogonalized with respect to firm size versus equal-weighted and not size-orthoganalized.

The demonstrated reliability of yearly and quarterly e-loadings has at least two implications for future research. First, e-loadings impose fewer sampling restrictions, relative to measures of earnings quality calculated using a time-series of either accounting data (such as persistence, predictability and smoothness) or a combination of accounting and market data (such as value relevance, timeliness or conservatism). The longer the time-series required to obtain a reliable estimate of these other measures, the fewer firms with the required series (leading to smaller samples), the greater is the survivorship bias (leading to samples skewed toward larger, successful firms), and the more static (that is, the less time-specific) is the resulting measure. In contrast, because reliable e-loadings can be calculated using only returns data, the resulting samples are more representative of the population of firms (that is, the samples can include smaller, younger firms that lack a time-series of Compustat data) and the e-loadings themselves are not mechanically constrained to be constant or slow to change. The greater representativeness of the sample affords both greater power (due to more variation in attributes of sample firms) and greater generalizability. In addition, because e-loadings can be estimated reliably over intervals as short as a quarter, they can be used to examine the effects of events (such as the promulgation of new accounting standards) or actions (such as earnings management) that might be expected to influence earnings quality at points in time or over fairly short intervals.

A second implication derives from the significant associations between e-loadings and other measures of earnings quality, such as persistence, smoothness, predictability and, to a lesser extent, value relevance, timeliness and conservatism. To the extent these other measures are viewed as capturing aspects of a firm's "true" earnings quality, positive associations between them and e-loadings suggest that investors' *perceptions* of earnings quality are rational. Stated differently, if investors were unable to discern (or unable to discern correctly) the pricing implications of earnings quality, we would not expect e-loadings to correlate with non-perception-based measures of earnings quality or with innate factors that have been shown to explain such measures.

We believe that returns-based measures of earnings quality are most appropriate in research settings that analyze changes in financial reporting quality, either because of a required change in reporting or because of a voluntary reporting or disclosure decision. Returns-based measures are likely to be less appropriate, or inapplicable, in research settings that are characterized by: shifts in both information risk and fundamental risks (such as mergers); explicit predictions about signs and/or magnitudes of effects (there is no theory which predicts the signs or magnitudes of e-loadings); or a requirement to separate total earnings quality into discretionary and innate portions. We also acknowledge that our results do not speak to whether e-loadings based on accruals quality are the pre-eminent returns-based measure of earnings quality. It seems unlikely that any single measure would be best in all research settings and, in fact, our results suggest that e-loadings based on other measures of earnings quality, such as persistence and smoothness, also exhibit substantial reliability. However, our results also indicate that e-loadings based on accruals quality perform at least as well as other measures in the contexts that we examine.

Table 1

Descriptive Information on AQ Sample and Returns Sample[a]

Panel A: Firm representation in AQ Sample and Returns sample

	AQ Sample		Returns Sample			AQ Sample		Returns Sample	
		% traded		% traded			% traded		% traded
Year	# firms	firms	# firms	firms	Year	# firms	firms	# firms	firms
1970	708	28.4%	2,405	96.3%	1988	2,194	28.4%	7,228	93.4%
1971	790	30.3%	2,496	95.6%	1989	2,155	28.9%	7,006	93.9%
1972	873	15.2%	2,648	46.0%	1990	2,249	30.8%	6,951	95.3%
1973	949	16.0%	5,634	94.7%	1991	2,375	32.4%	6,830	93.2%
1974	1,047	19.5%	5,191	96.5%	1992	2,436	31.9%	7,067	92.7%
1975	1,197	23.1%	5,005	96.4%	1993	2,583	31.8%	7,455	91.6%
1976	1,336	25.4%	5,048	96.1%	1994	2,779	31.9%	8,225	94.5%
1977	1,444	27.6%	5,057	96.6%	1995	2,792	30.7%	8,431	92.7%
1978	1,480	28.5%	4,941	95.1%	1996	2,785	28.7%	8,967	92.5%
1979	1,865	36.3%	4,895	95.2%	1997	2,773	27.9%	9,298	93.6%
1980	2,381	44.7%	4,939	92.7%	1998	2,758	28.1%	9,251	94.3%
1981	2,673	46.8%	5,370	94.0%	1999	2,703	28.5%	8,703	91.8%
1982	2,599	43.6%	5,448	91.4%	2000	2,751	30.0%	8,571	93.3%
1983	2,545	38.6%	5,919	89.8%	2001	2,803	33.1%	7,889	93.2%
1984	2,436	35.5%	6,472	94.3%	2002	2,880	37.1%	7,331	94.3%
1985	2,336	33.4%	6,454	92.4%	2003	2,920	39.8%	6,877	93.8%
1986	2,229	30.2%	6,661	90.3%	Mean	2,147	30.9%	6,408	92.4%
1987	2,168	28.1%	7,219	93.6%	Median	2,378	30.2%	6,746	93.7%

