# Forecasting risk in earnings*

**Theodosia Konstantinidi,** *Cass Business School, City University London*

**Peter F. Pope,** *London School of Economics and Political Science*

October 2014

*Contemporary Accounting Research, Forthcoming*

### Abstract

Conventional measures of risk in earnings based on historical standard deviation require long time series data and are inadequate when the distribution of earnings deviates from normality. We introduce a methodology based on current fundamentals and quantile regression to forecast risk reflected in the shape of the distribution of future earnings. We derive measures of dispersion, asymmetry and tail risk in future earnings using quantile forecasts as inputs. Our analysis shows that a parsimonious model based on accruals, cash flow, special items and a loss indicator can predict the shape of the distribution of earnings with reasonable power. We provide evidence that out-of-sample quantile-based risk forecasts explain incrementally analysts' equity and credit risk ratings, future return volatility, corporate bond spreads and analyst-based measures of future earnings uncertainty. Our study provides insights into the relations between earnings components and risk in future earnings. It also introduces risk measures that will be useful for participants in both the equity and credit markets.

*Keywords: Earnings; accruals; fundamentals-based risk forecasts; quantile regression.*

JEL classifications: M41; C13

*Accepted by K.R. Subramanyam. We especially thank K.R. Subramanyam (editor) and two anonymous reviewers for substantially improving this paper. We also appreciate helpful comments from Mark Bradshaw, Peter Easton, Peter Joos, Per Olsson, Stephen Penman, Scott Richardson, Jerold Zimmerman and participants at the Tel Aviv Accounting Research Symposium, 2011, the JP Morgan Quant Conference 2011, the INQUIRE (UK) Seminar 2011, the RSM/ESE Conference on Accounting and Valuation 2012, the 35th EAA Annual Congress 2012 and the 13th FRAP Conference 2013. The programming assistance in Matlab by Torben Latza and Riccardo Pianeti is gratefully acknowledged.

## 1. Introduction

Estimates of future corporate earnings are important inputs to security valuation models because future payoffs (e.g. dividends and debt repayments) depend on future earnings. Yet, although the identification and estimation of risk is a dominant theme in the finance literature, research establishing connections between future earnings outcomes, current accounting fundamentals and risk is relatively under-developed (Penman 2010). In this paper we contribute to closing this gap in the accounting literature by introducing a methodology for forecasting the shape of the conditional distribution (or density) of future earnings. We measure properties of the predicted distribution which are relevant to decision makers interested in assessing the fundamental risk in future earnings. We show that a set of robust risk metrics requiring forecasts of only seven quantiles is useful in explaining a number of outcomes where risk is important.

We characterize risk in earnings in terms of the predicted higher moments of the conditional distribution of future earnings (i.e. dispersion, skewness and kurtosis), noting that the potential relevance of each distributional attribute depends on the decision context and the decision makers' loss functions (Tay and Wallis 2000). We decompose estimated kurtosis into upside and downside tail risk components because some investors may put greater weight on low earnings outcomes than on high outcomes. For example creditors are more concerned with downside risk, because default risk increases when earnings are low. Our approach involves two steps; we first forecast a set of conditional quantiles of one-year-ahead earnings using quantile regressions (Koenker and Bassett 1978), and second we use the quantile forecasts as inputs to the calculation of robust risk measures. Our approach assumes that future earnings of firms with identical current characteristics are drawn from the same distribution, and that heterogeneity in predicted future earnings outcomes can be summarized in a set of risk measures that together adequately describe the conditional distribution of earnings.

The statistics and econometrics literatures suggest a number of parametric approaches for estimating the conditional moments of a distribution.[1] Our approach establishes robust risk measures that do not rely on strong assumptions concerning the parametric distribution describing earnings. It is similar in spirit to Kim and White (2004) who estimate the unconditional skewness and kurtosis of stock returns, and to White et al. (2008) who model the conditional moments of stock returns in a time series econometrics framework. The estimation of quantile-based risk measures is computationally very efficient because it requires forecasts of only seven representative quantiles of the distribution. Moreover, it does not suffer from practical limitations of applying time series methods to forecasting risk in earnings, where available data are inherently low frequency and typically have short historical time series.

The quantile forecasting model utilizes a parsimonious set of accounting variables including cash flow, accruals, special items and a loss indicator. We expect that the earnings decomposition will be useful in predicting the earnings distribution because the three earnings components contain differential information about future performance. Accruals contain intentional or unintentional errors that ultimately must reverse; and the dynamics of earnings and earnings components are different for loss firms as a result of more timely recognition of economic losses compared to economic gains (Basu 1997; Ball and Shivakumar 2005). The forecasting model includes as inputs only fundamental accounting information and therefore risk forecasts are independent of stock prices.[2] Despite the restricted set of predictor variables, we find that our model predicts future earnings beyond the conditional mean with reasonable power. Forecasting coefficients vary in clear patterns across quantiles, as must be the case if higher moments of the distribution of earnings depend on the predictors. They also vary significantly across the predictors, indicating that disaggregated financial statement items contain independent information relevant to assessing risk in earnings.

[1] We discuss these methods in section 2.

[2] While the focus of this paper is on constructing risk estimates based exclusively on fundamental accounting information, stock price-based inputs could be included as predictor variables in other applications of this methodology.

We test the empirical validity of our risk metrics using a number of outcome variables that relate to risk. Our results indicate that our risk metrics are strongly associated with equity and credit analysts' risk ratings, and with market outcomes including future stock return volatility and corporate bond spreads. In addition, the risk metrics predict analysts' exclusions from GAAP earnings and the absolute value of analysts' forecasts errors, a commonly used measure of earnings uncertainty. In all of our tests, we control for an extensive set of risk proxies used in the prior literature. Our study suggests that risk is multi-dimensional and reveals which dimensions of risk are relevant to market participants. For example, downside tail risk appears to be more important than upside tail risk in explaining the risk ratings of equity and credit analysts.

Overall, we contribute to the accounting literature by showing that a parsimonious set of fundamental accounting items can capture risk in future earnings with reasonable power. Despite the relative simplicity of our model, the risk metrics are capable of explaining variation in market outcomes and analysts' decisions where risk is important. Our tests bring new insights to Lui et al. (2007), who show that equity analysts' risk ratings are explained by risk proxies used in the prior literature and are incrementally informative in predicting equity return volatility; and Joos et al. (2014) who develop an equity risk measure based on the spread in analysts' target prices between bull and bear states. Our approach offers the prospect of developing more comprehensive analyses linking a broader set of fundamental and contextual variables to risk in earnings. This is turn should be informative for security valuation.

## 2. Research Design

### *2.1 Forecasting the shape of the earnings distribution*

A number of approaches to density forecasting can be identified in the statistics and econometrics literatures (see Tay and Wallis 2000 for a review). Time series approaches are the most popular in forecasting equity risk and typically rely on ARCH/GARCH-class volatility models using daily (or higher frequency) stock returns.[3] While less common in practice, time series models can include exogenous firm-specific or macroeconomic state variables to enhance predictability and to capture changes in regimes, e.g. across economic cycles. In accounting, Sheng and Thevenot (2012) were the first to forecast earnings volatility using GARCH-class time series volatility models, complementing standard approaches to estimating earnings volatility from historical data (e.g. Baginski and Wahlen 2003). However, in a comprehensive review of the time series approach to modeling volatility, Poon and Granger (2003) conclude that "naïve" estimates of historical volatility perform as well as conditional forecasts derived from more sophisticated volatility models (Poon and Granger 2003, p. 507).

Earnings volatility forecasts obtained from time-series models only capture accounting numbers to the extent that relevant information is reflected in individual firms' earnings histories. Our approach to modeling risk is complementary to a time series approach as it directly exploits cross-sectional heterogeneity in accounting fundamentals to predict cross-sectional differences in risk.[4] In light of the prominence of time series risk measures in the prior literature, we include measures of historical return volatility and earnings volatility as control variables in our empirical validation tests of our risk metrics.

We model the firm-specific density of one-year-ahead earnings using quantile regression (Koenker and Bassett 1978; Angrist and Pischke 2009, ch.7). This method is well-suited to cross-sectional modelling and does not require us to make distributional assumptions. If the conditional quantile function is linear in the predictor variables,[5] quantile regression solves the following minimization problem (Angrist and Pischke 2009, p.271):

$$Q_{i\tau}(EARN_{it+1}|X_{it}) = \arg\min_{\alpha_\tau} E[\rho_\tau(EARN_{it+1} - X'_{it}\alpha_\tau)] \quad (1)$$

[3] See Poon and Granger (2003) for a review.
[4] For identical reasons our paper is also complementary to recent attempts to forecast conditional skewness in high frequency stock returns, by modifying the symmetric distributional assumptions in GARCH-class models of conditional skewness (Harvey and Siddique 1999; León et al. 2005; Bauwens and Laurent 2005).
[5] Theoretical results in the literature show that the linear approximation assumed in quantile regression approximates the true conditional quantile function, analogous to OLS approximating the true conditional expectation (Angrist and Pischke 2009, p. 278).

where $EARN_{it+1}$ is earnings for firm i in year t+1, $X_{it}$ is a vector of predictor variables, in our case earnings and earnings components for firm i in year t, and $\alpha_\tau$ is a vector of quantile regression coefficients specific to quantile $\tau$. Denoting the quantile regression error $u_{it+1} = EARN_{it+1} - X_{it}'\alpha$, the so-called check function $\rho_\tau(u_{it+1}) = 1(u_{it+1} > 0)\tau|u_{it+1}| + 1(u_{it+1} \leq 0)(1-\tau)|u_{it+1}|$ weights the quantile forecast errors $u_{it+1}$ asymmetrically depending on their sign and on the respective quantile value $\tau$.[6] Estimating equation (1) for a range of quantiles allows us to forecast the shape of the distribution of earnings for firm i in year t+1 as a function of the predictor variables. As the predictor variables change for a different firm and year, the firm-year-specific conditional distribution of future earnings changes position and shape.

In principle the number of conditional quantile functions we could estimate is very large, but in our empirical tests we report quantile function estimates in the set $\tau \in \{0.01, 0.05, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 0.95, 0.99\}$. This range of quantiles is sufficient to capture the shape of the conditional distribution of future earnings and to describe the moments of the distribution using robust non-parametric measures from the statistics literature. Slope coefficients on predictor variables will only differ across quantiles if there is conditional heteroskedasticity, but coefficients will also differ across quantiles due to more general dependence between higher moments and predictor variables.[7]

We note that the quantile regression framework is not the only cross-sectional modeling approach to forecasting the conditional density of earnings. For example, one could model the conditional distribution of future earnings using the asymmetric least squares (ALS) method (Newey and Powell 1987; Yao and Tong 1996).[8] Alternatively, one could capture conditional variance (dispersion) in future earnings by

[6] Quantile regression can be implemented in various statistical software packages including SAS, using proc quantreg, and STATA, using the qreg command.
[7] See Angrist and Pischke (2009, p. 274-5) for a discussion of conditional heteroskedasticity and quantile regression properties.
[8] This approach estimates the *expectiles* of a distribution defined as the expectation of the exceedances beyond an expected quantile. Expectiles and quantiles are related (Efron 1991; Taylor 2008) but in contrast to quantiles, marginal changes in one part of a distribution affect all estimated expectiles (Taylor 2008). Since our main objective in this paper is to characterize risk in terms of robust measures of the moments of the distribution requiring quantile estimates, we believe that the use of quantile regression offers advantages in our setting. However the applicability

regressing the squared (or absolute) value of the residuals from an earnings forecasting model on predictor variables (Granger and Ding 1995).[9] We expect estimates of dispersion to be highly correlated across alternative methods, if they are conditioned on the same information set. As a consequence we do not claim to have identified the *best* method for modelling the shape of the earnings distribution, but the widespread application of quantile regression in other areas of economics and statistics does suggest that it is a useful technique.

*2.2 Definition of Risk Metrics*

Our risk metrics summarize the shape of the distribution of earnings in terms of robust measures of higher moments. We use a small set of seven out-of-sample conditional quantile forecasts of earnings to compute measures of conditional dispersion, skewness and kurtosis.

We capture conditional dispersion in the future earnings distribution using the predicted interquartile range (IQR), defined as:

$$\mathrm{IQR}_i = Q_{i75} - Q_{i25} \tag{2}$$

where for notational simplicity we use $\mathrm{Q}_{i\tau} = Q_{i\tau}(EARN_{it+1}|X_{it})$ to denote the estimated conditional τ'th quantile of earnings for firm i in year t+1. IQR is a commonly used and robust measure of dispersion which is proportional to the variance of EARN if the conditional variance of EARN is linear in the predictor variables (Koenker and Bassett 1982; Angrist and Pischke 2009, p. 274-5). Higher values of IQR reflect higher levels of uncertainty in future earnings realizations. We note that use of dispersion measures such as IQR in economic decisions should be most informative when benchmarked against a

of ALS estimates of expectiles is a potentially interesting avenue for future research extending our work. We thank an anonymous reviewer for pointing us to this literature.

[9] Tests on our data reveal that an uncertainty measure based on predictions of squared or absolute residuals has a high correlation with our uncertainty measure IQR. We thank an anonymous reviewer for this suggestion. A similar approach could in principle be followed for modeling higher moments of the distribution, using higher powers of the residuals as dependent variables in the second stage. Estimates of skewness and kurtosis using this approach have low correlation with our quantile-based risk measures, perhaps because higher order powers of first stage regression residuals are severely affected by outliers.

measure of central tendency. In regression tests we use the IQR dispersion measure but control directly for the conditional median expectation, $Q_{i50}$.

We estimate conditional skewness in the future earnings distribution using the following robust measure (see Bowley 1920; Hinkley 1975; Kim and White 2004):

$$SKEW_i = [(Q_{i75} - Q_{i50}) - (Q_{i50} - Q_{i25})]/IQR_i. \tag{3}$$

SKEW captures the balance between upside risk relative to downside risk in future earnings within the two middle quartiles. It is normalized by IQR to vary between −1 and +1, with a value of zero indicating a symmetric distribution within the interquartile range. In empirical tests we include SKEW together with IQR because the two measures are complementary.

We estimate conditional kurtosis in the future earnings distribution using the Moors (1988) statistic computed as:

$$KURT_i = [(Q_{i87.5} - Q_{i62.5}) + (Q_{i37.5} - Q_{i12.5})]/IQR_i. \tag{4}$$

KURT measures the density of the distribution close to $Q_{25}$ and $Q_{75}$ relative to the density close to $Q_{50}$ and it is normalized by IQR. Therefore KURT can be thought of as a robust measure of conditional kurtosis in future earnings. In empirical tests we include KURT together with IQR because the two measures are complementary. We also capture asymmetry in the relative densities of the distribution surrounding $Q_{25}$ and $Q_{75}$ by decomposing the numerator of KURT as follows:

$$UP_i = (Q_{i87.5} - Q_{i62.5})/IQR_i \tag{5}$$

$$DOWN_i = (Q_{i37.5} - Q_{i12.5})/IQR_i. \tag{6}$$

UP and DOWN are more sensitive to asymmetry in the outer quantiles of the distribution than SKEW, which ignores information in the conditional quantiles below $Q_{25}$ and above $Q_{75}$.

Quantile-based risk measures have a number of advantages over existing firm-specific measures of risk and uncertainty in the literature. First, unlike time-series estimates of volatility, skewness and kurtosis, they impose no survivorship requirement and can be estimated for all firms for which current predictor variables are observable.[10] Second, they have the advantage of being robust to outliers compared to analogous conventional moment-based measures and are known to perform well for data from a wide range of probability distributions (Kim and White 2004; Ghysels et al. 2011). Finally, the information set that can be used to estimate the quantile regressions is theoretically unlimited and can be expanded to increase the power of earnings risk forecasts. We do not claim to have identified the most powerful set of predictor variables in the tests that follow.

