The Higher Moments of Future Return on Equity*

Woo-Jin Chang
INSEAD
Boulevard de Constance
77305, Fontainebleau
France
woo-jin.chang@insead.edu
+33 (0)1 60 72 92 48 (Telephone)

Steven J. Monahan
INSEAD
Boulevard de Constance
77305, Fontainebleau
France
steven.monahan@insead.edu
+33 (0)1 60 72 92 14 (Telephone)

Amine Ouazad
INSEAD
Boulevard de Constance
77305, Fontainebleau
France
amine.ouazad@insead.edu
+33 (0)1 60 72 48 49 (Telephone)

September 2013

Keywords: Book-to-price ratio, credit ratings, earnings-to-price ratio, forecasting, higher moments, quantile regressions, return on equity, valuation multiples

JEL classification codes: C21, C53, G17, M41

* We benefited from conversations with Lucie Tepla and the comments of Daniel Bens, Richard Frankel, Joseph Gerakos, Bjorn Jorgensen, Maria Ogneva, Stephen Penman, Gil Sadka and seminar participants at INSEAD's Finance Brownbag, the 2013 INSEAD Accounting Symposium, the 2013 London Business School Accounting Symposium, and The Ohio State University.

The Higher Moments of Future Return on Equity

Abstract

We use quantile regressions to evaluate the higher moments of future return on equity, ROE. First, we evaluate the in-sample relations between current firm-level attributes and the moments of lead ROE. We show that: (1) as current profitability increases lead ROE tends to increase, become more disperse, and more leptokurtic; (2) loss firms tend to have lower, more disperse, and more left-skewed lead ROE; (3) as accruals increase lead ROE tends to decrease and become more disperse; and, (4) firms with higher leverage and/or lower payout ratios tend to have greater dispersion in lead ROE. Second, we show that the in-sample relations generate reliable out-of-sample estimates of the probability of a loss as well as the standard deviation, skewness, and kurtosis of lead ROE. Moreover, when compared to estimates obtained via alternative approaches, our out-of-sample estimates: (1) always contain incremental information content and (2) are typically more reliable. Finally, we evaluate the role that the higher moments of future ROE play in determining valuation multiples and credit ratings.

1. Introduction

Return on equity, ROE, is a key economic variable. Hence, numerous studies in accounting and finance evaluate the relation between ROE and firm-level attributes; and, there is a large literature that evaluates different approaches for forecasting ROE.[1]

A limitation of extant studies is that virtually all of them focus (implicitly or explicitly) on expected—i.e., the mean of—future ROE.[2] Although the mean is an important moment of the distribution, higher moments are important too. For example, as the variance of future firm-level payoffs increases so does the likelihood of default (e.g., Merton [1974]); and, a number of theoretical asset-pricing models imply that equity prices depend on higher *firm-level* moments of future payoffs (e.g., Johnson [2004], Brunnermeier et al. [2007], Mitton and Vorkink [2007], and Barberis and Huang [2008]).

As discussed above and in section two, a number of empirical questions can be asked about the economic relevance of the higher moments of future ROE. However, meaningful answers to these questions cannot be obtained unless there is a reliable approach for estimating higher moments. Hence, in this study, we focus on two more-fundamental questions. First, what are the in-sample relations between firm-level attributes and the higher moments of lead ROE? Second, can these in-sample relations be used to develop reliable out-of-sample estimates of the higher moments of lead ROE?

We use a novel research design that is based on quantile regressions. As discussed in section three, quantile regressions are particularly appropriate in our setting for two reasons.

[1] Examples of studies that focus on ROE or related earnings metrics include: Freeman et al. [1982], Fairfield and Yohn [2000], Fama and French [2000, 2006], Nissim and Penman [2001, 2003], Banker and Chen [2006], Hou and Robinson [2006], Soliman [2008], Fairfield et al., [2009], Esplin et al. [2011], Hou et al. [2012], Li and Mohanram [2012], Penman [2012], and Gerakos and Gramercy [2013]. This list is not exhaustive and we apologize in advance to the authors of the papers that are omitted.