Panel B: Descriptive statistics for selected financial variables

	AQ Sample						
	mean	std. dev.	10%	25%	median	75%	90%
Market value	1294.97	4938.84	7.97	24.95	114.08	617.69	2539.33
Total assets	1403.51	4123.69	15.53	43.90	170.94	800.05	3137.31
Sales	1298.30	3451.51	16.99	53.93	210.98	872.37	3109.41
ROA	0.037	0.086	-0.043	0.015	0.046	0.078	0.114
ROE	0.073	0.183	-0.106	0.035	0.105	0.157	0.213

	Returns Sample						
	mean	std. dev.	10%	25%	median	75%	90%
Market value	942.49	4133.90	6.01	17.44	72.41	375.47	1588.85
Total assets	1004.68	3488.89	10.19	27.51	98.75	443.31	1985.70
Sales	907.51	2874.54	9.60	30.40	114.79	483.29	1896.36
ROA	0.026	0.106	-0.078	0.008	0.044	0.078	0.118
ROE	0.055	0.208	-0.179	0.019	0.100	0.156	0.218

Sample descriptions: The AQ Sample consists of all firms with at least seven years of time-series *Compustat* data on working capital accruals and lead, lag and current cash flows to estimate Dechow-Dichev regressions. The Returns Sample includes all firms with at least 100 *CRSP* daily returns in year T.

[a] Panel A reports the distribution of the AQ Sample and, separately, the Returns Sample by year. Panel B reports summary financial data on the size and profitability of the firms in each sample. Market value, total assets and sales are in millions of dollars. Variables are measured over the entire sample period 1970-2003.

Table 2
Associations between e-loadings and Innate Determinants of Earnings Quality
and Other Proxies for Earnings Quality

Panel A: Regressions of e-loadings on innate firm characteristics[a]

	Base model: CAPM		Base model: 3-factor model	
	Coefficient	t-statistic	Coefficient	t-statistic
log(Assets)	-0.0584	-15.86	-0.0339	-13.99
σ(CFO)	0.7026	13.54	0.6550	11.86
σ(Sales)	0.2304	16.10	0.1707	11.51
log (OperCycle)	0.0539	13.62	0.0383	12.13
NegEarn	0.1248	8.70	0.1490	13.55
Adj. R^2	0.2108		0.1154	

Panel B: Correlations between e-loadings and other proxies for earnings quality[b]

	Base model: CAPM		Base model: 3-factor model	
	Pearson	Spearman	Pearson	Spearman
Accruals Quality	0.4397	0.5008	0.4459	0.4918
p-value	*<.0001*	*<.0001*	*<.0001*	*<.0001*
Persistence	0.1653	0.1872	0.2148	0.2575
p-value	*<.0001*	*<.0001*	*<.0001*	*<.0001*
Predictability	0.0207	0.0120	0.0102	0.0760
p-value	*<.0001*	*0.0047*	*0.0162*	*<.0001*
Smoothness	0.2226	0.2023	0.2457	0.2337
p-value	*<.0001*	*<.0001*	*<.0001*	*<.0001*
Value Relevance	0.0088	0.0245	0.0416	0.0278
p-value	*0.0379*	*<.0001*	*<.0001*	*<.0001*
Timeliness	0.0264	0.0383	0.0138	0.0007
p-value	*<.0001*	*<.0001*	*0.0011*	*0.8701*
Conservatism	0.0081	0.0137	0.0034	0.0197
p-value	*0.0553*	*0.0013*	*0.4234*	*<.0001*

[a] Panel A reports the results of annual regressions of e-loadings (based on the CAPM and, separately, 3-factor model) on variables capturing innate features of earnings quality: *log(Assets)* = firm j's log of total assets in year T (a measure of size), $\sigma(CFO)$ = standard deviation of cash flows over year T through T-6, $\sigma(Sales)$ = standard deviation of sales calculated over years T-6 to T, *log(OperCycle)* = log of the sum of days receivables and days inventory), *NegEarn* equals 1 if year T earnings are negative, 0 otherwise. We report the mean value of the 34 annual coefficients; t-statistics are based on the standard errors of the 34 annual coefficients.