*2.3 Forecasting model specifications*

We forecast the quantiles of future earnings based on the following model:[11]

$$EARN_{t+1} = \alpha_{01}d^{+} + \alpha_{02}d^{-} + \alpha_{11}ACC_{t}.d^{+} + \alpha_{21}OCF_{t}.d^{+} + \alpha_{31}SI_{t}.d^{+} + \alpha_{12}ACC_{t}.d^{-} + \alpha_{22}OCF_{t}.d^{-} + \alpha_{32}SI_{t}.d^{-} + v_{t+1} \quad (7)$$

where EARN is earnings, ACC is total accruals, OCF is operating cash flow, SI is special items, $d^{+}$ is an indicator variable equal to one if $EARN_t \geq 0$, $d^{-}$ is an indicator variable equal to one if $EARN_t < 0$ and $v$ is the forecast error. All accounting variables are scaled by average total assets. We include industry fixed effects to allow for differences in the distributions of earnings across industries.

Prior literature has shown that accruals and cash flow are useful in predicting the mean of future earnings (Sloan 1996; Richardson et al. 2005). We build on this literature by testing whether the two components are capable of forecasting the shape of the distribution of earnings. Accruals contain information about future cash flow, but they also contain errors due to their dependence on estimates and judgments which affects their reliability and persistence. They can also be manipulated in favor of current earnings outcomes, but at the expense of future earnings outcomes. For these reasons, we expect that

[10] For example, in our empirical tests, the time-series estimate of earnings volatility (EarnVol) can be computed for only 24,559 firm-years out of 36,232 firm-years, despite requiring data for a minimum of only five years.
[11] For simplicity in the rest of the paper we drop the firm subscript i in all symbolic notation. However it is important to keep in mind that conditional quantile measures and the risk metrics derived from the quantile estimates are firm-specific.

accruals and cash flow will have separate roles in forecasting the distribution of future earnings, i.e. $\alpha_{11} \neq \alpha_{21} \neq 0,\ \alpha_{12} \neq \alpha_{22} \neq 0$. We also expect that special items, included in both accruals and cash flow, will have an incremental role in forecasting the future earnings distribution. Special items are likely to be associated with risk-related events, including financial distress and divestment of noncore lines of business, and therefore they might capture incremental information about the risk in future earnings, i.e. $\alpha_{31} \neq 0,\ \alpha_{32} \neq 0$.

Prior literature focusing on explaining the conditional mean of future earnings has shown that the persistence of earnings is asymmetric, with losses being more transitory than profits (Hayn 1995). Because shareholders have the option to liquidate the firm, losses cannot persist indefinitely. Additionally, timely loss recognition implies that accruals will have large transitory components when a firm suffers economic loss (Basu 1997). Assuming that accounting loss is a proxy for economic loss, we expect that the roles of ACC, OCF and SI in predicting the earnings distribution will vary based on whether the firm reports current losses or profits. In the empirical analysis we test whether $\alpha_{11} = \alpha_{12}$ and $\alpha_{21} = \alpha_{22}$.

Our model intentionally adopts a parsimonious set of predictor variables to illustrate application of the risk estimation methodology and its implications for equity and credit markets. However, as noted above, the set of predictor variables could be enhanced. Subsequent work by Chang et al. (2013) and Correia et al. (2013) has expanded our model to include a broader set of instruments using both earnings and financial policy attributes.[12] Chang et al. (2013) also use a more computationally intensive approach than ours to estimate the cumulative conditional distribution function of future earnings based on quantile forecasts from 125 quantile regressions. They use the quantile approach to estimate the probability of future losses and moments of the distribution of future earnings using conventional statistical expressions. Correia et al. (2013) use both conventional and robust quantile-based estimates of conditional dispersion in earnings to predict bankruptcy risk and credit spreads.

[12] Chang et al. (2013) estimate quantiles of future return on equity and Correia et al. (2013) estimate quantiles of future return on net operating assets.

## 3. Sample Selection and Data

### *3.1 Forecasting model data*

We obtain accounting and pricing data from the CRSP/Compustat merged database (CCM). Our initial sample covers all ordinary common stocks listed on NYSE, AMEX and Nasdaq for the period 1988–2009. The beginning of the sample period reflects the availability of cash flow data based on *SFAS 95* Statement of Cash Flows. Cash flow data are required to compute accruals, as recommended by Hribar and Collins (2002). We exclude all firm-year observations with SIC codes in the range 6000–6999 (financial firms) because the behavior of earnings and other financial statement numbers for these firms is different.[13]

The main variables used in our forecasting model are defined as follows. EARN is earnings before extraordinary items taken from the cash flow statement (*IBC*)[14], OCF is operating cash flow (*OANCF*) minus extraordinary items and discontinued operations (*XIDOC*) and SI is special items (*SPI*). Following Hribar and Collins (2002), we compute accruals as ACC ≡ EARN – OCF. We deflate earnings, accruals, cash flow and special items in each year t by the average total assets in the year. To mitigate the effect of outliers in our forecasting regressions, we delete observations in the extreme top and bottom percentiles of the distributions of deflated EARN, OCF, ACC and SI in each year. We also restrict the sample to firms with full data to estimate our forecasting equations and with total assets in excess of $100 million to avoid the influence of small firms, as in Dichev and Tang (2009). Our requirements result in an initial sample of 43,526 firm-year observations. The sample formation is summarized in Panel A of Table 1. Panel B of Table 1 reports the industry decomposition of the sample, using the 12 industry definitions in Barth et al. (1999) excluding financial firms.

[13] We use the historical *SICH* code from Compustat when it is available and the current *SIC* code when *SICH* is missing.
[14] Compustat item labels (*XFP* names) for accounting variables are in parentheses.

### *3.2 Validation test data*

We use several additional data sources to obtain outcome variables as well as control variables for our validation tests. We employ data from the I/B/E/S summary files to obtain earnings forecasts and pro forma earnings used in estimating analysts' exclusions and analysts' forecast errors. In tests using analysts' risk ratings, we use a proprietary dataset of monthly ratings of stock-specific investment risk reported by financial analysts in a large securities firm. Data are available to us from January 2003. Tests based on stock return volatility use data from CRSP and tests on credit ratings use data from Compustat. Finally, tests based on bond yields utilize data from the Fixed Investment Securities Database (FISD) and TRACE databases. From FISD we obtain issue- specific information, such as issuance date, issuance size, maturity date, coupon rate as well as variables that identify bonds with special features. From TRACE we obtain bond yields and prices for each transaction during the period 2002 – 2010. All variables used in the empirical tests are defined in Appendix A.

## 4. Results

### *4.1 Descriptive statistics*

In Table 2 we present descriptive statistics for each of the main variables used in estimating the quantile regression models. Panel A reports distributional statistics. Consistent with previous research (e.g. Sloan 1996; Barth et al. 2001) accruals are negative on average because they include depreciation and amortization, while operating cash flow are positive on average because they exclude investing cash flow. The distribution of earnings is negatively skewed and displays excess kurtosis, and these properties are largely due to the distributional properties of accruals. The correlations in Panel B of Table 2 also confirm that the relations between our main variables in our forecasting models are consistent with prior research.

*4.2 Quantile regression estimates*

We estimate forecasting equation (7) for eleven quantile values $\tau \in \{0.01, 0.05, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 0.95, 0.99\}$, although we set aside the four most extreme quantiles for the purpose of calculating the risk measures described in Section 2.2. We also show results for direct estimates of the earnings uncertainty measure IQR. Since IQR $= Q_{75} - Q_{25}$, these direct estimates are useful in showing whether quantile function coefficients at $Q_{25}$ and $Q_{75}$ are significantly different, as well as showing how our measure of earnings uncertainty depends on predictor variables. For reasons explained earlier, the significance of predictor variables in the IQR regression is informative about the nature of conditional heteroskedasticity. To facilitate comparisons with prior research we also report OLS regression results for the same model.[15] While we include industry fixed effects in all estimations, to conserve space we do not report them. Rather we report the mean of the effective industry-specific intercepts (equal to the intercept plus the industry fixed effect).[16] We also report pseudo $R^2$ statistics for the quantile regressions and adjusted $R^2$ statistics for the OLS regression, as well as the corresponding incremental $R^2$ statistics relative to a model containing only industry fixed effects.

Panel A of Table 3 contains coefficient estimates for forecasting equation (7). The results from the quantile regressions are contained in the columns labelled 1% − 99% and the direct estimates for predicting our earnings uncertainty metric IQR are contained in the final column. Panel B contains tests of whether the coefficients on ACC and OCF are equal for profit or loss firms and whether coefficients on ACC and OCF differ between profit and loss firms.

The following examples based on the $Q_{75}$ and $Q_{25}$ regressions illustrate how to appropriately interpret the quantile regression results. The forecasting coefficients at the 75$^{th}$ and 25$^{th}$ quantiles for the profit sample suggest that a change of one unit in ACC (OCF) is associated with a change of 0.860 (0.917) units

[15] Reported results relating to the forecasting model are based on panel regressions for the full sample period. In unreported robustness tests we replace the panel regressions by Fama-MacBeth (1973) estimates with the Newey-West autocorrelation adjustment, in order to control for cross-sectional and serial correlation in the residuals of the quantile regressions. Results are qualitatively the same. Details are available from the authors on request.
[16] Details are available from the authors on request.

in the predicted value of $Q_{75}$ and a change of 0.700 (0.809) units in the predicted value of $Q_{25.}$ Consider a representative profit firm having median levels of ACC and OCF within the profit sample equal to −0.044 and 0.102 respectively and special items equal to zero (i.e. EARN = −0.044 + 0.102 = 0.058). For this firm, the predicted value of $Q_{75}$ is 0.067 [= 0.012 + 0.860 × (−0.044) + 0.917 × 0.102] while the predicted value of $Q_{25}$ is 0.032 [= −0.019 + 0.700 × (−0.044) + 0.809 × 0.102]. In other words, there is a 25 percent probability that earnings next year will be above 0.067 and a 25 percent probability that earnings next year will be below 0.032. In this case the earnings uncertainty metric IQR has a value of 0.035, equal to the difference between the predicted values of $Q_{75}$ and $Q_{25}$.

The interpretation of the quantile estimates for the loss sample is similar. The forecasting coefficients at the 75$^{th}$ and 25$^{th}$ quantiles suggest that a one unit change in ACC (OCF) is associated with a change of 0.359 (0.516) units in the predicted value of $Q_{75}$ and a change of 0.784 (0.984) units in the predicted value of $Q_{25.}$ Consider now a representative loss firm having median levels of ACC and OCF within the loss sample of −0.095 and 0.032 respectively and special items equal to zero (i.e. EARN = −0.095 + 0.032 = −0.063). For this firm, the predicted value of $Q_{75}$ is −0.005 [= 0.013 + 0.359 × (−0.095) + 0.516 × 0.032] while the predicted value of $Q_{25}$ is −0.089 [= −0.046 + 0.784 × (−0.095) + 0.984 × 0.032]. In this case, there is a 25 percent probability that earnings next year will be above −0.005 and a 25 percent probability that earnings next year will be below −0.089. IQR is then equal to 0.084 showing that there is considerably more uncertainty about future earnings outcomes for the loss firm than for the profit firm.

Comparison of the quantile estimates between profit firms ($d_1$ = 1) and loss firms ($d_2$ = 1) in Panel A reveals interesting differences in the behavior of the ACC and OCF forecasting coefficients. The monotonically declining pattern of both coefficients for loss firms is consistent with the conditional variance of the distribution of future earnings being negatively related to both ACC and OCF for loss firms (Angrist and Pischke 2009, p.274−5). This prediction is confirmed by the direct quantile regression estimate for IQR where the coefficients on both $ACC.d_2$ and $OCF.d_2$ are negative. The positive coefficient on $SI.d_2$ indicates that when special items are present, the effects of ACC and OCF on earnings uncertainty are partially mitigated.

Contrary to the results for loss firms, the patterns of coefficients on ACC and OCF for profit firms are not monotonic across all quantiles. However, with the exceptions of the most extreme (and noisiest) $Q_1$ and $Q_{99}$ regressions, coefficients are generally lower for quantiles below $Q_{50}$ than for quantiles above $Q_{50}$. This suggests that for profit firms, the conditional variance of the distribution of future earnings is positively related to both ACC and OCF. The direct IQR quantile regression estimate in the final column confirms that our estimated earnings uncertainty metric is indeed positively and significantly related to both $ACC.d_1$ and $OCF.d_1$, with the sensitivity of IQR to ACC being almost 60% higher than for OCF. The marginal effect of SI is again opposite in sign, indicating that the presence of special items within accruals and/or cash flow mitigates the effects of ACC and OCF on earnings uncertainty. Comparisons of the coefficient estimates for profit and loss firms in the IQR regression indicate that in addition to the opposite signs in the dependence of IQR on ACC and OCF, the marginal sensitivities of IQR to both ACC and OCF are lower for profit firms than for loss firms.

We provide visual evidence on the quantile regression results in Figure 1, where we plot the cumulative distribution function conditional on ACC and OCF. In Panel A, we hold OCF constant at its median level in the profit sub-sample (median OCF = 0.102) and the loss sub-sample (median OCF = 0.032) and we allow ACC to vary. In Panel B, we hold ACC constant at its median level in the profit sub-sample (median ACC = −0.044) and the loss sub-sample (median ACC = −0.095) and we allow OCF to vary. SI is assumed to be zero. The figure clearly shows that for all three levels of ACC (Panel A) and all three levels of OCF (Panel B) that we consider in each subsample, the distribution of future earnings for loss firms is considerably more dispersed than the distribution for profit firms. Furthermore, it is noticeable that the distribution for loss firms becomes wider as ACC or OCF decrease, consistent with our quantile regression results relating negatively IQR with ACC and OCF. In contrast, the distribution for the profit sample becomes narrower as ACC and OCF decrease, consistent with the positive relation between IQR and both ACC and OCF.

Further comparison of quantile regression estimates reveals that the forecasting coefficients on OCF are uniformly higher than for ACC across all quantiles irrespective of the sign of $EARN_t$ and differences between OCF and ACC coefficients are statistically significant except for $Q_{99}$ when $EARN_t < 0$. Hence decomposing earnings into accruals and cash flow components enhances the ability of the model to predict the shape of the distribution of future earnings. Results also indicate that the slope coefficients on both ACC and OCF are significantly different between profit and loss firms across all but one quantiles (Panel B). This indicates that forecasting power improves by conditioning coefficient estimates on profits or losses. With one exception at $Q_1$ when $EARN_t \geq 0$, Panel A shows that the coefficient on SI is negative and significant. This indicates that the power of the model is enhanced by the inclusion of special items.

*4.3 Risk metrics*

In Table 4 we report the properties of our risk metrics that use the quantile forecasts as inputs. We also include descriptive statistics for the median forecast $Q_{50}$ to provide an indication of the distribution of the predicted central moment of future earnings. We estimate our risk metrics on a recursive basis, expanding the dataset used in quantile regression annually. We require a minimum of five years' data in estimation, restricting the out-of-sample forecast period to 1993−2009. Our first regression consists of 7,178 observations and includes the years 1988−1992. Panel A contains descriptive statistics and Panel B reports correlations between the risk metrics.

The descriptive statistics reveal that IQR is positively skewed and displays considerable variation across the sample. Consistent with the asymmetric timeliness of earnings, SKEW is negative except for a small proportion of cases. However there is wide variation in the predicted value of SKEW. The distribution of KURT indicates that the predicted tail density is generally higher than 1.23, the value expected under the normal distribution. The descriptive statistics for UP and DOWN indicate that this is due to higher left tail density (DOWN) and lower right tail density (UP) compared to a normal distribution (in which case UP and DOWN would equal 0.615).

The correlations in Table 4 Panel B indicate that IQR, SKEW and KURT are not highly correlated, despite IQR being in the denominator of the other two metrics. The relatively high correlations of SKEW with UP and DOWN are not surprising because the quantile ranges over which UP and DOWN are computed partially overlap with the numerator of SKEW. For this reason, when using UP and DOWN as risk metrics in subsequent tests we drop SKEW from the relevant models. The high correlation between KURT and DOWN compared to the correlation between KURT and UP appears important, keeping in mind that $\mathrm{KURT} \equiv \mathrm{UP} + \mathrm{DOWN}$. It indicates that DOWN is more variable than UP, and hence lower tail behavior appears to be driving variation in KURT more. As a consequence DOWN appears a plausible alternative measure of downside risk to the degree of negative SKEW.