[2] A recent study by Konstantinidi and Pope [2012], who evaluate the quantiles of return on assets, is a notable exception. We elaborate on the relative contributions of the two studies in section two.

First, as discussed in Buchinsky [1998], when a regression is estimated for the q^{th} quantile, the coefficient on a particular regressor is a consistent estimate of the change in the q^{th} conditional quantile of the dependent variable given a marginal change in the regressor of interest. Hence, we use the coefficients obtained from regressions estimated for a sequence of quantiles to infer the relation between firm-level attributes and the location and shape of the distribution of lead ROE.

Second, the fitted value obtained from a regression estimated for the q^{th} quantile is a consistent estimate of the q^{th} conditional quantile of the dependent variable. Hence, for each firm-year in our sample we calculate out-of-sample estimates of the conditional quantiles of lead ROE for a sequence of quantiles. Next, we combine these estimates to form an estimate of the conditional cumulative distribution function, cdf, of lead ROE. Finally, we use the cdf to infer the conditional probability that lead ROE is negative—i.e., the probability of a future loss—as well as the conditional mean, standard deviation, skewness, and kurtosis of lead ROE.

In our first set of analyses, we document the in-sample relations between the moments of lead ROE and firm-level attributes that fall into two categories: (1) attributes of current ROE and (2) attributes of current financial policy. Regarding the relations between lead ROE and the attributes of current ROE, three results are noteworthy. First, as current ROE increases lead ROE tends to increase, become more disperse, and more leptokurtic—i.e., fat-tailed. Hence, although higher current ROE implies higher expected lead ROE, it also implies riskier lead ROE. Second, loss firms tend to have lower, more disperse, and more left-skewed lead ROE; and, as current losses become larger in magnitude, lead ROE tends to decrease and become more disperse. Hence, current losses are associated with lower, riskier lead ROE. Finally, as accruals increase lead ROE tends to decrease and become more disperse. This implies that accruals, which are a measure of growth, are positively associated with the riskiness of lead ROE.

Regarding the relations between the moments of lead ROE and firms' financial policies, we show that the distribution of lead ROE becomes more disperse and more leptokurtic as current leverage increases. This is consistent with the well-known result described in Modigliani and Miller [1958]: equity becomes riskier as leverage increases. We also show that dividend-paying firms have distributions of lead ROE that are less disperse and less left-skewed. This supports results in Brav et al. [2005], who show that 70 percent of managers they surveyed view the "…stability and sustainability of future earnings" as a key determinant of payout-policy. It also supports the dividend-smoothing hypothesis described in Lintner [1956].

In our second set of analyses we evaluate whether the in-sample relations described above can be used to develop reliable out-of-sample predictions. We begin by conducting tests at the firm-year level in which we evaluate our estimates of the probability of a future loss, Q_PROB, and our estimates of the standard deviation of lead ROE, Q_STD. We show that Q_PROB and Q_STD are positively associated with future realized losses and future unsigned forecast errors, respectively.[3] These associations remain after controlling for alternative estimates.[4] Moreover, our quantile-based estimate of the standard deviation of lead ROE is a more reliable predictor of future unsigned forecast errors than the alternative estimates that we evaluate.

Next, we conduct tests at the industry-year level. We do this because industry attributes are relevant *per se* and because realized moments are not observable at the firm-year level. Hence, it is not possible to use firm-level tests to obtain *direct* evidence about reliability.

[3] A firm's unsigned forecast error in year t+h equals the absolute value of the difference between the firm's realized ROE in year t+h and the year t estimate of the mean of the firm's ROE in year t+h.
[4] We elaborate on our alternative estimates and how we calculate them in section six.

However, it is possible to use industry-level tests. In particular, as discussed in section six, we use the law of total moments described in Brillinger [1969] to construct estimates of industry-year moments from contemporaneous estimates of firm-year moments. We then compare the industry-level estimates to the future realized industry-level moments.