[b] Panel B reports the pairwise correlation between e-loadings and several measures of earnings quality: accruals quality (*AQ*), *Persistence* (the AR1 parameter from firm-specific regressions of current earnings per share on lagged earnings per share), *Predictability* (the standard deviation of the residuals from firm-specific AR1 models of earnings), *Smoothness* (the ratio of the standard deviation of earnings to the standard deviation of cash flows), *Value Relevance* (the explained variability of a regression of annual returns on the level and change in earnings per share), *Timeliness* (the explained variability from a Basu (1997) regression of earnings on returns controlling for the sign of those returns), and *Conservatism* (the coefficient on negative returns from the aforementioned reverse regression). Each of these proxies is measured using rolling 10-year windows, and is ordered such that larger values denote poorer quality (that is, following Francis et al. (2004), we use the negatives of the calculated measures for *Persistence, Value Relevance, Timeliness* and *Conservatism*). To compare these 10-year based measures of earnings

quality with our firm- and year-specific e-loadings, we average the e-loadings across the same rolling 10 year windows.

Table 3

Association between e-loadings and Investor Responses to Unexpected Earnings[a]

Panel A: Regression of abnormal returns on unexpected earnings news, conditional on e-loadings

	Base model: CAPM		Base model: 3-factor model	
Indep. variable	Coefficient	t-statistic	Coefficient	t-statistic
UE	0.1575	10.20	0.1487	10.78
UE × e-loading	-0.0415	-3.96	-0.0370	-4.32

Panel B: Regression of abnormal returns on unexpected earnings, conditional on e-loadings and control variables

	Base model: CAPM		Base model: 3-factor model	
	Coefficient	t-statistic	Coefficient	t-statistic
UE	0.5512	14.55	0.5457	14.84
UE × e-loading	-0.0219	-2.16	-0.0239	-3.14
UE × NegEarn	-0.3989	-13.72	-0.3976	-13.76
UE × ln(Size)	-0.0023	-0.75	-0.0017	-0.61
UE × MB	0.0147	3.77	0.0142	3.89
UE × Lev	-0.1155	-4.07	-0.1163	-4.13

[a] This table reports the mean coefficient estimates obtained from 80 quarterly estimations of 2-day market reactions to earnings announcements on unexpected earnings, unexpected earnings interacted with e-loadings, and other variables known to affect the market reaction. *UE* = unexpected earnings conveyed in quarterly earnings announcement, measured using analysts' consensus forecast as the measure of expected earnings. Panel A shows the results of regressing cumulative abnormal returns in days (-1,0) on *UE* and *UE* interacted with e-loadings. Panel B shows similar regressions which include *UE* interacted with other variables known to affect the earnings response coefficient: negative reported earnings (*NegEarn* = 1 if reported earnings in quarter q are negative, 0 otherwise); firm size (*ln(Size)* = log of sales revenues); market-to-book ratio (*MB*); and the ratio of debt to equity (*Leverage*). T-statistics are based on the standard errors of the 80 quarterly coefficients.

Table 4

Relation Between e-Loadings and Forecast Dispersion and Forecast Accuracy[a]

Panel A: Regression of forecast dispersion on e-loadings and control variables

	Base model: CAPM				Base model: 3-factor			
Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
e-loading	0.00159	7.99	0.00080	6.31	0.00160	12.07	0.00081	8.40
Age	-0.00001	-6.34	-0.00001	-6.55	-0.00001	-6.1	-0.00001	-6.47
ln(size)	0.00001	0.22	0.00006	3.03	-0.00011	-7.12	0.00001	0.57
AbsUE	n/a	n/a	0.04593	8.82	n/a	n/a	0.04624	8.79
NegUE	n/a	n/a	0.00073	13.37	n/a	n/a	0.00072	12.88
Loss	n/a	n/a	0.00360	17.76	n/a	n/a	0.00353	17.07
Adj. R^2	0.0337		0.2835		0.0374		0.2848	

Panel B: Regression of forecast accuracy on e-loadings and control variables

	Base model: CAPM				Base model: 3-factor			
Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
e-loading	0.00400	5.62	0.00138	2.92	0.00352	5.39	0.00137	3.49
Age	0.00003	6.25	0.00004	7.84	0.00003	6.25	0.00004	7.86
ln(size)	0.00065	-7.57	0.00040	4.93	-0.00081	-9.64	0.00037	4.62
AbsUE	n/a	n/a	0.39921	10.68	n/a	n/a	0.39775	10.67
NegUE	n/a	n/a	0.00436	11.05	n/a	n/a	0.00436	11.13
Loss	n/a	n/a	0.01718	19.27	n/a	n/a	0.01730	18.60
Adj. R^2	0.0171		0.4540		0.0168		0.4534	