Figure 2 provides graphical insights into how risk forecasts depend on OCF and ACC and the sign of EARN. In each panel we plot the median estimated risk metrics for 5x5 sequentially sorted portfolios, sorted first within industry, second on OCF and then on ACC. Portfolios are formed separately for firms with EARN $\geq 0$ (left-hand panels in each row) and EARN $< 0$ (right-hand panels), conditional on having at least 25 observations within each industry-year-profit and industry-year-loss group. The histograms in Figure 2 take account of the actual distributions and dependence between OCF and ACC in the sample data and of how these properties vary between profit and loss firms. Graphs for each risk measure are plotted on the same scale to facilitate comparison between profit and loss firms.

The first row of Figure 2 confirms previous results showing that profit firms have generally lower earnings uncertainty than loss firms. However, within the set of profit firms, there is significant variation in IQR due to positive dependence on ACC – on average IQR increases monotonically by 28% from low to high ACC portfolios, holding OCF constant. This indicates that greater levels of accruals contribute to greater uncertainty for profit firms. For loss firms, overall levels of IQR are on average approximately twice as high, and the pattern of dependence between IQR and ACC reverses, with high ACC firms having approximately half the earnings uncertainty compared to low ACC firms, holding OCF constant. Another notable insight is that IQR falls by approximately 50% as OCF increases. Taken together these results indicate that future earnings uncertainty depends crucially on the mix of combination of cash flow and accruals underlying reported losses or profits.

The second row of Figure 2 reveals that the degree of negative SKEW in future earnings is higher for profit than for loss firms and increases monotonically with ACC in both cases. Noting that the denominator of SKEW is IQR which hence affects the magnitude of SKEW, the higher levels of SKEW for profit firms and the negative dependence between SKEW and ACC for loss firms are not surprising. However, for profit firms the monotonic negative dependence between SKEW and ACC cannot be attributed to IQR, because this increases with ACC.

The third row of Figure 2 also reveals patterns of significant and monotonic dependence between KURT and ACC for both profit and loss firms. Again the patterns reverse between profit and loss firms. Similar to SKEW, we note that IQR is in the denominator of KURT and hence the relation between KURT and ACC can be at least partially attributed to the relation between IQR and ACC documented in the first row of Figure 2. However, further insights into the behavior of KURT can be gained by examining the decompositions of KURT into UP and DOWN, both of which again contain IQR in the denominator.

The fourth and fifth rows of Figure 2 reveal that the behavior of UP and DOWN is very different for both profit and loss firms. For profit firms UP displays a sharply negative and monotonic relation with ACC, while DOWN is relatively insensitive to ACC. There is also evidence that UP increases with OCF, especially at lower levels of ACC, and there is mild negative dependence between DOWN and OCF. The dependence between UP and ACC for profit firms suggests that firms achieving profits through ACC have lower right-tail density in future earnings, consistent with the lower persistence of accruals which reverse in the future. In contrast, for loss firms UP is relatively insensitive to ACC and has a mild negative dependence on OCF. DOWN is now sharply and monotonically positively related to ACC and exhibits a modest negative relation with OCF.

## 5. Are quantile-based risk metrics relevant to equity and bond markets?

We now investigate the relevance of higher moments in future earnings for the equity and corporate bond markets. We examine whether our risk metrics explain analysts' equity risk ratings, credit risk ratings and corporate bond yield spreads, and whether they predict future equity return volatility. We also study whether our risk metrics are associated with analyst-based measures of future earnings uncertainty, captured by analysts' exclusions when defining pro forma earnings and the absolute value of analysts' forecast errors.

We estimate the following equations:

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_2 SKEW + \gamma_3 KURT + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon \qquad (8)$$

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_{31} UP + \gamma_{32} DOWN + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon \qquad (9)$$

Outcome is defined in successive tests as equity risk rating (Equity_Rating), credit rating (Credit_Rating), an indicator variable capturing whether a credit rating is speculative grade or investment grade (Speculative_Grade), future equity return volatility (PostVol), corporate bond yield spread (Spread), the absolute value of analysts' earnings exclusions when defining pro forma earnings (|Exclusions|) and the absolute value of analysts' forecast errors (|FE|). In different specifications of (8) we combine SKEW and KURT along with IQR and the predicted median of future earnings $Q_{50}$, obtained from the median regression. The inclusion of $Q_{50}$ controls for the projected level of future earnings around which the distribution of earnings is located. In different specifications of (9) we combine UP and DOWN together with IQR and $Q_{50}$. We include year fixed effects in all of the regressions, which correspond to 17 intercept dummies for the period 1993 - 2009. We suppress the reporting of intercepts and fixed effects in order to conserve space. All variables are defined in Appendix A.[17]

*5.1 Equity risk ratings and future return volatility*

Prior research establishes that equity analysts produce useful information in the form of earnings forecasts and forecast revisions, target prices, stock recommendations and text-based discussions (see Joos et al. (2014) for a recent review of the literature). Recent research also shows that analysts produce information useful in assessing equity risk. Lui et al. (2007) take a first step in examining the role of financial analysts as providers of information concerning investment risk. They find that analysts' risk ratings are determined by commonly used risk proxies including idiosyncratic risk, size, book-to-market, leverage, accounting losses and accounting quality. They also show that analysts' risk ratings are informative about predicting future return volatility. A recent paper by Joos et al. (2014) complements Lui et al. (2007) by showing that the spread in target price estimates between the analysts' bull and bear scenarios is a useful risk proxy, being significantly associated with risk-related firm fundamentals and predicting future long-run return outcomes.

We build on this prior literature on equity risk in two ways. First, we examine whether our earnings risk metrics help to explain equity risk ratings and future return volatility. Second, we test whether our risk metrics contribute incrementally to explaining those outcomes, beyond a broad set of empirical risk proxies used in the literature. This analysis has the potential to indicate the extent to which the risk assessment process used by both analysts and investors is based on deeper fundamental attributes than those identified in the previous literature.

We regress each Outcome variable on our earnings risk metrics and control variables based on equations (8) and (9). In the case of equity risk ratings, Outcome is defined as Equity_Rating and is

[17] To mitigate possible effects of outliers, we winsorize all of our control variables at the top and bottom 1% of their distributions in each year. EarnVol, TRANSP and EarnQual are not winsorized because they are computed based on trimmed data. We also winsorize quantile-based risk metrics to avoid the small denominator problem, when IQR is close to zero.

measured on a scale from 1 (low risk) to 4 (speculative risk). The regression is estimated using ordered logit regression and standard errors are clustered by firm. In the case of future return volatility, Outcome is defined as the realized volatility of daily stock returns over the 12 months after our risk metrics are formed (PostVol). The regression is estimated using OLS and standard errors are clustered by firm and year. The included control variables proposed in Lui et al. (2007) are defined in Appendix A. In all regressions, we include year fixed effects. We winsorize PostVol in each year at the top and bottom 1% of its distribution, similar to Lui et al. (2007).

Results from estimating equation (8) are reported in Table 5 Panel A. Columns (1)−(3) contain results for equity risk ratings and columns (4)−(6) show results for future return volatility. Results for equity risk ratings indicate that analysts view firms with higher IQR as riskier. This finding is robust to the inclusion of the full array of risk proxies from the prior literature, suggesting that IQR captures incremental information about fundamental risk, which is understood by analysts. After controlling for predicted earnings ($Q_{50}$) and combinations of other control variables, SKEW is never significant in explaining risk ratings. However the coefficient on KURT is reliably positive in all specifications. Overall, the results indicate that while dispersion in the center of the distribution captured by IQR is a significant determinant of equity risk ratings, greater density in the tails of the distribution is viewed as an additional driver of equity risk by analysts.

Results for future return volatility indicate that IQR is strongly associated with PostVol, even after controlling for alternative risk proxies including return volatility in the prior period. KURT is also a significant predictor of future volatility before controlling for other risk proxies, but its significance disappears after controls are introduced. This suggests that one or more of the risk proxy controls from the prior literature are capturing extreme returns realizations, although none is explicitly motivated from this perspective. In summary, results indicate that our proxy for earnings uncertainty IQR contains independent information useful in predicting return volatility over the next 12 months.

In Panel B we report results when KURT is decomposed into UP and DOWN as in equation (9). The analyst results show that the significance of KURT is largely attributable to DOWN, the coefficient being highly significant in all specifications. In contrast, the coefficient on UP is only marginally significant in column (1), although it gains its significance when control variables are introduced in columns (2) and (3). The results suggest that when adverse predicted earnings outcomes are large, stocks are regarded by analysts as riskier. Viewed another way, it appears that analysts are able to identify stocks with higher downside risk and they penalize such stocks in their risk ratings. Results for future return volatility show that both DOWN and UP become insignificant after including other risk proxies. However, when they are used alone, DOWN is related positively and UP negatively to future return volatility.

In summary, both equity risk ratings and future volatility appear to reflect unique information captured by our earnings risk metrics. This information is incremental to a wide range of characteristics-based risk proxies considered by the prior literature. One implication of our results is that our earnings risk metrics, in particular IQR, KURT and its components UP and DOWN, will be useful to investors and other decision makers in estimating equity risk, especially when equity analyst risk ratings are unavailable.

*5.2 Credit ratings*

A number of studies have shown that credit ratings are related to a number of common risk proxies, including fundamental accounting attributes such as earnings quality and earnings volatility (Francis et al. 2005; Cheng and Subramanyam 2008). We consider the relation between our risk proxies and long-term issuer credit ratings provided by Standard & Poor's, which range from AAA to D (debt in default). Similar to Ashbaugh et al. (2006), we recode the 22 ratings classifications employed by Standard and Poor's into seven categories (Credit_Rating)[18] and subsequently in the form of an indicator variable capturing whether a rating indicates investment grade (BBB or better) or speculative grade (Speculative_Grade).

[18] Credit_rating takes the value of 1 if S&P rating is AAA, the value of 2 if S&P rating is AA+, AA, or AA-, the value of 3 if S&P rating is A+, A, or A-, the value of 4 if S&P rating is BBB+, BBB, or BBB-, the value of 5 if S&P rating is BB+, BB, or BB-, the value of 6 if S&P rating is B+, B, or B- and the value of 7 if S&P rating is CCC+, CCC, CCC-, CC, C, D or SD.

Similar to tests based on equity risk ratings, we test whether our earnings risk metrics explain credit ratings and whether they contribute information incrementally beyond a set of market-based and accounting-based credit risk proxies considered in the prior literature based on Cheng and Subramanyam (2008) and Francis et al. (2005). These proxies are described in detail in Appendix A. We estimate equations (8) and (9) first when Outcome is defined as Speculative_Grade, coded as 1 for speculative grade and 0 for investment grade; and second when Outcome is defined as Credit_Rating, taking values from 1 (low risk) to 7 (high risk). In both cases positive coefficients indicate that a risk metric is positively associated with default risk assessed by credit analysts. Regressions are estimated using logit regression for Speculative_Grade and ordered logit regression for Credit_Rating. All regressions include year fixed-effects and two-way clustered standard errors.

Results for the credit ratings tests are reported in Table 6. Consider first the Speculative_Grade models (Panels A and B). When we estimate equation (8), the earnings uncertainty proxy IQR is consistently significant in predicting whether a firm is rated as investment grade or speculative grade, with IQR predicting firms to have higher credit risk. The inclusion of a battery of control variables from the prior literature does not affect the significance of IQR, indicating that IQR contains incremental information for explaining credit ratings. SKEW is significant in predicting Speculative_Grade only after controlling for common risk proxies. Similar to Equity_Rating results, KURT is a significant predictor of Speculative_Grade in column (1). However, it is only marginally significant in column (2) and it is rendered insignificant in the final specification (column 3) when earnings volatility (EarnVol), earnings quality (EarnQual) and a loss proxy (Loss) are added as risk factors. Finally, when we estimate equation (9) using Speculative_Grade, IQR remains robust under all specifications and DOWN is seen to be a highly significant predictor of Speculative_Grade, even after controlling for standard risk proxies including current losses. It appears from these results that the magnitude of potential downside losses, as captured by either SKEW (equation (8)) or DOWN (equation (9)) is an economic attribute that complements IQR in discriminating between investment grade and speculative grade credit ratings.

Tests involving Credit_Rating as the dependent variable are more demanding because the challenge is to discriminate between credit ratings within investment grade and speculative grade classifications using a linear logit function and the same variables. Panel A in Table 6 presents results for estimates of equation (8) and Panel B for equation (9). Despite the demanding nature of these tests the significance of IQR is quite robust. Only when we include EarnVol, EarnQual and Loss is IQR rendered insignificant. This is partly due to the severe sample attrition resulting from the time-series data demands of these risk proxies.[19] Further, untabulated correlations between our IQR metric and these risk proxies are relatively high. SKEW is again important in Panel A, indicating that conditional skewness is significant in discriminating between credit ratings. Similar to the Speculative_Grade results, DOWN is a significant determinant of credit ratings in Panel B, even in this more demanding test.

In summary, credit analysts appear to reflect information captured by IQR, SKEW and DOWN in their credit risk ratings. This information is incrementally relevant after controlling for a wide range of characteristics-based risk proxies considered in the prior literature. One implication of our findings is that our earnings risk metrics, especially IQR, SKEW and DOWN, could be useful to investors and other decision makers in estimating credit risk when credit ratings are unavailable.

*5.3 Corporate bond yield spreads*

We also explore the ability of our fundamental risk metrics to explain corporate bond yields. Similar to credit ratings, bond yields capture the market's assessment of the probability of default. Therefore, this test examines whether risk in future earnings captures information about default risk that is perceived by the market. Because we control for credit ratings, which we showed earlier to also depend on our risk metrics, these tests can be viewed as a test of whether credit ratings fully incorporate the information in our risk metrics that is relevant for market pricing.

[19] Untabulated results reveal that projected losses ($Q_{50} < 0$) comprise 17.2% of the sample. In this sub-sample, earnings uncertainty is irrelevant in all tests except those based on analysts' exclusions. Therefore the inclusion of projected loss firms biases coefficients on IQR towards zero in full sample tests.

The relation between equity risk and credit risk has already been demonstrated in prior literature. Campbell and Taksler (2003) provide evidence that equity return volatility is a significant determinant of corporate bond yield spreads, after controlling for other issue- and issuer- specific factors, such as credit ratings. While equity volatility reflects uncertainty in future cash flow or asset values, our earnings risk metrics may capture new information on risk in future payoffs and possible future default. Furthermore, higher moments of earnings, such as skewness and kurtosis, are likely to influence incrementally the market's assessment of default.

Following prior literature, we restrict our sample to fixed-rate US dollar corporate bonds that are non-puttable and non-convertible. In addition, we exclude bonds with odd frequency of coupon payments[20] (Elton et al. 2001). Callable bonds constitute a large fraction of the TRACE sample and therefore we include them in our analysis but control for callability, i.e. an indicator variable equal to one if a bond is callable and zero otherwise (Bao et al. 2011). To obtain monthly bond yields from the TRACE database, we follow Bessembinder et al. (2009). After eliminating cancelled, corrected, and commission trades from the data, we eliminate all trades under $100,000 and rely on the last yield posted in the database each month. The sample formation is shown in Panel A of Table 7.

We define Outcome in equations (8) and (9) as the treasury spread (Spread), which is calculated as the difference between the yield to maturity on a bond and the contemporaneous yield on a benchmark US treasury. As in Campbell and Taksler (2003), we use the CRSP Fixed Term indices to obtain monthly yields on bonds of 1, 2, 5, 7, 10, 20 and 30 years to maturity. Therefore we assume that each bond transaction occurs at the end of each month. Consistent with Campbell and Taksler (2003) we delete the top and bottom 1% of the spread variable each month, in order to reduce the effects of potential data errors in the TRACE database.[21] We ensure that all accounting data are known to the market when a bond purchase or sale takes place by allowing a three months lag after the fiscal year-end.

---

[20] In particular, we delete observations with frequency of coupon payments equal to -1 or 99. The remaining bond issues of our sample have frequency of coupon payments equal to 0, 1, 2, 4 or 12.
[21] The results are robust if we winsorize Spread instead.