The results of the industry-level tests lead to the conclusion that the quantile-based estimates are reliable out-of-sample predictors. In particular, each of the estimated industry-level moments is positively associated with its future realized industry-level counterpart and these associations remain after controlling for alternative estimates. Moreover, with the exception of our estimate of the kurtosis of lead ROE, our quantile-based estimates of the moments of lead industry-level ROE are more reliable predictors of the future realized moments than any of the alternative estimates that we evaluate.

Although our primary research objective is to document the in-sample relations described above and to show that they can be used to generate reliable out-of-sample estimates, we also provide initial evidence about the economic relevance of the higher moments of ROE. Specifically, in our final set of analyses we evaluate the relations between the higher moments of lead ROE and two important economic variables: (1) valuation multiples and (2) credit ratings.

We show that the current, firm-level book-to-price and earnings-to-price ratios are both negatively associated with our quantile-based estimates of the standard deviation and skewness of lead ROE and positively associated with our quantile-based estimate of the kurtosis of lead ROE. These results remain after controlling for the moments of historical stock returns. Hence, we show that, *ceteris paribus*, investors place higher prices on securities with future payoffs that are relatively volatile, positively-skewed, and thin-tailed, which is consistent with extant analytical and empirical results (e.g., Johnson [2004], Ang et al. [2006, 2009], Brunnermeier et al. [2007], Mitton and Vorkink [2007], Barberis and Huang [2008], Boyer et al. [2010], and Conrad et al. [2013]).

In our analyses of credit ratings we show that, *ceteris paribus*, firms with worse credit ratings tend to have lead ROE that is relatively volatile, negatively-skewed, and fat-tailed. These results remain after controlling for other determinants of credit ratings and the moments of historical stock returns. The association between credit ratings and the standard deviation of lead ROE is consistent with results in Merton [1974], who shows that the probability of default is increasing in volatility. The results related to kurtosis (skewness) are consistent with arguments made in practice about the implications of tail risk for optimal bond-portfolio formation (e.g., see chapters 16 and 17 of Dynkin et al. [2007]). These results are also intuitive as they suggest that when extreme negative payoffs are more likely ratings agencies assume a higher expected loss.

We make four contributions to the extant literature. First, we develop a general approach based on quantile regressions that can be used to study the economic relevance of the higher moments of future ROE as well as other variables. Second, we provide pertinent evidence about the relations between the higher moments of lead ROE and key firm-level attributes. Third, we show that our methodology yields reliable out-of-sample estimates of probability of a future loss as well as the standard deviation, skewness, and kurtosis of lead ROE. These are nontrivial results given the difficulty involved in predicting higher moments; and, the particular challenges related to predicting higher moments of ROE, which is a function of aggregate data that are reported fairly infrequently. Finally, we provide initial evidence about the relevance of the higher moments of future ROE within the context of equity and credit valuation.

2. Related Literature

The higher moments of firm-level payoffs are potentially relevant in various economic contexts. Consider equity valuation, a number of extant analytical models suggest that firm-level moments matter. For example, Johnson [2004] shows that, for levered firms, the option value of equity is increasing in the volatility of future firm-level payoffs. On the other hand, Merton [1987] shows that, when incomplete information leads to market segmentation, equity prices are a decreasing function of firm-level volatility. The skewness of future firm-level payoffs may also affect equity prices. For instance, Brunnermeier et al. [2007] show that when optimistic "…investors hold beliefs that optimally trade off the *ex ante* benefits of anticipatory utility against the *ex post* costs of basing investment decisions on biased beliefs," equity prices are increasing in firm-level skewness. A similar result is obtained by Barberis and Huang [2008], who assume that investors make decisions according to cumulative prospect theory. Finally, Mitton and Vorkink [2007] show that, when rational investors have heterogeneous preferences, equity prices are increasing in the skewness of future firm-level payoffs.

Firm-level moments are also potentially relevant in the context of debt valuation. As shown in Merton [1974], debt values are decreasing in firm-level volatility. He obtains this result by assuming that asset returns are normally distributed; hence, higher moments do not matter in his model. However, if the assumption of normality is relaxed, higher moments are likely relevant (e.g., Dynkin et al. [2007]). In particular, given debt holders face relatively high exposure to downside risk while benefitting little from positive shocks, they should assign lower values to debt instruments issued by firms with future payoffs that exhibit negative skewness or positive kurtosis.