Sample description and variable definitions: The tests in this table require data on both e-loadings and analyst forecasts. For tests of forecast dispersion (Panel A), we require the firm to have at least three forecasts for quarter q. For tests of forecast accuracy (Panel B), we require the firm to have at least one forecast for quarter q. $Dispersion_{j,q}$ = the price-scaled standard deviation of analysts' earnings forecasts for firm j's quarter q made in the three months preceding the quarter q earnings announcement. $|FE|_{j,q}$ = the average value of the price-scaled absolute forecast error for all quarter q forecasts about firm j made in the three months preceding the quarter q earnings announcement. $Age_{j,q}$ = the average number of days between the forecast and the earnings announcement date; $ln(Sales)_{j,q}$ = log of firm j's sales for quarter q; $|UE^{SRW}_{j,q}|$ = the absolute of the seasonal random walk (SRW) measure of the unexpected earnings revealed in the quarter q earnings announcement, scaled by share price; $NegUE^{SRW}_{j,q} = 1$ if $UE^{SRW}_{j,q} < 0$, 0 otherwise; $Loss_{j,q}$ = 1 if quarter q reported earnings are negative, 0 otherwise.

[a] We report the mean coefficient estimates obtained from quarterly regressions (Q=80) of each forecast measure ($Dispersion_{j,q}$ and $|FE|_{j,q}$) on the e-loadings (CAPM and 3-factor), *Age*, and *ln(Sales)*. We also report results which include the earnings surprise measures ($|UE^{SRW}_{j,q}|$, $NegUE^{SRW}_{j,q}$ and $Loss_{j,q}$) as independent variables. T-statistics are based on the standard errors of the 80 coefficient estimates.

Table 5
Yearly Distributions of Restatement, Lawsuit and Bankruptcy Samples

Panel A: Restatement Sample

Year	Original sample	Firms with returns data	Year	Original sample	Firms with returns data
1997	89	82	2000	190	162
1998	100	87	2001	216	199
1999	168	151	2002	119	107
			Total	882	788

Panel B: Lawsuit Sample

Year	Original sample	Firms with returns data	Year	Original sample	Firms with returns data
1990	64	61	1997	72	68
1991	38	35	1998	99	93
1992	56	53	1999	83	74
1993	39	37	2000	80	74
1994	61	59	2001	68	64
1995	41	40	2002	108	94
1996	48	41	Total	857	793

Panel C: Bankruptcy Sample

Year	Original sample	Firms with returns data	Year	Original sample	Firms with returns data
1970	6	4	1987	7	5
1971	4	1	1988	25	12
1972	2	1	1989	14	11
1973	8	6	1990	30	18
1974	6	4	1991	66	43
1975	6	4	1992	53	34
1976	1	0	1993	27	14
1977	4	4	1994	7	3
1978	8	3	1995	10	6
1979	3	1	1996	6	2
1980	5	3	1997	13	12
1981	9	5	1998	25	16
1982	16	13	1999	24	15
1983	17	6	2000	42	25
1984	12	8	2001	52	38
1985	11	6	2002	53	26
1986	15	10	2003	31	12
			Total	618	371

Sample descriptions: The Restatement sample consists of firms with financial statement restatements over 1997-2002 (source, GAO (2002) study on financial restatements). The Lawsuit sample consists of firms who experienced securities class action litigation over the period 1990-2002 where the lawsuit alleged one or more accounting violations. The Bankruptcy sample consists of firms which delisted for bankruptcy reasons over the period 1970-2003. For each sample we report the number of firms in the original sample, and the number of firms with the necessary returns data (and so are available for our tests).

Table 6

Comparison of e-Loadings of Event Firms (Restatement, Lawsuit and Bankruptcy) with Non-Event Firms[a]

Panel A: Restatement Sample

	Base model: CAPM					Base model: 3-Factor				
	Event Sample	Non-Event Sample	Diff.	t-stat.	%-tile	Event Sample	Non-Event Sample	Diff.	t-stat.	%-tile
Year -5	0.2994	0.2691	0.0302	1.00	91	0.1018	0.1292	-0.0274	-0.76	18
Year -4	0.3183	0.2884	0.0299	1.08	94	0.1116	0.1433	-0.0317	-0.98	10
Year -3	0.3390	0.2955	0.0435	1.63	99	0.1321	0.1521	-0.0200	-0.68	20
Year -2	0.3518	0.2773	0.0745	2.89	100	0.1591	0.1636	-0.0045	-0.17	42
Year -1	0.3502	0.2851	0.0651	2.57	100	0.1951	0.1902	0.0049	0.18	58
Year 0	0.3899	0.2681	0.1218	4.32	100	0.3153	0.2072	0.1081	3.57	100
Year +1	0.3368	0.2603	0.0766	2.46	100	0.2694	0.1963	0.0730	2.22	100