The choice of risk proxy controls, other firm-level controls and macroeconomic control variables is based largely on Elton et al. (2001) and Campbell and Taksler (2003). Control variables are defined in Appendix A. We show results before and after controlling for the S&P issuer credit rating (Credit_Rating). We include year and issuer fixed effects and we cluster standard errors by firm. Issuer fixed effects are important to control for issuer influences on yields, because firms have multiple bonds in the dataset.

Panel B of Table 7 shows results from estimating equation (8). Results confirm our earlier findings that IQR is a highly significant determinant of economic outcomes, even after controlling for a comprehensive range of risk proxies and other factors predicted to be important based on the prior literature, including credit ratings. KURT is similarly important and captures incremental information relevant to bond pricing and again after controlling for credit rating. When we decompose KURT into UP and DOWN in Panel C, UP appears to be significant in explaining Spread, whereas DOWN is less consistently significant especially after including Credit_Rating. This finding suggests that credit analysts fully incorporate relevant information in DOWN into their credit ratings.

Overall, these results present new evidence that our earnings risk metrics, especially IQR, KURT and UP explain corporate credit spreads incrementally to the determinants previously identified in the literature. While our earlier results show that IQR is a significant determinant of Credit_Rating, we note that the results in Table 7 suggest that IQR contains relevant information beyond credit ratings for explaining the credit spread.

*5.4 Sell-side Analysts' Decisions*

Our final set of tests examines the relations between our risk metrics and observable outcomes associated with sell-side analysts' decisions. We focus on two aspects of analysts' forecast activity studied in the prior literature where earnings risk can be important. First, we study earnings forecast accuracy, which is often used as a proxy for information uncertainty (e.g. Barron et al. 1998; Horton et al. 2013). Our results provide new insights into how the distributional properties of future earnings affect earnings forecast accuracy. Second, we study the pro forma earnings numbers reported by analysts in the I/B/E/S database. The use of pro forma earnings is often motivated by suggestions that more permanent earnings constructs are useful to investors for forecasting cash flow and firm value (Bradshaw and Sloan 2002). We focus on the component of as-reported earnings that is excluded from analysts' pro forma numbers, in order to develop new understanding of how pro forma earnings adjustments are related to the distributional properties of future earnings.

### *5.4.1 Analysts' forecast errors*

A large body of research has focused on the determinants of analysts' forecast errors. When information uncertainty is high, earnings are more difficult to predict and the accuracy of analysts' forecasts is expected to be lower. The degree of information uncertainty faced by analysts depends on the quality of a firm's general information environment and on the transparency and timeliness with which news is reported in earnings (Lim 2001, p.377-8). We contribute to the literature by testing whether our fundamentals-based risk measures explain incrementally forecast accuracy, after controlling for proxies of information uncertainty (Lim 2001), GAAP informativeness (Lougee and Marquardt 2004) and other known determinants of forecast accuracy (Horton et al. 2013; Tan et al. 2011). We define Outcome as the absolute forecast error (|FE|) measured in relation to pro forma actual earnings reported in I/B/E/S and we estimate equations (8) and (9) before and after including other controls. FE is winsorized at the top and bottom 1% of its distribution in each year.

Results in Table 8 indicate that IQR, SKEW, UP and DOWN contain new information explaining forecast accuracy, beyond historical earnings volatility (EarnVol) and other known determinants of forecast accuracy. The association between IQR and one-year-ahead absolute forecast errors is significantly positive, indicating that forecast accuracy is lower for firms with higher estimated IQR. Additionally, results in Panel A indicate that SKEW is positively associated with absolute forecast errors, while Panel B indicates that, after including control variables, the coefficient on UP is significantly positive and on DOWN is significantly negative. Taken together, these results suggest the following insights: (i) IQR captures new information about uncertainty in future earnings and is an important determinant of forecast accuracy; and (ii) SKEW, UP and DOWN capture new information concerning asymmetry and extreme outcomes in future earnings beyond the information contained in known determinants of forecast accuracy. Holding earnings uncertainty constant, our evidence is consistent with analysts being more efficient in anticipating the realization of unfavorable earnings outcomes relative to favorable earnings outcomes.

*5.4.2 Analysts' pro forma earnings*

The use of non-GAAP earnings constructs is widespread and reflects both supply-side and demand-side pressures. On the supply-side, reporting firms may choose to emphasize alternative measures in performance reporting (e.g. Vincent et al. 1999; Bradshaw and Sloan 2002; Lougee and Marquardt 2004; Serafeim 2011), especially when the informativeness of GAAP earnings is low or when firms wish to promote strategic interests (Doyle et al. 2003, 2013; Lougee and Marquardt 2004). On the demand-side, analysts often choose to forecast "Street earnings" (pro forma earnings) that typically exclude transitory items which are less value relevant (Bradshaw and Sloan 2002).

The exclusions of analysts from their earnings definitions can be motivated in the following ways. If analysts have incentives to provide informative earnings forecasts, they will tend to use earnings constructs that exclude relatively uninformative GAAP items. If, on the other hand, analysts have incentives to provide accurate earnings forecasts, they will seek to exclude earnings components which are more difficult to predict. In both cases, the nature of excluded items will be similar (e.g. special items, write-offs). In our last set of tests we examine links between ex ante risk properties of GAAP earnings and pro forma exclusions. Although we do not attempt to distinguish between competing motivations underlying analysts' exclusions, our results are potentially useful in developing a richer understanding of the origins of the use of pro forma earnings numbers by analysts.

We define Outcome as the absolute value of the difference between pro forma earnings and GAAP earnings (|Exclusions|). Consistent with Doyle et al. (2003), we winsorize Exclusions at the top and bottom 1% of its distribution in each year. We estimate equations (8) and (9) before and after including proxies for information uncertainty (Lim 2001), GAAP informativeness (Lougee and Marquardt 2004), and other known determinants of forecast accuracy (Horton et al. 2013; Tan et al. 2011).

Results in Table 8 Panel A indicate that the magnitude of one-year-ahead exclusions is strongly and positively associated with IQR, even after including EarnVol and the full array of our control variables. KURT (Panel A) and its components UP and DOWN (Panel B) are also positively related to future exclusions before and after including other risk controls. In summary, these results suggest that pro forma exclusions from GAAP numbers are predictable using our risk metrics. Our evidence is consistent with analysts' exclusions being higher when GAAP earnings informativeness is low, but also when GAAP earnings are more difficult to predict.

## 6. Robustness/Additional Tests

### *6.1 Controlling for the level of earnings*

The modelling of the earnings distribution mechanically causes any risk measures derived from quantile estimates to be functions of the predictor variables. While SKEW, KURT, UP and DOWN are non-linear functions of the predictor variables, IQR continues to be a linear function (which we explicitly report in the final column of Table 3). Prior research links profitability to risk exposure. To ensure that our risk measures do not capture EARN, we repeat all of our analysis (untabulated) by replacing $Q_{50}$ with EARN. Our main inferences remain the same.

### *6.2 Outliers treatment of continuous dependent variables*

Consistent with the prior literature, we have dealt with outliers by winsorizing or trimming the 1%, 99% of continuous dependent variables. We did so to mitigate the impact of potential data errors on our regressions. If however the observations in the top and bottom 1% of the dependent variables' distribution were free from error, winsorization would artificially reduce total variance and inflate the t-statistics in our regressions. To test the robustness of our results, we repeat our tests in Tables 5−8, without winsorizing the raw value of PostVol, Spread, Exclusions and FE. Results with respect to PostVol, |Exclusions| and |FE| are qualitatively the same. Bond tests are also robust in relation to IQR, but less so with respect to kurtosis and its components. This is perhaps not surprising given that kurtosis captures tail risk and therefore it is particularly sensitive to errors in the tails of the *Outcome* variable.[22]

*6.3 Model specification*

In modelling the earnings distribution, we have disaggregated earnings into accruals, cash flow and special items and we have also conditioned model coefficients on the presence of losses. To obtain a better understanding of the importance of different model elements for constructing valid risk metrics we perform two analyses. First we conduct pairwise comparisons of forecast accuracy between our main forecasting model in equation (7) and five restricted versions of equation (7);[23] [24] and (ii) we repeat all analyses in Tables 5−8 using risk metrics derived from each restricted model in turn.

Untabulated results show that our model results in significantly lower forecast errors than the restricted models considered across all quantiles used in risk metric estimation. This suggests that a relatively more complex model, based on accruals, cash flow, special items and losses, maximizes forecast accuracy. Risk prediction results show that different elements of the model are important for different risk outcomes. For example, if the focus is on predicting bond yield spreads, then allowing for losses in the forecasting model is necessary and sufficient for the relevance of IQR. If, on the other hand, one is interested in predicting both bond spreads and future volatility, allowing for losses and

---

[22] The maximum value of Spread without trimming is 668%, whereas it reduces to 128% after deleting outliers.
[23] The five restricted models tested are as follows:

*Model 1:* $EARN_{t+1} = \alpha_0 + \alpha_1 EARN_t + v_{t+1}$

*Model 2:* $EARN_{t+1} = \alpha_0 + \alpha_1 ACC_t + a_2 OCF_t + v_{t+1}$

*Model 3:* $EARN_{t+1} = \alpha_0 + \alpha_1 ACC_t + a_2 OCF_t + a_3 SI_t + v_{t+1}$

*Model 4:* $EARN_{t+1} = \alpha_{01} d^+ + \alpha_{02} d^- + \alpha_1 d^+ EARN_t + \alpha_2 d^- EARN_t + v_{t+1}$

*Model 5:* $EARN_{t+1} = \alpha_{01} d^+ + \alpha_{02} d^- + \alpha_1 d^+ ACC_t + \alpha_2 d^- ACC_t + \alpha_3 d^+ OCF_t + \alpha_4 d^- OCF_t + v_{t+1}$

[24] We evaluate the forecast accuracy (FA) of the various forecasting models with reference to the mean loss appropriate to each forecast. Loss is defined as the weighted absolute forecast error corresponding to each quantile, consistent with the minimization problem of the quantile regression defined in expression (1) above.

decomposing earnings into accruals and cash flow are both necessary. Enhancing the model with special items is useful for predicting bond yield spreads, not only with IQR but also with KURT.

*6.4 Definition of risk metrics*

In defining our tail risk metrics, we use the quantile estimates $Q_{12.5}$, $Q_{37.5}$, $Q_{25}$, $Q_{50}$, $Q_{62.5}$, $Q_{75}$ and $Q_{87.5}$, following prior literature. In doing so, we avoid relying on the extreme tails of the distribution, estimates of which are noisier as one can see from the decreasing pseudo $R^2$ statistics of the quantile regressions in the extreme tails. Nevertheless we investigate whether alternative measures of kurtosis using extreme tails are significant in risk prediction results. We re-define KURT, UP and DOWN, as follows:

$$KURT_i' = [(Q_{i95} - Q_{i62.5}) + (Q_{i37.5} - Q_{i5})]/IQR_i \tag{10}$$

$$UP_i' = (Q_{i95} - Q_{i62.5})/IQR_i \tag{11}$$

$$DOWN_i' = (Q_{i37.5} - Q_{i5})/IQR_i. \tag{12}$$

The definitions of IQR and SKEW remain unchanged.[25]

Results with respect to IQR remain unaffected under all risk prediction tests. Our inferences also remain unchanged with respect to kurtosis and its upside and downside components in tests based on equity risk ratings, analysts' exclusions and earnings forecast accuracy. These results suggest that extreme tail risk is relevant for equity analysts' exclusions from their earnings definition, forecast accuracy, and equity risk ratings. However, when testing credit ratings, speculative grade, bond spreads and future return volatility, results for kurtosis and its components are sensitive to the inclusion of ten percent of observations with the lowest earnings uncertainty (IQR). For these cases, extreme downside risk appears to be irrelevant or to be weighted negatively. These results could imply that credit market analysts fail to understand or discount extreme tail risk when IQR is very low. Alternatively it is possible that extreme quantile estimates are more sensitive to measurement error in cases of low earnings uncertainty, or that they are subject to a small denominator measurement issue (IQR being close to zero).

---

[25] Defining SKEW as $[(Q_{95} - Q_{50}) - (Q_{50} - Q_5)]/[(Q_{95} - Q_5)$ does not affect inferences.

## 7. Conclusion

We show that a parsimonious forecasting model based only on current accruals, cash flow, special items and a loss indicator is capable of forecasting quantiles of the distribution of future earnings. Our results reveal interesting dependencies between the conditional shape of the distribution of future earnings and current earnings components. These dependencies are hidden behind the conditional mean effects documented in the existing earnings persistence literature.

We introduce risk metrics based on the forecasted distribution capturing dispersion, skewness and kurtosis in future earnings. We show evidence that quantile-based risk forecasts are associated with equity and credit risk ratings, future return volatility, credit spreads and analyst-based measures of earnings uncertainty. Our results hold even after controlling for other risk proxies, suggesting that analysts' assessments of fundamental risk go beyond risk metrics commonly employed in the prior literature. They also suggest that predictions of market outcomes are enhanced by information contained in our risk metrics. In short, our earnings risk metrics capture new risk information.

Future research could ascertain whether the predictive ability of our deliberately parsimonious earnings quantile model can be further enhanced by the inclusion of further fundamental and non-financial predictor variables. It could also address the relative power of alternative approaches to forecasting dispersion and higher moments of the distribution of future earnings.

Appendix A: Definitions of variables[26]

| | |
|---|---|
| *Main Variables* | |
| EARN | Income before extraordinary items (*IBC*), scaled by average total assets. |
| OCF | Operating cash flows (*OANCF*) minus extraordinary items and discontinued operations (*XIDOC*), scaled by average total assets. |
| ACC | Accruals defined as EARN minus OCF. |
| SI | Special items (*SPI*), scaled by average total assets. |
| $d^{+}(d^{-})$ | An indicator variable equal to one if $EARN_t \geq 0$ ($EARN_t < 0$). |
| *Risk Measures* | |
| IQR | The 0.5 interquartile range calculated as $(Q_{75} - Q_{25})$. |
| SKEW | Skewness calculated as $[(Q_{75} - Q_{50}) - (Q_{50} - Q_{25})]$/IQR. |
| KURT | Kurtosis calculated as $[(Q_{87.5} - Q_{62.5}) + (Q_{37.5} - Q_{12.5})]$/IQR. |
| UP | The upside component of KURT, calculated as $(Q_{87.5} - Q_{62.5})$/IQR. |
| DOWN | The downside component of KURT, calculated as $(Q_{37.5} - Q_{12.5})]$/IQR. |
| *Equity Ratings and Future Volatility* | |
| Equity_Rating | A discrete variable taking the values 1, 2, 3, 4 (low risk, medium risk, high risk and speculative). |
| PostVol | The standard deviation of a stock's daily return for a period of 12 months, starting three months after the firm's fiscal year-end. A minimum of 11 calendar months of daily return observations is required. In our regression tests, we use the logarithm of PostVol following Lui et al. (2007). |
| PreVol | The standard deviation of a stock's daily return for a period of 12 months ending at fiscal year-end. A minimum of 60 daily return observations is required. In our regression tests, we use the logarithm of PreVol following Lui et al. (2007). |
| Beta | The market beta, estimated by regressing firm-level daily stock returns on the value-weighted CRSP market index over a window of 12 months ending at fiscal year-end. A minimum of 60 daily return observations is required (Lui et al. 2007). |
| IVol | The idiosyncratic volatility of a stock, i.e. the portion of total stock return volatility unexplained by the market. IVol is calculated as the standard deviation of the residuals obtained from the regression used to calculate the market beta. |
| Illiquidity | The daily ratio of absolute stock return to its dollar volume, averaged over the fiscal year (Amihud 2002). |
| MV | The logarithm of the market value of equity three months after fiscal year-end. |
| NegBV | A dummy variable that takes the value of 1 if book value of equity is negative. |

[26] Compustat item labels (*XFP* names) for accounting variables are in parentheses.