Panel B: Lawsuit Sample

	Base model: CAPM					Base model: 3-Factor				
	Event Sample	Non-Event Sample	Diff.	t-stat.	%-tile	Event Sample	Non-Event Sample	Diff.	t-stat.	%-tile
Year -5	0.2756	0.2971	-0.0215	-0.59	20	0.0280	0.1277	-0.0996	-2.61	1
Year -4	0.3028	0.3056	-0.0028	-0.09	49	0.0653	0.1265	-0.0613	-1.74	3
Year -3	0.3116	0.2970	0.0147	0.49	76	0.0313	0.1266	-0.0953	-2.98	1
Year -2	0.3568	0.2795	0.0773	2.81	100	0.0917	0.1433	-0.0516	-1.73	3
Year -1	0.4208	0.2725	0.1483	5.70	100	0.1573	0.1552	0.0021	0.08	60
Year 0	0.4792	0.2722	0.2070	7.15	100	0.2773	0.1631	0.1142	3.48	100
Year +1	0.4742	0.2525	0.2217	6.56	100	0.2857	0.1519	0.1338	3.74	100

Panel C: Bankruptcy Sample

	Base model: CAPM					Base model: 3-Factor				
	Event Sample	Non-Event Sample	Diff.	t-stat.	%-tile	Event Sample	Non-Event Sample	Diff.	t-stat.	%-tile
Year -5	0.4877	0.3319	0.1558	3.81	100	0.2428	0.1135	0.1293	2.86	99
Year -4	0.4957	0.3378	0.1579	4.22	100	0.2431	0.1092	0.1340	3.15	100
Year -3	0.5116	0.3209	0.1908	5.24	100	0.2748	0.1177	0.1571	3.90	100
Year -2	0.5589	0.3035	0.2554	7.13	100	0.3478	0.1284	0.2194	5.33	100
Year -1	0.6186	0.2947	0.3239	7.75	100	0.5158	0.1365	0.3793	7.74	100
Year 0	0.6388	0.2830	0.3558	4.24	100	0.5154	0.1265	0.3888	4.06	100

Sample descriptions: See Table 5 for descriptions of the three Event Samples (Restatement, Lawsuit and Bankruptcy). For each Event Sample we randomly select 100 Non-Event Samples. Each of the 100 Non-Event Samples contains the same number of firms as the Event Sample; observations are selected from all non-event firms in year T so as to replicate the year-by-year distribution in the Event Sample.

[a] The columns labeled "Event Sample" show the mean e-loading for all Event firms in year T=-5,…,+1 (for the Bankruptcy Sample, we end with year 0). The columns labeled "Non-Event Sample" show the average value of the e-loading calculated across the 100 Non-Event Samples. The columns labeled "Diff." show the difference between the mean values of the e-loadings for the Event and Non-Event Samples; the columns labeled "t-stat." show the t-statistic assessing the statistical significance of this difference (from zero). Finally, the columns labeled "%-tile" show the percentile rank of the mean e-loading of the Event Sample within the empirical distribution of mean e-loadings for the Non-Event Samples. For example, a percentile rank of 100 implies that the Event Sample e-loading is the largest e-loading value of the distribution of mean e-loadings for the Non-Event Samples.

Table 7

Tests of E-Loadings Orthogonalized With Respect to Spreads and PIN Scores[a]

Panel A: Regression of forecast dispersion on e^{Orthog} and control variables

	Base model: 3-factor			
Variable	Coefficient	t-statistic	Coefficient	t-statistic
e^{Orthog}	0.0014	10.15	0.0007	6.21
Age	0.0000	-4.68	0.0000	-4.21
ln(size)	-0.0002	-7.98	0.0000	-1.38
AbsUE	n/a	n/a	0.0457	8.79
NegUE	n/a	n/a	0.0008	10.66
Loss	n/a	n/a	0.0035	17.01
Adj. R^2	0.0342		0.2817	

Panel B: Comparison of e^{Orthog} of Restatement Firms with Non-Event Firms

	Base model: 3-Factor				
	Event Sample	Non-Event Sample	Diff.	t-stat.	%-tile
Year -5	0.0065	-0.0043	0.0107	0.2825	61
Year -4	0.0058	-0.0027	0.0085	0.2461	66
Year -3	0.0155	-0.0024	0.0178	0.5978	76
Year -2	0.0307	-0.0035	0.0343	1.2275	97
Year -1	0.0292	-0.0021	0.0313	1.1491	89
Year 0	0.1137	-0.0027	0.1163	3.6674	100
Year +1	0.1012	0.0005	0.1007	2.5855	100