| | |
|---|---|
| B/M | Book value of common equity (*CEQ*) divided by market value of equity at fiscal year-end. |
| D/E | Long-term debt (*DLTT*) plus debt in current liabilities (*DLC*) divided by book value of common equity (*CEQ*). |
| IPO | A dummy variable that takes the value of 1 if a firm had its initial public offering within two years prior to the fiscal year end. |
| EarnVol | The time-series standard deviation of earnings before extraordinary items (*IBC*) scaled by average total assets, computed recursively using at least 5 years including the current year. |
| EarnQual | Earnings quality computed as the standard deviation of the residuals from the regression $ACC_t = \delta_0 + \delta_1 \Delta REV_t + \delta_2 GPPE_t + \delta_3 CF_{t-1} + \delta_4 CF_t + \delta_5 CF_{t+1} + \varepsilon_t$, over a period of 5 years (Francis et al. 2005). The regression is estimated by industry and year, conditional on having a minimum of 10 observations within industry-year group. All data used in the estimation are available 3 months after fiscal year-end and. GPPE is gross property, plant and equipment (*PPEGT*) and ΔREV is the change in revenue (*SALE*), both of which are scaled by average total assets. |
| Loss | A dummy variable that takes the value of 1 if current earnings is negative and 0 otherwise. |
| <u>*Credit Ratings*</u> | |
| Investment_Grade | A dummy variable equal to one for speculative grade bonds and zero for investment grade bonds. |
| Credit_Rating | The S&P'S long-term issuer credit rating, taking values from 1 (low risk) to 7 (high risk). |
| PreVol | The standard deviation of a stock's daily return for a period of 12 months ending at fiscal year-end. A minimum of 60 daily return observations is required. In our regression tests, we use the logarithm of PreVol following Lui et al. (2007). |
| Beta | The market beta, estimated by regressing firm-level daily stock returns on the value-weighted CRSP market index over a window of 12 months ending at fiscal year-end. A minimum of 60 daily return observations is required (Lui et al. 2007). |
| RET | Buy and hold raw returns over the fiscal year. |
| Mn_PRC | The mean daily closing price per share over the fiscal year. |
| MV | The logarithm of the market value of equity three months after fiscal year-end. |
| B/M | Book value of common equity (*CEQ*) at fiscal year-end divided by market value of equity at fiscal year-end. |
| D/A | Long-term debt (*DLTT*) plus short-term debt (*DLC*) to total assets (*AT*) at the end of the year. |
| Cover | Pre-tax interest coverage defined as the ratio of operating income after depreciation (*OIADP*) plus interest expense (*XINT*) to interest expense (*XINT*) (Campbell and Taksler 2003). |
| Intan | Intangibles' intensity measured as research and development expense (*XRD*) plus advertising expense (*XAD*) scaled by total assets at the end of the year. |

| | |
|---|---|
| ΔEQ | A dummy variable equal to 1 if change in shareholder equity (*SSTK*) during the year is greater than zero, zero otherwise. |
| Transp | Financial Transparency, measured as negative one times the squared residual from the regression ARET = $\beta_0 + \beta_1$NIBX + $\beta_2$Loss + $\beta_3$NIBX*Loss + $\beta_4\Delta$NIBX + ε, where ARET is the market adjusted return over the fiscal year, NIBX is income before extraordinary items (*IBC*) scaled by beginning of year market value of equity and ΔNIBX is the change in NIBX. The regression is estimated by industry and year conditional on having at least 10 observations within industry-year group (Cheng and Subramanyam 2008). |
| EarnVol | The time-series standard deviation of earnings before extraordinary items (*IBC*) scaled by average total assets, computed recursively using at least 5 years including the current year. |
| \|ABACC\| | Absolute value of abnormal accruals. Abnormal accruals are estimated using the cross-sectional Jones model and comprise the residuals from the following intercept-suppressed regression: $ACC_t = \alpha_0(1/AVTA_t) + \alpha_1 GPPE_t + \alpha_2\ \Delta REV_t$, where AVTA is average total assets, GPPE is gross property, plant and equipment (*PPEGT*) scaled by AVTA and ΔREV is the change in revenue (*SALE*) scaled by AVTA. The regression is estimated by industry and year conditional on having at least 10 observations within each industry-year group (Cheng and Subramanyam 2008). |
| Loss | A dummy variable that takes the value of 1 if current earnings is negative and 0 otherwise. |
| *Bond Yields* | |
| Spread | The treasury spread, i.e. the difference between the yield to maturity on each bond and the yield on a benchmark US treasury in a particular month (in percentage). |
| Issue_size | Issue size of each bond. |
| Coupon | The coupon rate (in percentage). |
| YtM | Years to maturity. |
| Call | A dummy that is set to 1 if a bond is callable and zero otherwise. |
| T-Note | The closest benchmark Treasury rate (in percentage). |
| Term_Slope | The slope of the term structure, calculated as the difference between the 10- and 2-year Treasury rates (in percentage). |
| Euro_TBill | The difference between the 30-day Eurodollar and Treasury yield (in percentage). |
| Prevol | The standard deviation of a stock’s daily return for a period of 12 months ending at fiscal year-end. A minimum of 60 daily return observations is required. In our regression tests, we use the logarithm of PreVol following Lui et al. (2007). |
| Mn_RET | The mean of firm-level daily stock returns over the fiscal year. |
| Cover | Pre-tax interest coverage defined as the ratio of operating income after depreciation (*OIADP*) plus interest expense (*XINT*) to interest expense (*XINT*) (Campbell and Taksler 2003). |
| Margin | Profit margin defined as operating income before depreciation (*OIBDP*) to sales (*SALE*). |

| | |
|---|---|
| LD/A | Long-term debt (*DLTT*) to total assets (*AT*) at the end of the year. |
| D/Cap | Total debt to capitalization computed as long-term debt (*DLTT*) plus debt in current liabilities (*DLC*) plus average short-term borrowings (*BAST*) to total liabilities (*LT*) plus market value of equity (from CRSP) at fiscal year-end. |
| Credit_Rating | The S&P'S long-term issuer credit rating, taking values from 1 (low risk) to 7 (high risk). |
| <u>*Analysts*</u> | |
| $\text{Forecast}^{\text{IBES}}$ | The (median) consensus analyst forecast of earnings per share three months after fiscal year end, divided by average total assets per share at the forecast date. |
| $\text{Actual}^{\text{IBES}}$ | $\text{Actual}^{\text{IBES}}$ is the actual earnings per share reported by IBES, divided by average total assets per share at the forecast date. Shares are adjusted for stock splits occurring between the forecast date and fiscal year end. |
| $\text{Actual}^{\text{GAAP}}$ | The applicable basic or diluted earnings per share from Compustat (matched to the IBES definition) before extraordinary items (EPSPX or EPSFX) divided by average total assets per share at the analysts' forecast date. Shares are adjusted for stock splits occurring between the forecast date and fiscal year end. |
| Exclusions | $\text{Actual}^{\text{IBES}} - \text{Actual}^{\text{GAAP}}$. |
| FE | $\text{Actual}^{\text{IBES}} - \text{Forecast}^{\text{IBES}}$. |
| Prevol | The standard deviation of a stock's daily return for a period of 12 months ending at fiscal year-end. A minimum of 60 daily return observations is required. In our regression tests, we use the logarithm of PreVol following Lui et al. (2007). |
| ARET | Market adjusted buy and hold returns over the fiscal year. |
| TA | The logarithm of total assets (Horton et al. 2013). |
| B/M | Book value of common equity (*CEQ*) divided by market value of equity at fiscal year-end. |
| Numan | The total number of I/B/E/S analysts covering a firm three months after fiscal year end. |
| Turnover | Number of shares traded in year t, divided by the firm's average number of shares outstanding in year t (Tan et al. 2011). |
| DSEC | A dummy variable equal to 1 if the firm issued equity or debt greater than 5% of total assets in year t (Tan et al. 2011). |
| Intan_asset | The ratio of intangible assets to total assets at the beginning of the year (Tan et al. 2011). |
| EarnVol | The time-series standard deviation of earnings before extraordinary items (*IBC*) scaled by average total assets, computed recursively using at least 5 years including the current year. |
| \|ACC\| | The absolute value of ACC. |
| Loss | A dummy variable that takes the value of 1 if current earnings is negative and zero otherwise. |

## References


Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5 (1): 31-56.

Ashbaugh-Skaife, H., D. W. Collins, and R. LaFond. 2006. The effects of corporate governance on firms' credit ratings. *Journal of Accounting and Economics* 42 (1-2): 203-243.

Angrist, J. D., and J. S. Pischke. 2009. Mostly harmless econometrics: an empiricist's companion. *Princeton University Press*.

Baginski, S. P., and J. M. Wahlen. 2003. Residual income risk, intrinsic values, and share prices. *The Accounting Review* 78 (1): 327-351.

Ball, R., and L. Shivakumar. 2006. The role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research* 44 (2): 207-242.

Bao, J., J. Pan, and J. Wang. 2011. The illiquidity of corporate bonds. *The Journal of Finance* 66 (3): 91-946.

Barth, M. E., D. P. Cram, and K. K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (1): 27-58.

Barth, M. E., W. H. Beaver, J. R. M. Hand, and W. R. Landsman. 1999. Accruals, cash flows and equity values. *Review of Accounting Studies* 4 (3-4): 205-229.

Barron, O. E., O. Kim, S. C. Lim, and D. E. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73 (4): 421-433.

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

Bauwens, L., and S. Laurent. 2005. A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics* 23 (3): 346-354.

Bessembinder, H., K. M. Kahle, W. F. Maxwell, and D. Xu. 2009. Measuring abnormal bond performance. *The Review of Financial Studies* 22 (10): 4219-4258.

Bowley, A. L. 1920. Elements of statistics. Scribner's, New York.

Bradshaw, M. T., and R. G. Sloan. 2002. GAAP versus the Street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (1): 41-66.

Campbell, J. Y. and G. B. Taksler. 2003. Equity volatility and corporate bond yields. *The Journal of Finance* 58 (6): 2321-2350.

Chang, W., S. Monahan, and A. Ouazad. 2013. The higher moments of future return on equity. *Working Paper*, INSEAD.

Cheng, M. and K. R. Subramanyam. 2008. Analyst following and credit ratings. *Contemporary Accounting Research* 25 (4): 1007-1043.

Correia, M., J. Kang, and S. Richardson. 2013. Asset volatility. *Working Paper*, London Business School.

Dichev, I. D., and V. W. Tang. 2009. Earnings volatility and earnings predictability. *Journal of Accounting and Economics* 47 (1-2): 160-181.

Doyle, J. T., R. J. Lundholm, and M. T. Soliman. 2003. The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8 (2-3): 145-174.

Doyle, J. T., J. N. Jennings, and M. T. Soliman. 2013. Do managers define non-GAAP earnings to meet or beat analyst forecasts? Journal of Accounting and Economics 56 (1): 40-56.

Efron, B. 1991. Regression percentiles using asymmetric squared error loss. *Statistica Sinica* 1. 93-125.

Elton, E., M. J. Gruber, D. Agrawal, and C. Mann. 2001. Explaining the rate spread on corporate bonds. *The Journal of Finance* 56 (1): 247-277.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607-636.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39 (2), 295-327.

Ghysels, E., A. Plazzi, and R. Valkanov. 2011. Conditional skewness of stock market returns in developed and emerging markets and its economic fundamentals. *Swiss Finance Institute Research Paper* 11-06.

Granger, C. W. J., and Z. Ding. 1995. Some properties of absolute return: An alternative measure of risk. *Annales d'Économie et de Statistique* (40): 67-91.

Harvey, C. R., and A. Siddique. 1999. Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis* 34 (4): 465-487.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (2): 125-153.

Hinkley, D. V. 1975. On power transformations to symmetry. *Biometrika* 62 (1): 101-111.

Hribar, P., and D. W. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (1): 105-134.

Horton, J., G. Serafeim, and I. Serafeim. 2013. Does mandatory IFRS adoption improve the information environment? *Contemporary Accounting Research* 30 (1): 388-423.

Joos, P., J. D. Piotroski, and S. Srinivasan. 2014. Can analysts assess fundamental risk and valuation uncertainty? An empirical analysis of scenario-based value estimates. *Working Paper*.

Kim, T. H., and H. White. 2004. On more robust estimation of skewness and kurtosis. *Finance Research Letters* 1 (1): 56-73.

Koenker, R., and G. Bassett Jr. 1978. Regression quantiles. *Econometrica* 46 (1): 33-50.

Koenker, R., and G. Bassett Jr. 1982. Robust tests for heteroscedasticity based on regression quantiles. *Econometrica: Journal of the Econometric Society* 50 (1): 43-61.

León, A., G. Rubio, and G. Serna. 2005. Autoregressive conditional volatility, skewness and kurtosis. *The Quarterly Review of Economics and Finance* 45 (4): 599-618.

Lim, T. 2001. Rationality and analysts' forecast bias. *The Journal of Finance* 56 (1): 369-385.

Lougee, B., and C. Marquardt. 2004. Earnings informativeness and strategic disclosure: An empirical examination of 'pro forma' earnings. *The Accounting Review* 79 (3): 769-795.

Lui, D., S. Markov, and A. Tamayo. 2007. What makes a stock risky? Evidence from sell-side analysts' risk ratings. *Journal of Accounting Research* 45 (3): 629-664.

Newey, W. K., and J. L. Powell. 1987. Asymmetric least squares estimation and testing. *Econometrica: Journal of the Econometric Society* 55 (4): 819-847.

Nissim, D., and S. H. Penman. 2001. Ratio analysis and equity valuation: from research to practice. *Review of Accounting Studies* 6 (1): 109 – 154.

Moors, J. J. A. 1988. A quantile alternative for kurtosis. *The Statistician* 37, 25-32.

Penman, S. H. 2010. Financial forecasting, risk and valuation: Accounting for the future. *Abacus* 46 (2): 211-228.

Poon, S. H., and C. W. Granger. 2003. Forecasting volatility in financial markets: A review. *Journal of Economic Literature* 41 (2): 478-539.

Richardson, S. A., R. G. Sloan, M. T. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437-485.

Serafeim, G. 2011. Consequences and Institutional Determinants of Unregulated Corporate Financial Statements: Evidence from Embedded Value Reporting. *Journal of Accounting Research* 49 (2): 529-571

Sheng, X., and M. Thevenot. 2012. A new measure of earnings forecast uncertainty. *Journal of Accounting and Economics* 53 (1-2): 21-33.

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

Tan, H., Wang, S. and M. Welker. 2011. Analyst following and forecast accuracy after mandated IFRS adoptions. *Journal of Accounting Research* 49 (5): 1307-1357.

Tay, A. S., and K. F. Wallis. 2000. Density forecasting: A survey. *Journal of Forecasting* 19 (4): 235-254.

Taylor, J. W. 2008. Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics* 6 (2): 231-252.

Vincent, L. 1999. The information content of funds from operations (FFO) for real estate investment trusts (REITs). *Journal of Accounting and Economics* 26 (1-3): 69-104

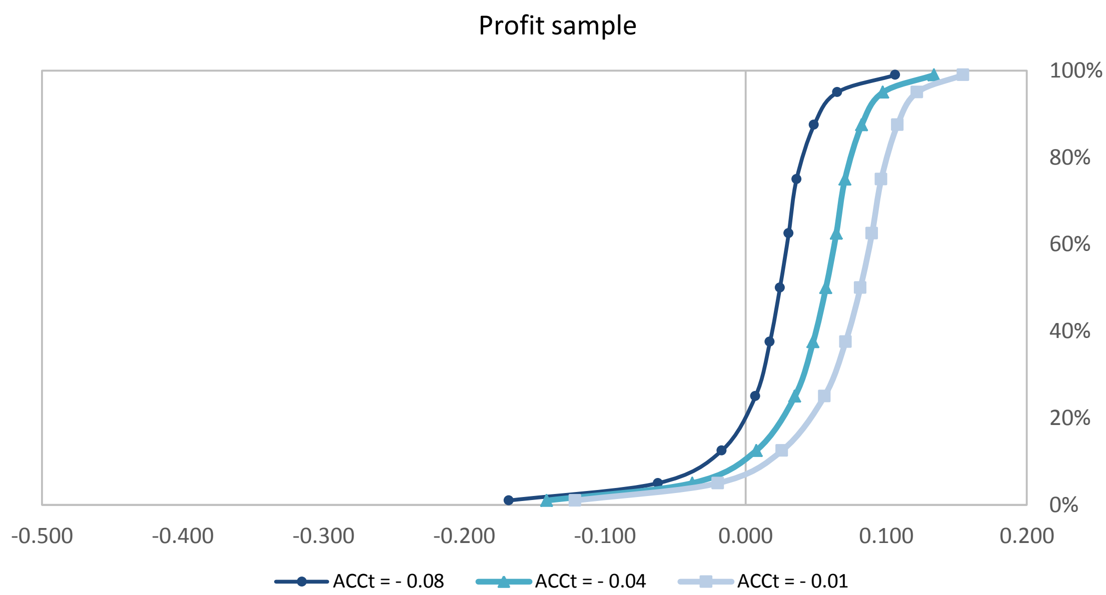
White, H., T. H. Kim, and S. Manganelli. 2008. *Modeling autoregressive conditional skewness and kurtosis with multi-quantile CAViaR* (No. 957). European Central Bank.