Panel C: Regression of e^{Orthog} on firm age

	Coefficient	t-statistic
Intercept	0.0365	5.58
Age	-0.0031	-10.89
Adj. R^2	0.7509	

Sample description and variable definition: The sample consists of all firm-year observations with data on $\hat{\varepsilon}_{j,T}^{Orthog}$ and the other variables in each test, over 1983-2003. $\hat{\varepsilon}_{j,T}^{Orthog}$ is firm j's e-loading in year T, orthogonalized with respect to bid-ask spreads or PIN scores.

Panel A shows the results of regressing analyst forecast dispersion on $\hat{\varepsilon}_{j,T}^{Orthog}$ as well as control variables. Panel B shows values of $\hat{\varepsilon}_{j,T}^{Orthog}$ for event and non-event firms in the five years leading up to, the year of, and the year after restatement announcements. Panel C reports the results of regressing $\hat{\varepsilon}_{j,T}^{Orthog}$ on firm age.

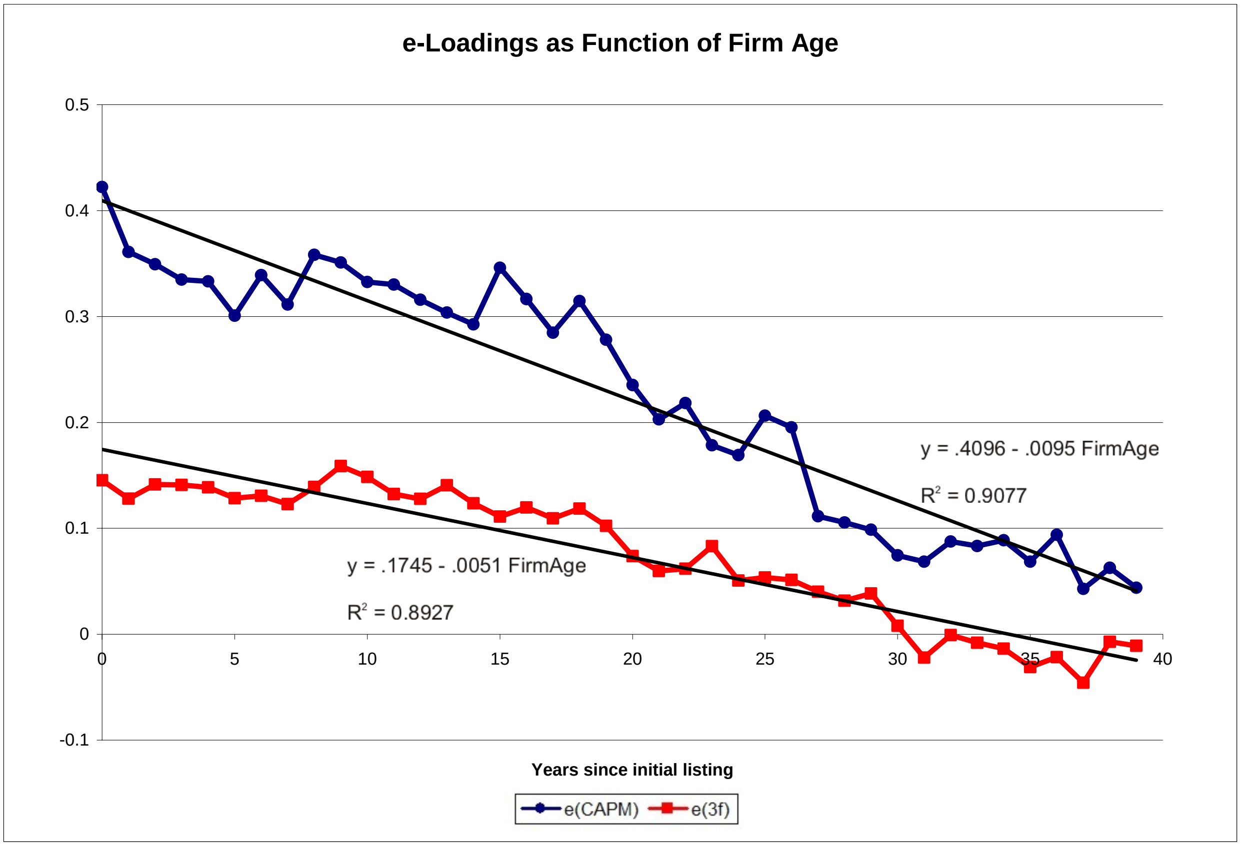

Figure 1 shows the relation between firm age (measured as years since year 0, the year when the firm first listed on any one of the US stock exchanges) and the mean e-loading calculated across all firms in event year T. The graph demonstrates that e-loadings are highest in the early years of a firm's life and decline thereafter. The downward trends are reliably different from zero at the 0.0001 level.

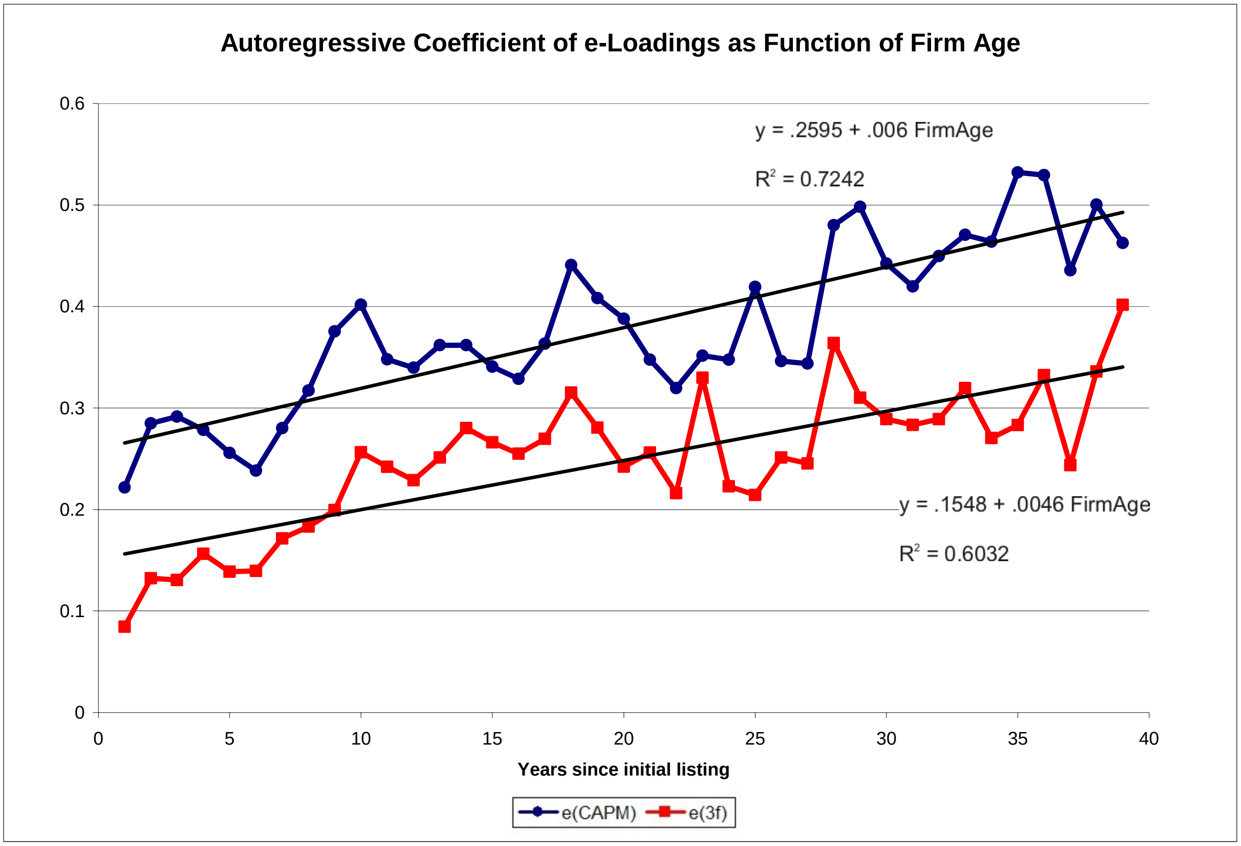

Figure 2 shows the relation between firm age and the stability of the year-over-year e-loading. Stability is measured as the AR1 parameter from a regression of $e^{\#}_{j,T}$ on $e^{\#}_{j,T-1}$, estimated in cross-section for all firms of *FirmAge*=1,…,39+. We report the AR1 parameter for each of the *FirmAge* groups. The graph demonstrates that the stability of e-loadings is lowest in early years of a firm's life and increase thereafter. The upward trends are reliably different from zero at the 0.0001 level.

References

Alford, A., J. Jones, R. Leftwich and M. Zmijewski. 1993. The relative informativeness of accounting disclosures in different countries. *Journal of Accounting Research* 31 (Supplement): 183-223.