Yao, Q., and H. Tong. 1996. Asymmetric least squares regression estimation: A nonparametric approach. *Journal of Nonparametric Statistics* 6 (2-3): 273-292.

Figure 1

Cumulative probability distribution of future earnings

Panel A: OCF is held constant



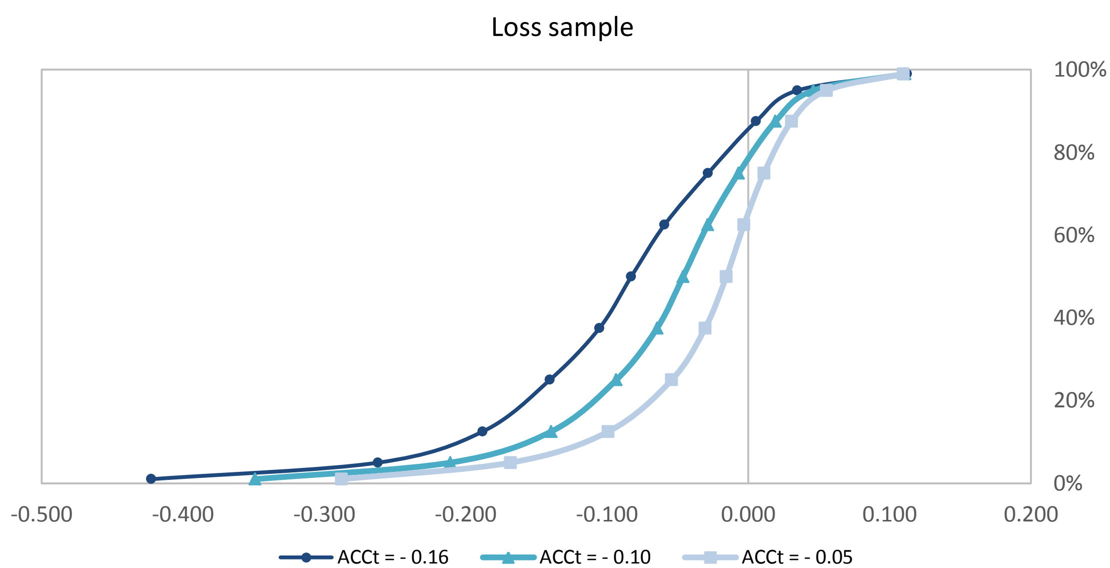
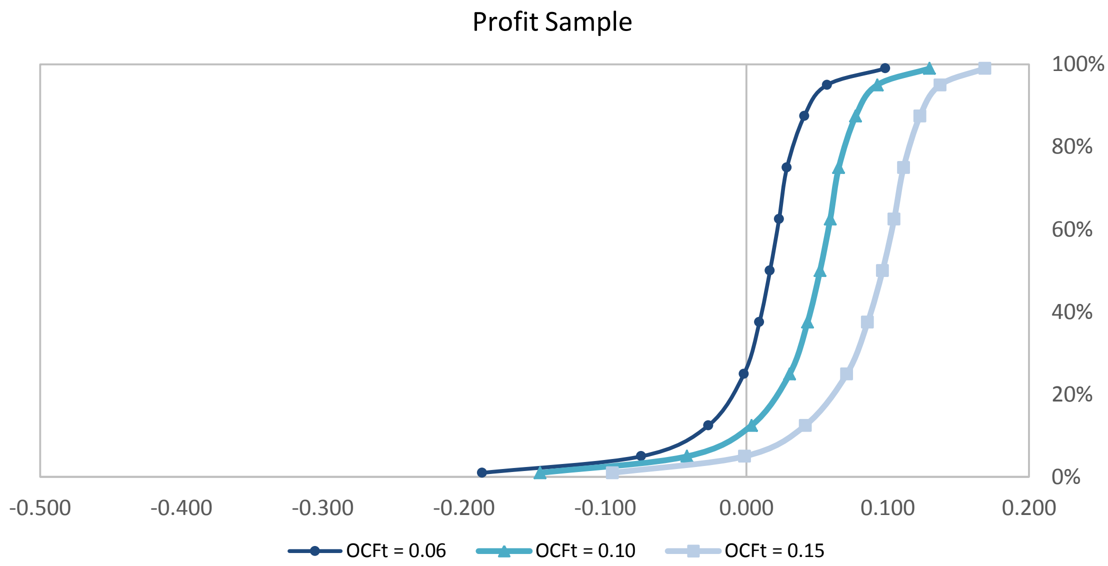
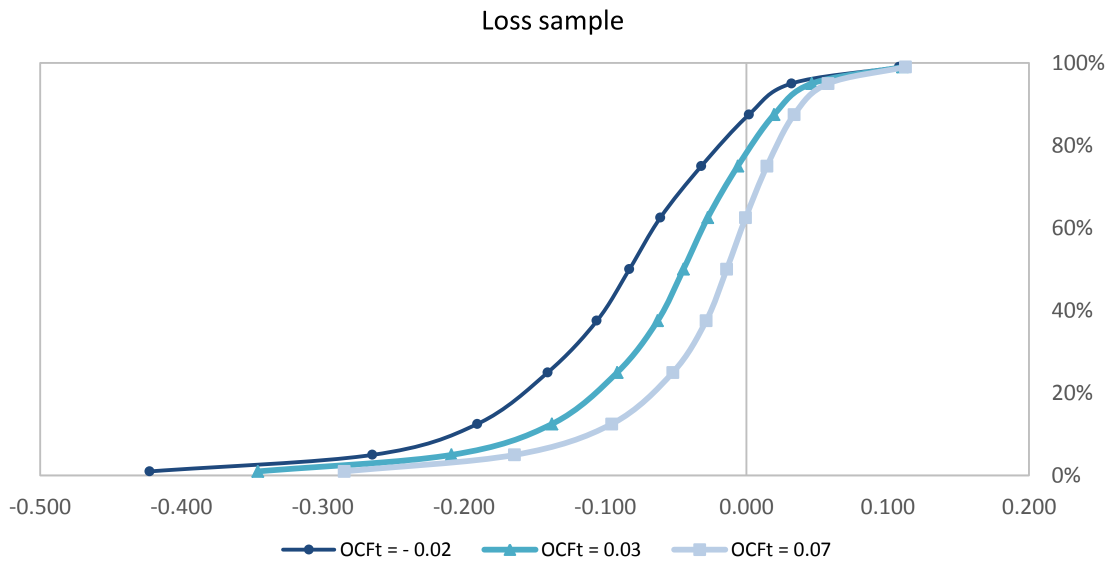



Figure 1 (Continued)

Panel B: ACC is held constant




The figure illustrates the cumulative probability distribution function of $EARN_{t+1}$ conditional on accruals and cash flow. In Panel A, the value of $OCF_t$ is held constant at its median level in the profit (Median $OCF_t$ = 0.102) and loss sub-samples (Median $OCF_t$ = 0.032) while $ACC_t$ varies based on the $Q_{25}$, $Q_{50}$ and $Q_{75}$ values of the actual ACC distribution in each sub-sample. In Panel B, the value of $ACC_t$ is held constant at its median level in the profit (Median $ACC_t$ = -0.044) and loss sub-samples (Median $ACC_t$ = -0.095) while $OCF_t$ varies based on the $Q_{25}$, $Q_{50}$ and $Q_{75}$ values of the actual OCF distribution in each sub-sample. The figure uses 11 quantile estimates in the set $\tau \in$ {0.01, 0.05, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 0.95, 0.99}.The variables are defined in Appendix A.

Figure 2

Out-of-sample forecasts of risk in future earnings

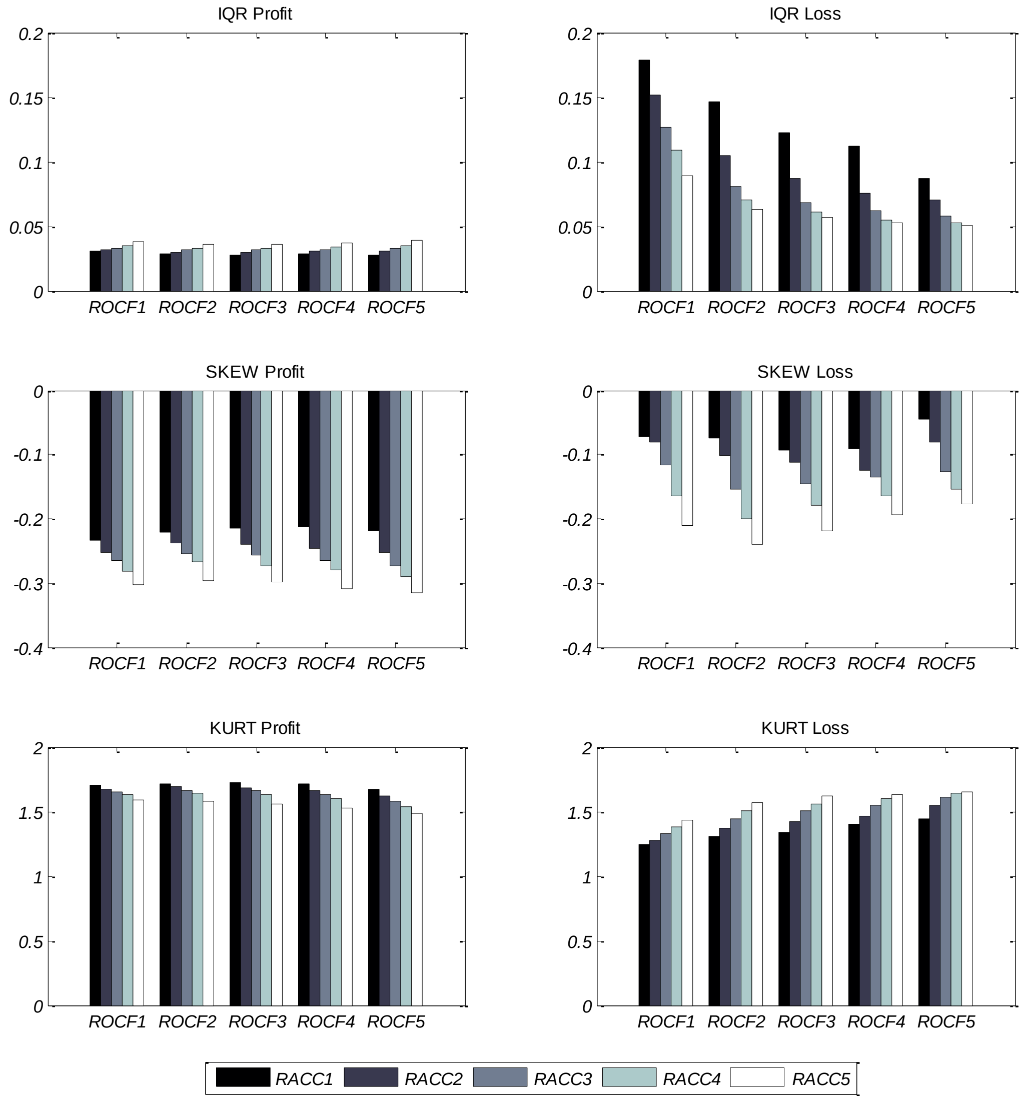
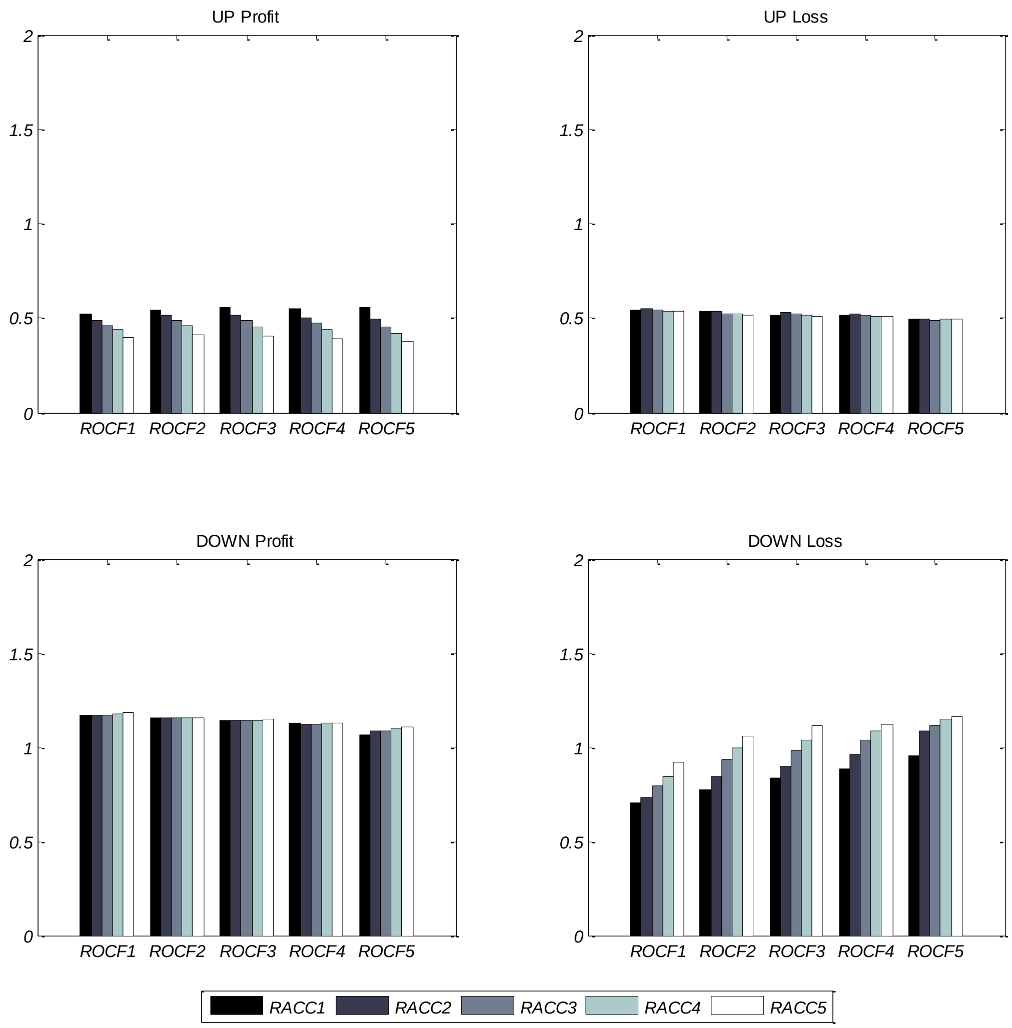


Figure 2 (continued)




The figure provides insights into how risk forecasts depend on OCF and ACC and the sign of EARN. In each panel, we plot the median of each estimated risk metric (IQR, SKEW, KURT, UP and DOWN) for 5×5 sequentially sorted portfolios, sorted first by industry, second by OCF and third by ACC. Portfolios are formed separately for profit and loss firms. We require at least 25 observations within each industry-year-profit and industry-year-loss group. RACC1 – RACC2 and ROCF1 – ROCF5 denote the 5 portfolio groups based on ACC and OCF respectively. The variables are defined in Appendix A.