Altman, E. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23: 589-609.

Altman, E. 1993. *Corporate financial distress and bankruptcy*. 2nd ed., John Wiley & Sons, New York.

Anderson, K. and T. Yohn. 2003. The effect of 10-K restatements on firm value, information asymmetries, and investors' reliance on earnings. Georgetown University working paper.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24(1): 3-37.

Botosan, C. and M. Plumlee. 2005. Assessing alternative proxies for the expected risk premium. *The Accounting Review* 80(1): 21-53.

Bowen, R., A. Davis, and D. Matsumoto. 2002. Do conference calls affect analysts' forecasts? *The Accounting Review* 77(2): 285-316.

Carhart, M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52: 57-82.

Chen, S., T. Shevlin and Y. Tong. 2004. What is the information content of dividend changes? A new investigation of an old puzzle. Working paper, University of Washington.

Daniel, K. and S. Titman. 1997. Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance* 52: 1-33.

Dechow, P. and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 35-59.

Easley, D., N. Kiefer and M. O'Hara. 1996. Cream-skimming or profit-sharing? The curious role of purchased order flow. *Journal of Finance* 51: 811-833.

Easley, D., N. Kiefer, M. O'Hara, and J. Paperman. 1996. Liquidity, information, and less-frequently traded stocks. *Journal of Finance* 51: 1405-1436.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* LIX, 4: 1553-1583.

Fama, E. and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3-56.

Fama, E and K. French. 1995. Size and book-to-market in earnings and returns. *Journal of Finance* 50: 131-155.

Fama, E. and K. French. 2005. New lists: fundamentals and survival rates. Forthcoming, *Journal of Financial Economics*.

Fama, E. and J. MacBeth. 1973. Risk, return and equilibrium: empirical tests. *Journal of Political Economy* 81: 607-636.

Francis, J., R. LaFond, P. Olsson and K. Schipper. 2004. Costs of equity and earnings attributes. *The Accounting Review* 79(4): 967-1010.

Francis, J., R. LaFond, P. Olsson and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39(2).

Hasbrouck, J. 2004. Trading Costs and Returns for US Equities: The Evidence from Daily Data. New York University working paper.

Heflin, F., K.R. Subramanyam and Y. Zhang. 2003. Regulation FD and the financial information environment: early evidence. *The Accounting Review* 78(1): 1-37.

Heninger, W. 2001. The association between auditor litigation and abnormal accruals. *The Accounting Review* 76, 1: 111-126.

Hribar, P. and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40: 105-134.

Huang, R. and H. Stoll. 1997. The Components of the Bid-Ask Spread: A General Approach. *The Review of Financial Studies* 10: 995-1034.

Imhoff, E. and G. Lobo. 1992. The effect of ex ante earnings uncertainty on earnings response coefficients. *The Accounting Review* 67: 427-439.

Lamont, O., C. Polk, and J. Saá-Requejo 2001. Financial constraints and stock returns. *The Review of Financial Studies* 14: 529-554.

Lang, M. 1991. Time-varying stock price response to earnings induced by uncertainty about the time-series properties of earnings. *Journal of Accounting Research* 29(2): 229-257.

Lang, M. and R. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71(4): 467-492.

Leuz, C., and R. Verrecchia. 2005. Firms' capital allocation choices, information quality, and the cost of capital. University of Pennsylvania working paper.

Lewellen, J. 1999. The time-series relations among expected return, risk and book-to-market. *Journal of Financial Economics* 54: 5-43.

Lys, T. and R. Watts. 1994. Lawsuits against auditors. *Journal of Accounting Research* 32 (Supplement): 65-93.

McNichols, M. 2002. Discussion of 'The quality of accruals and earnings: the role of accrual estimation errors.' *The Accounting Review* 77 (Supplement): 61-69.

Moore, E. and R. Pfeiffer. 2004. The effects of financial statement restatements on firms' financial reporting strategies. University of Massachusetts working paper.

Newey, W. and K. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703-708

Pastor, L., and P. Veronesi. 2003. Stock valuation and learning about profitability. *Journal of Finance* 58: 1749-1789.

Richardson, S., I. Tuna and M. Wu. 2002. Predicting earnings management: the case of earnings restatements. University of Pennsylvania and Hong Kong University of Science and Technology working paper.

Stice, J. 1991. Using financial and market information to identify pre-engagement factors associated with lawsuits against auditors. *The Accounting Review* 66(3): 516-553.

Subramanyam, K.R. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22: 249-281.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76: 357-373.