TABLE 1

Sample formation

**Panel A**: Data selection

| | Firm-years | Firms |
|---|---|---|
| Matched Compustat/CRSP for the period 1987 – 2009 | 147,285 | 17,541 |
| Less stocks other than NYSE, AMEX or Nasdaq stocks | (2,978) | (304) |
| Sample with stocks listed on NYSE, AMEX or Nasdaq | 144,307 | 17,237 |
| Less stocks other than ordinary common stocks | (18,308) | (2,187) |
| Sample with ordinary common stocks listed on NYSE, AMEX or Nasdaq | 125,999 | 15,050 |
| Less financial firms | (24,389) | (2,710) |
| Non-financial firms with ordinary common stocks listed on NYSE, AMEX or Nasdaq | 101,610 | 12,340 |
| Less observations with missing Barth et al. (1999) industry classification | (719) | (107) |
| Less observations with TA ≤ $100 million | (47,274) | (5,714) |
| Less observations with missing deflated EARN, ACC, OCF, SI | (7,735) | (838) |
| Less observations at the extreme 1% of the distribution of EARN, ACC, OCF, SI by year | (2,356) | (165) |
| Final sample with non-missing EARN, ACC, OCF and SI | 43,526 | 5,516 |

**Panel B:** Industry composition

| | Industry | Primary SIC codes | Firm-years | %of obs |
|---|---|---|---|---|
| 1 | Mining and construction | 1000–1999, excluding 1300–1399 | 1,185 | 2.72 |
| 2 | Food | 2000–2111 | 1,408 | 3.23 |
| 3 | Textiles, printing/pub | 2200–2780 | 3,090 | 7.10 |
| 4 | Chemicals | 2800–2824, 2840–2899 | 1,542 | 3.54 |
| 5 | Pharmaceuticals | 2830–2836 | 1,296 | 2.98 |
| 6 | Extractive industries | 2900–2999, 1300–1399 | 1,903 | 4.37 |
| 7 | Durable manufacturers | 3000–3999, excluding 3570–3579 and 3670–3679 | 10,332 | 23.74 |
| 8 | Computers | 7370–7379, 3570–3579, 3670–3679 | 5,218 | 11.99 |
| 9 | Transportation | 4000–4899 | 3,357 | 7.71 |
| 10 | Utilities | 4900–4999 | 3,394 | 7.80 |
| 11 | Retail | 5000–5999 | 6,404 | 14.71 |
| 12 | Services | 7000–8999, excluding 7370–7379 | 4,397 | 10.10 |
| Total | | | 43,526 | 100 |

TABLE 2

Descriptive statistics for the variables used in forecasting models

**Panel A:** Distributional statistics

| | Mean | St.dev | Skewness | Kurtosis | 1% | 5% | 25% | Median | 75% | 95% | 99% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EARN | 0.030 | 0.090 | -2.160 | 10.722 | -0.307 | -0.136 | 0.006 | 0.041 | 0.076 | 0.145 | 0.202 |
| ACC | -0.059 | 0.076 | -1.431 | 7.668 | -0.317 | -0.185 | -0.090 | -0.052 | -0.019 | 0.047 | 0.115 |
| OCF | 0.089 | 0.081 | -0.271 | 1.398 | -0.141 | -0.046 | 0.045 | 0.088 | 0.137 | 0.223 | 0.291 |
| SI | -0.013 | 0.041 | -5.256 | 43.191 | -0.192 | -0.080 | -0.010 | 0.000 | 0.000 | 0.008 | 0.034 |

**Panel B:** Pearson (Spearman) Correlations above (below) the diagonal

| | EARN | ACC | OCF | SI |
|---|---|---|---|---|
| EARN | 1 | 0.533 | 0.614 | 0.554 |
| ACC | 0.300 | 1 | -0.340 | 0.501 |
| OCF | 0.617 | -0.458 | 1 | 0.149 |
| SI | 0.332 | 0.271 | 0.106 | 1 |

Panel A reports the distribution of the main variables and Panel B reports Pearson (above diagonal) and Spearman (below diagonal) correlation statistics. Descriptives are based on a sample of 43,526 firm-year observations. The variables are defined as follows: EARN is income before extraordinary items scaled by average total assets, OCF is operating cash flow minus extraordinary items and discontinued operations scaled by average total assets, ACC is accruals defined as EARN minus OCF and SI is special items scaled by average total assets.

TABLE 3

Forecasting the distribution of future earnings – OLS and quantile regressions

$$EARN_{t+1} = \alpha_{01}d^{+} + \alpha_{02}d^{-} + \alpha_{11}ACC_{t}.d^{+} + \alpha_{21}OCF_{t}.d^{+} + \alpha_{31}SI_{t}.d^{+} + \alpha_{12}ACC_{t}.d^{-} + \alpha_{22}OCF_{t}.d^{-} + \alpha_{32}SI_{t}.d^{-} + v_{t+1}$$

**Panel A:** Coefficient estimates

| | 1% | 5% | 12.5% | 25% | 37.5% | 50% | 62.5% | 75% | 87.5% | 95% | 99% | OLS | IQR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean($\alpha_{01}$) | -0.219 | -0.096 | -0.046 | -0.019 | -0.008 | 0.000 | 0.006 | 0.012 | 0.024 | 0.040 | 0.082 | -0.010 | 0.031 |
| Mean($\alpha_{02}$) | -0.276 | -0.161 | -0.093 | -0.046 | -0.024 | -0.009 | 0.001 | 0.013 | 0.031 | 0.056 | 0.107 | -0.028 | 0.059 |
| ACC . $d^{+}$ | 0.673 | 0.609 | 0.612 | 0.700 | 0.768 | 0.817 | 0.846 | 0.860 | 0.846 | 0.812 | 0.687 | 0.734 | 0.160 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| $OCF_t$ .$d^{+}$ | 1.022 | 0.812 | 0.761 | 0.809 | 0.851 | 0.885 | 0.905 | 0.917 | 0.910 | 0.889 | 0.782 | 0.848 | 0.108 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| SI .$d^{+}$ | 0.045 | -0.141 | -0.170 | -0.269 | -0.368 | -0.424 | -0.474 | -0.504 | -0.486 | -0.441 | -0.348 | -0.311 | -0.236 |
| | (0.848) | (0.005) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| ACC .$d^{-}$ | 1.223 | 0.853 | 0.808 | 0.784 | 0.680 | 0.611 | 0.511 | 0.359 | 0.228 | 0.187 | -0.025 | 0.526 | -0.425 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.822) | (0.000) | (0.000) |
| OCF .$d^{-}$ | 1.530 | 1.119 | 1.057 | 0.984 | 0.859 | 0.765 | 0.671 | 0.516 | 0.355 | 0.289 | 0.052 | 0.737 | -0.468 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.592) | (0.000) | (0.000) |
| SI .$d^{-}$ | -0.594 | -0.427 | -0.474 | -0.547 | -0.537 | -0.520 | -0.458 | -0.359 | -0.287 | -0.255 | -0.064 | -0.428 | 0.189 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.432) | (0.000) | (0.000) |
| Overall $R^2$ | 0.273 | 0.310 | 0.313 | 0.306 | 0.320 | 0.347 | 0.378 | 0.405 | 0.418 | 0.399 | 0.297 | 0.454 | |
| Incremental $R^2$ | 0.180 | 0.229 | 0.258 | 0.285 | 0.311 | 0.336 | 0.355 | 0.367 | 0.367 | 0.341 | 0.234 | 0.433 | |

TABLE 3 (continued)

$$EARN_{t+1} = \alpha_{01}d^{+} + \alpha_{02}d^{-} + \alpha_{11}ACC_{t}.d^{+} + \alpha_{21}OCF_{t}.d^{+} + \alpha_{31}SI_{t}.d^{+} + \alpha_{12}ACC_{t}.d^{-} + \alpha_{22}OCF_{t}.d^{-} + \alpha_{32}SI_{t}.d^{-} + v_{t+1}$$

**Panel B:** Restrictions

| | 1% | 5% | 12.5% | 25% | 37.5% | 50% | 62.5% | 75% | 87.5% | 95% | 99% | OLS | IQR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ACC v. OCF | | | | | | | | | | | | | |
| $\alpha_{11} - \alpha_{21}$ | -0.349 | -0.203 | -0.149 | -0.109 | -0.083 | -0.068 | -0.059 | -0.057 | -0.064 | -0.077 | -0.095 | -0.114 | 0.052 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| $\alpha_{12} - \alpha_{22}$ | -0.307 | -0.266 | -0.249 | -0.200 | -0.179 | -0.154 | -0.160 | -0.157 | -0.127 | -0.102 | -0.077 | -0.211 | 0.043 |
| | (0.021) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.174) | (0.000) | (0.041) |
| Profit v. Loss | | | | | | | | | | | | | |
| $\alpha_{11} - \alpha_{12}$ | -0.550 | -0.244 | -0.196 | -0.084 | 0.088 | 0.206 | 0.335 | 0.501 | 0.618 | 0.625 | 0.712 | 0.208 | 0.585 |
| | (0.003) | (0.001) | (0.001) | (0.011) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) | (0.000) |
| $\alpha_{21} - \alpha_{22}$ | -0.508 | -0.307 | -0.296 | -0.175 | -0.008 | 0.120 | 0.234 | 0.401 | 0.555 | 0.600 | 0.730 | 0.111 | 0.576 |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.784) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.070) | (0.000) |

The table is based on a sample of 36,544 firm-year observations. Panel A reports regression results using current ACC, CF, SI and future EARN. Panel B reports restrictions between coefficient estimates. Standard errors in the OLS regressions are clustered by firm and year and in the quantile regressions they are estimated via bootstrapping. Industry dummies are included and the average of industry-specific intercepts is reported. P-values are reported in parentheses. Reported $R^2$ statistics include the Overall $R^2$, which incorporates the contribution of the industry fixed effects, and the Incremental $R^2$s contributed by ACC, OCF, $d_1$ and $d_2$ beyond industry fixed effects. The variables are defined as follows: EARN is income before extraordinary items scaled by average total assets, OCF is operating cash flow minus extraordinary items and discontinued operations scaled by average total assets, ACC is accruals defined as EARN minus OCF and SI is special items scaled by average total assets. $d^{+}(d^{-})$ is an indicator variable equal to one if $EARN_t \geq 0$ ($EARN_t < 0$).

TABLE 4

Descriptive statistics of risk measures

**Panel A:** Distributional statistics

| | Mean | St. Dev. | 1% | 5% | 25% | Median | 75% | 95% | 99% |
|---|---|---|---|---|---|---|---|---|---|
| IQR | 0.048 | 0.037 | 0.008 | 0.010 | 0.028 | 0.035 | 0.056 | 0.121 | 0.200 |
| SKEW | -0.235 | 0.099 | -0.430 | -0.381 | -0.298 | -0.249 | -0.184 | -0.043 | 0.053 |
| KURT | 1.608 | 0.185 | 1.185 | 1.311 | 1.488 | 1.601 | 1.729 | 1.912 | 2.090 |
| UP | 0.493 | 0.094 | 0.305 | 0.351 | 0.430 | 0.487 | 0.547 | 0.659 | 0.777 |
| DOWN | 1.114 | 0.179 | 0.623 | 0.783 | 1.019 | 1.109 | 1.227 | 1.399 | 1.514 |
| $Q_{50}$ | 0.040 | 0.064 | -0.179 | -0.071 | 0.015 | 0.042 | 0.075 | 0.137 | 0.188 |

**Panel B:** Correlations

| | IQR | SKEW | KURT | UP | DOWN | $Q_{50}$ |
|---|---|---|---|---|---|---|
| IQR | 1 | 0.275 | -0.485 | -0.088 | -0.449 | -0.643 |
| SKEW | -0.024 | 1 | -0.265 | 0.645 | -0.611 | -0.597 |
| KURT | -0.470 | -0.195 | 1 | 0.288 | 0.861 | 0.227 |
| UP | -0.305 | 0.728 | 0.214 | 1 | -0.228 | -0.436 |
| DOWN | -0.278 | -0.528 | 0.858 | -0.256 | 1 | 0.463 |
| $Q_{50}$ | -0.217 | -0.568 | 0.040 | -0.540 | 0.253 | 1 |

The table refers to the period 1993 to 2009. Panel A reports the distribution of quantile-based estimates and Panel B reports Pearson (above diagonal) and Spearman (below diagonal) correlation statistics. IQR, SKEW, KURT, UP and DOWN are out-of-sample estimates of the spread, skewness, kurtosis, upside risk and downside risk in future earnings. $Q_{50}$ is the median forecast of earnings.

TABLE 5

Equity risk

**Panel A:** Relation with earnings risk

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_2 SKEW + \gamma_3 KURT + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon$$

| | Equity_Rating | | | PostVol | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| IQR | 42.685 | 29.485 | 27.224 | 4.944 | 1.236 | 1.440 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| SKEW | 0.398 | 1.555 | 0.866 | 0.051 | 0.075 | -0.020 |
| | (0.710) | (0.146) | (0.508) | (0.415) | (0.154) | (0.804) |
| KURT | 2.355 | 2.310 | 2.391 | 0.169 | 0.033 | -0.008 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.105) | (0.771) |
| $Q_{50}$ | -8.264 | -1.222 | -7.112 | -0.633 | 0.008 | -0.120 |
| | (0.000) | (0.512) | (0.011) | (0.000) | (0.954) | (0.479) |
| PreVol | | | | | 0.638 | 0.575 |
| | | | | | (0.000) | (0.000) |
| IVol | | 148.771 | 154.255 | | | |
| | | (0.000) | (0.000) | | | |
| Beta | | 0.847 | 0.673 | | 0.042 | 0.066 |
| | | (0.000) | (0.008) | | (0.080) | (0.137) |
| Illiquidity | | 3.690 | 6.296 | | 0.000 | 0.000 |
| | | (0.188) | (0.018) | | (0.879) | (0.995) |
| MV | | -0.563 | -0.471 | | -0.039 | -0.036 |
| | | (0.000) | (0.000) | | (0.000) | (0.007) |
| NegBV | | 0.750 | -0.077 | | 0.002 | 0.082 |
| | | (0.258) | (0.937) | | (0.953) | (0.086) |
| B/M | | 0.357 | 0.744 | | 0.012 | 0.034 |
| | | (0.152) | (0.068) | | (0.159) | (0.000) |
| B/M*NegBV | | -0.669 | -0.568 | | -0.035 | -0.066 |
| | | (0.240) | (0.444) | | (0.182) | (0.027) |
| D/E | | 0.095 | 0.097 | | 0.004 | 0.005 |
| | | (0.086) | (0.179) | | (0.105) | (0.019) |
| D/E*NegBV | | -0.137 | -0.230 | | -0.012 | -0.002 |
| | | (0.379) | (0.326) | | (0.003) | (0.809) |
| IPO | | 0.081 | 0.000 | | 0.032 | 0.020 |
| | | (0.862) | N/A | | (0.005) | (0.568) |
| EarnVol | | | 7.956 | | | 0.367 |
| | | | (0.122) | | | (0.019) |
| EarnQual | | | 26.684 | | | 0.470 |
| | | | (0.000) | | | (0.249) |
| Loss | | | -0.338 | | | 0.040 |
| | | | (0.356) | | | (0.048) |
| $R^2$ | 0.125 | 0.337 | 0.350 | 0.146 | 0.460 | 0.416 |
| N | 2,083 | 2,075 | 1,400 | 32,361 | 32,226 | 13,477 |

TABLE 5 (Continued)

**Panel B:** Relation with earnings risk components

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_{31} UP + \gamma_{32} DOWN + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon$$

| | Equity_Rating | | | PostVol | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| IQR | 40.825 | 28.367 | 24.106 | 4.305 | 1.073 | 1.362 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| UP | 1.713 | 2.509 | 2.142 | -0.232 | -0.050 | -0.064 |
| | (0.069) | (0.007) | (0.048) | (0.038) | (0.541) | (0.440) |
| DOWN | 2.451 | 2.097 | 2.462 | 0.215 | 0.034 | 0.004 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.177) | (0.892) |
| $Q_{50}$ | -9.323 | -2.231 | -7.644 | -1.225 | -0.174 | -0.186 |
| | (0.000) | (0.223) | (0.007) | (0.000) | (0.268) | (0.282) |
| Controls | No | Yes | Yes | No | Yes | Yes |
| $R^2$ | 0.125 | 0.337 | 0.350 | 0.150 | 0.460 | 0.416 |
| N | 2,083 | 2,075 | 1,400 | 32,361 | 32,226 | 13,477 |

The table reports results from regressing the *Outcome* variable on our risk metrics and a set of control variables. In the equity rating results, the sample is based on 2,083 observations covering the period 2003 – 2009. *Outcome* is defined as Equity_Rating and the results are obtained from ordered logit regressions. Reported $R^2$s are pseudo $R^2$s incremental to those obtained when only year fixed effects are included in the model. Standard errors are clustered by firm. In the return volatility results, the sample is based on 32,361 observations covering the period 1993 – 2009. *Outcome* is defined as PostVol and the results are obtained from OLS regressions. Reported $R^2$s are adjusted $R^2$s incremental to those obtained when only year fixed effects are included in the model. Standard errors are clustered by firm and year. In both tests, the control variables include PreVol, Beta, Illiquidity, MV, NegBV, B/M, B/M*NegBV, D/E, D/E*NegBV, IPO, EarnVol, EarnQual and Loss. Year fixed effects are included in the models, but not reported. P-values are reported in parentheses. Equity_Rating is a discrete variable taking the values 1, 2, 3, 4 (low, medium, high and speculative risk) and PostVol is the standard deviation of a stock's daily return for a period of 12 months, starting three months after the firm's fiscal year-end. IQR, SKEW, KURT, UP and DOWN are out-of-sample estimates of the spread, skewness, kurtosis, upside risk and downside risk in future earnings. $Q_{50}$ is the median forecast of earnings. Control variables are defined in Appendix A.

TABLE 6

Credit ratings

**Panel A:** Relation with earnings risk

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_2 SKEW + \gamma_3 KURT + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon$$

| | Speculative_Grade | | | Credit_Rating | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| IQR | 40.840 | 34.641 | 25.377 | 19.851 | 14.984 | 5.197 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.150) |
| SKEW | -0.486 | -1.749 | -2.167 | -0.799 | -1.654 | -1.984 |
| | (0.367) | (0.007) | (0.003) | (0.095) | (0.001) | (0.000) |
| KURT | 0.824 | 0.547 | 0.522 | 0.476 | 0.404 | 0.251 |
| | (0.006) | (0.091) | (0.171) | (0.075) | (0.123) | (0.398) |
| $Q_{50}$ | -20.167 | -5.818 | -10.726 | -21.421 | -13.125 | -16.787 |
| | (0.000) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) |
| Prevol | | 3.489 | 3.130 | | 2.946 | 2.697 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| Beta | | -0.024 | -0.074 | | -0.017 | -0.056 |
| | | (0.913) | (0.723) | | (0.894) | (0.702) |
| RET | | 1.161 | 1.144 | | 0.802 | 0.771 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| Mn_Prc | | -0.012 | -0.009 | | -0.008 | -0.005 |
| | | (0.038) | (0.138) | | (0.009) | (0.134) |
| MV | | -1.022 | -1.012 | | -0.809 | -0.807 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| B/M | | -0.066 | 0.127 | | -0.031 | 0.111 |
| | | (0.690) | (0.537) | | (0.674) | (0.218) |
| D/A | | 5.253 | 5.926 | | 3.412 | 3.771 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| Cover | | 0.001 | 0.001 | | 0.001 | 0.001 |
| | | (0.054) | (0.038) | | (0.100) | (0.030) |
| Intan | | 1.871 | 0.827 | | 0.567 | -0.469 |
| | | (0.199) | (0.597) | | (0.575) | (0.649) |
| ΔEQ | | 0.174 | 0.092 | | 0.120 | 0.043 |
| | | (0.343) | (0.676) | | (0.257) | (0.722) |
| Transp | | -0.111 | -0.153 | | -0.007 | -0.071 |
| | | (0.535) | (0.559) | | (0.867) | (0.133) |
| EarnVol | | | 29.114 | | | 24.041 |
| | | | (0.000) | | | (0.000) |
| \|ABACC\| | | | -2.306 | | | -1.973 |
| | | | (0.129) | | | (0.066) |
| Loss | | | -0.492 | | | -0.121 |
| | | | (0.002) | | | (0.281) |
| Pseudo $R^2$ | 0.188 | 0.547 | 0.546 | 0.121 | 0.359 | 0.362 |
| N | 15,307 | 13,226 | 11,077 | 15,307 | 13,226 | 11,077 |

TABLE 6 (Continued)

**Panel B:** Relation with earnings risk components

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_{31} UP + \gamma_{32} DOWN + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon.$$

| | Speculative_Grade | | | Credit_Rating | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| IQR | 38.960 | 34.520 | 24.981 | 17.214 | 14.522 | 5.443 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.123) |
| UP | -0.029 | -0.186 | -1.241 | -1.146 | -0.428 | -1.090 |
| | (0.964) | (0.811) | (0.114) | (0.041) | (0.424) | (0.034) |
| DOWN | 1.086 | 0.903 | 1.186 | 0.867 | 0.751 | 0.757 |
| | (0.001) | (0.016) | (0.005) | (0.003) | (0.010) | (0.024) |
| $Q_{50}$ | -20.760 | -5.179 | -11.557 | -22.965 | -12.664 | -17.049 |
| | (0.000) | (0.002) | (0.000) | (0.000) | (0.000) | (0.000) |
| Controls | No | Yes | Yes | No | Yes | Yes |
| Pseudo $R^2$ | 0.189 | 0.546 | 0.546 | 0.123 | 0.359 | 0.362 |
| N | 15,307 | 13,226 | 11,077 | 15,307 | 13,226 | 11,077 |

The table is based on a sample of 15,307 firm-year observations over the period 1993 to 2009. Panels A and B report results from regressing the *Outcome* variable on our risk metrics and a set of controls. When *Outcome* is defined as Speculative_Grade, the results are based on logit regressions. When *Outcome* is defined as Credit_Rating the results are based on ordered logit regressions. The control variables include PreVol, Beta, RET, Mn_Prc, MV, B/M, D/A, Cover, Intan, ΔEQ, Transp, EarnVol, |ABACC| and Loss. Year fixed effects are included but not reported. P-values (in parentheses) are based on clustered standard errors by firm and year. Reported $R^2$s are pseudo $R^2$s incremental to those obtained when only year fixed effects are included in the model. Speculative_Grade is a dummy variable equal to one for speculative grade bonds and zero for investment grade bonds. Credit_Rating is the S&P'S long-term issuer credit rating, taking values from 1 (low risk) to 7 (high risk). IQR, SKEW, KURT, UP and DOWN are out-of-sample estimates of the spread, skewness, kurtosis, upside risk and downside risk in future earnings. $Q_{50}$ is the median forecast of earnings. Control variables are defined in Appendix A.

TABLE 7

Corporate Bond Yields

**Panel A:** Bond Sample formation

| | Bond-months | Bond-years | Bonds | Firms |
|---|---|---|---|---|
| Merged FISD&TRACE monthly files | 638,936 | 111,198 | 39,916 | 5,282 |
| Minus floating-rate and foreign currency bonds | (101,322) | (24,504) | (11,468) | (426) |
| Minus bonds with odd frequency of coupon payments | (94) | (40) | (26) | (0) |
| Minus puttable and convertible bonds | (39,750) | (5,364) | (1,528) | (568) |
| Sample excluding special features | 497,770 | 81,290 | 26,894 | 4,288 |
| Minus observations with missing YtM, Spread | (21,332) | (4,757) | (3,109) | (187) |
| Sample with minimum requirements | 476,438 | 76,533 | 23,785 | 4,101 |
| Minus +/- 1% of the yield spread by month | (9,433) | (1,225) | (348) | (96) |
| Monthly sample with available yield spread | 467,005 | 75,308 | 23,437 | 4,005 |
| Merged FISD-Compustat 3 months after fiscal year-end | | (66,423) | (20,521) | (3,291) |
| Final Sample | | 8,885 | 2,916 | 714 |

TABLE 7 (Continued)

**Panel B:** Bond yield spreads and earnings risk

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_2 SKEW + \gamma_3 KURT + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon$$

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| IQR | 79.301 | 76.370 | 74.963 | 70.962 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| SKEW | 0.361 | 0.132 | 1.823 | 1.020 |
| | (0.856) | (0.947) | (0.329) | (0.591) |
| KURT | 2.973 | 2.769 | 2.016 | 1.844 |
| | (0.001) | (0.001) | (0.012) | (0.028) |
| $Q_{50}$ | -3.097 | -3.666 | 2.481 | 3.907 |
| | (0.335) | (0.251) | (0.468) | (0.251) |
| Issue_size | | -0.145 | -0.214 | -0.192 |
| | | (0.109) | (0.010) | (0.008) |
| Coupon | | 0.034 | 0.024 | 0.025 |
| | | (0.000) | (0.000) | (0.000) |
| YtM | | 0.059 | 0.060 | 0.046 |
| | | (0.183) | (0.124) | (0.212) |
| Call | | -0.142 | -0.186 | -0.192 |
| | | (0.256) | (0.139) | (0.125) |
| T-Note | | -0.468 | -0.291 | -0.301 |
| | | (0.000) | (0.000) | (0.000) |
| Term_Slope | | 0.380 | 0.508 | 0.575 |
| | | (0.011) | (0.007) | (0.002) |
| Euro_TBill | | -0.368 | -0.237 | -0.326 |
| | | (0.009) | (0.090) | (0.023) |
| PreVol | | | 1.626 | 1.214 |
| | | | (0.002) | (0.009) |
| Mn_Ret | | | -3.919 | -4.530 |
| | | | (0.001) | (0.000) |
| Cover | | | -0.001 | 0.003 |
| | | | (0.907) | (0.671) |
| Margin | | | 3.987 | 3.469 |
| | | | (0.001) | (0.002) |
| LD/A | | | -5.534 | -5.646 |
| | | | (0.258) | (0.246) |
| D/Cap | | | 16.243 | 14.187 |
| | | | (0.000) | (0.001) |
| Credit_Rating | | | | 1.403 |
| | | | | (0.000) |
| Adj. $R^2$ | 0.048 | 0.055 | 0.104 | 0.111 |
| N | 8,885 | 8,885 | 8,840 | 8,776 |

TABLE 7 (Continued)

**Panel C:** Bond yield spreads and earnings risk components

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_{31} UP + \gamma_{32} DOWN + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon.$$

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| IQR | 83.805 | 80.753 | 80.835 | 75.927 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| UP | 4.310 | 3.958 | 4.880 | 4.371 |
| | (0.003) | (0.006) | (0.000) | (0.001) |
| DOWN | 3.411 | 3.281 | 1.747 | 1.645 |
| | (0.003) | (0.004) | (0.105) | (0.137) |
| $Q_{50}$ | -1.706 | -2.295 | 5.288 | 6.832 |
| | (0.595) | (0.469) | (0.144) | (0.057) |
| | | | | |
| Controls | No | Yes | Yes | Yes |
| Adj. $R^2$ | 0.048 | 0.055 | 0.105 | 0.112 |
| N | 8,885 | 8,885 | 8,840 | 8,776 |

The table is based on a sample of 8,885 bond-year observations over the period 2002 to 2009. Panel A shows the bond sample formation. Panels B and C report OLS estimates from regressing the *Outcome* variable on our risk metrics and a set of controls. *Outcome* is defined as Spread and the control variables include Issue_size, Coupon, YtM, Call, T-note, Term_slope, Euro_TBill, Prevol, Mn_Ret, Cover, Margin, LD/A, D/Cap and Credit_Rating. Firm and year fixed effects are included but not reported. P-values (in parentheses) are based on clustered standard errors by firm. Reported $R^2$s are adjusted $R^2$s incremental to those obtained when only firm and year fixed effects are included in the model. . Spread is defined as the treasury spread, i.e. the difference between the yield to maturity on each bond and the yield on a benchmark US treasury in a particular month (in percentage). IQR, SKEW, KURT, UP and DOWN are out-of-sample estimates of the spread, skewness, kurtosis, upside risk and downside risk in future earnings. $Q_{50}$ is the median forecast of earnings. Control variables are defined in Appendix A.

TABLE 8

Analyst-based uncertainty

**Panel A:** Relation with earnings risk

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_2 SKEW + \gamma_3 KURT + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon$$

| | \|FE\| | | \|Exclusions\| | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| IQR | 0.319 | 0.139 | 0.517 | 0.138 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| SKEW | 0.017 | 0.019 | 0.016 | 0.005 |
| | (0.005) | (0.000) | (0.284) | (0.342) |
| KURT | 0.007 | 0.000 | 0.028 | 0.012 |
| | (0.027) | (0.856) | (0.000) | (0.000) |
| $Q_{50}$ | 0.060 | 0.042 | 0.022 | 0.000 |
| | (0.000) | (0.012) | (0.188) | (0.972) |
| PreVol | | 0.008 | | 0.003 |
| | | (0.000) | | (0.175) |
| ARET | | -0.001 | | -0.004 |
| | | (0.567) | | (0.000) |
| TA | | -0.002 | | 0.000 |
| | | (0.000) | | (0.788) |
| B/M | | 0.001 | | 0.004 |
| | | (0.555) | | (0.021) |
| Numan | | 0.000 | | 0.000 |
| | | (0.824) | | (0.032) |
| Turnover | | 0.000 | | 0.000 |
| | | (0.000) | | (0.000) |
| DSEC | | 0.001 | | 0.000 |
| | | (0.009) | | (0.722) |
| Intan_asset | | -0.009 | | 0.021 |
| | | (0.000) | | (0.000) |
| EarnVol | | 0.103 | | 0.146 |
| | | (0.000) | | (0.000) |
| \|ACC\| | | -0.015 | | -0.007 |
| | | (0.026) | | (0.437) |
| Loss | | 0.002 | | 0.003 |
| | | (0.279) | | (0.024) |
| Adj. $R^2$ | 0.080 | 0.164 | 0.083 | 0.082 |
| N | 25,216 | 17,677 | 25,216 | 17,677 |

TABLE 8 (Continued)

**Panel B:** Relation with earnings risk components

$$Outcome = \gamma_0 + \gamma_1 IQR + \gamma_{31} UP + \gamma_{32} DOWN + \gamma_4 Q_{50} + \sum_k \delta_k CONTROL_k + \varepsilon.$$

| | \|FE\| | | \|Exclusions\| | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| IQR | 0.303 | 0.131 | 0.497 | 0.130 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| UP | 0.004 | 0.009 | 0.023 | 0.012 |
| | (0.468) | (0.004) | (0.002) | (0.004) |
| DOWN | 0.004 | -0.005 | 0.027 | 0.011 |
| | (0.235) | (0.020) | (0.000) | (0.000) |
| $Q_{50}$ | 0.041 | 0.040 | 0.000 | -0.004 |
| | (0.013) | (0.018) | (0.986) | (0.823) |
| | | | | |
| Controls | No | Yes | No | Yes |
| Adj. $R_2$ | 0.079 | 0.163 | 0.083 | 0.081 |
| N | 25,216 | 17,677 | 25,216 | 17,677 |

The table is based on a sample of 25,216 firm-year observations over the period 1993 to 2009. Panels A and B report OLS estimates from regressing the *Outcome* variable on our risk metrics and a set of controls. In the exclusions results, *Outcome* is defined as the absolute value of analysts' total exclusions (|Exclusions|). In the forecast error results, *Outcome* is defined as the absolute value of analysts' forecast errors (|FE|). The control variables include Prevol, ARET, TA, B/M, Numan, Turnover, DSEC, Intan_asset, EarnVol, |ACC| and Loss. Year fixed effects are included but not reported. P-values (in parentheses) are based on clustered standard errors by firm and year. Reported $R^2$s are adjusted $R^2$s incremental to those obtained when only year fixed effects are included in the model. |Exclusions| is defined as | $Actual^{IBES}$ – $Actual^{GAAP}$ | and |FE| as |$Actual^{IBES}$ – $Forecast^{IBES}$|. IQR, SKEW, KURT, UP and DOWN are out-of-sample estimates of the spread, skewness, kurtosis, upside risk and downside risk in future earnings. $Q_{50}$ is the median forecast of earnings. Control variables are defined in Appendix A.