In addition to equity and credit valuation, there are numerous other contexts in which the higher moments of future firm-level or industry-level payoffs are potentially relevant. For example, Givoly and Hayn [2000] argue that the skewness of firm-level earnings relative to the skewness of firm-level cash flows is an indicator of conditional conservatism. Studies in the industrial organization literature (e.g., Klepper [1996]) evaluate the causes and consequences of industry-level entry, exit, growth and innovation, which are likely associated with the *ex ante* distribution of industry-level payoffs.

The above implies there are a number of interesting empirical questions regarding the economic relevance of the higher moments of future ROE. Before these questions can be addressed, however, a reliable approach for estimating the higher moments of future ROE must be developed. Hence, our primary research objective is to develop an approach based on quantile regressions and to demonstrate that it yields reliable out-of-sample predictions. We also evaluate the in-sample relations between the estimated moments and observable firm-level attributes; and, we provide initial evidence about the relation between our out-of-sample estimates and two key economic variables: (1) valuation multiples and (2) credit ratings.

Regarding our research objective, it bears mentioning that forecasting higher moments, especially moments higher than the variance, is difficult. The reason for this is that skewness and kurtosis relate to rare events. For example, a firm, or a set of similar firms, with high *ex ante* skewness (kurtosis) may have exhibited low *historical* skewness (kurtosis). Hence, out-of-sample predictions of future skewness (kurtosis) obtained from historical data may be unreliable. Moreover, forecasting the higher moments of firm-level ROE presents additional challenges given that ROE is the ratio of two summary numbers that are reported fairly infrequently.

Consequently, whether we can develop reliable out-of-sample estimates of the moments of future firm-level ROE is an empirical question.

Finally, it is important that we compare our study to an earlier study by Konstantinidi and Pope [2012] (KP hereafter). KP evaluate the quantiles of lead return on assets. The key difference between the two studies is that KP primarily focus on the extreme quantiles whereas we develop a general approach for evaluating and estimating the moments of the distribution. This is important for several reasons. First, the moments (as opposed to the extremes) of the distribution are the relevant economic constructs of interest. Moreover, different analytical assumptions lead to different predictions about which moments matter and how they matter. Second, and related to the first point, interpreting the extreme quantiles is difficult because they are a function of the variance, skewness, and kurtosis of the distribution. Finally, by estimating the moments we are able to draw clear-cut and in-depth inferences about the: (1) in-sample relations between current, firm-level attributes and the higher moments of future ROE and (2) reliability of our out-of-sample estimates.

3. Research Design

We begin by providing a general overview of quantile regressions and how to use them to evaluate and estimate the conditional moments of a random variable. Next, we provide details about how we model the relation between firm-level attributes and the higher moments of lead ROE. Please note that in this section we focus on providing an intuitive description of quantile regressions. We relegate the discussion of technical details to Appendix A.

3.1 General Overview of Quantile Regressions

As discussed in Buchinsky [1998], the estimation of a quantile regression involves choosing the coefficient vector $B^q = \beta_0^q, \ldots, \beta_k^q$ that solves the minimization problem shown in equation (1).

$$\underset{B^q=\beta_0^q,\ldots,\beta_k^q}{\operatorname{argmin}} \frac{1}{N}\left\{ \sum_{i:y_{i,t} \geq \sum_{j=0}^{k} \beta_j^q x_{i,t-h,j}} q\left| y_{i,t} - \sum_{j=0}^{k} \beta_j^q x_{i,t-h,j} \right| + \sum_{i:y_{i,t} < \sum_{j=0}^{k} \beta_j^q x_{i,t-h,j}} (1-q)\left| y_{i,t} - \sum_{j=0}^{k} \beta_j^q x_{i,t-h,j} \right| \right\} \qquad (1)$$

In equation (1), $y_{i,t}$ is the year t—i.e., the lead—value of the dependent variable for observation $i \in [1,N]$, $x_{i,t-h,j}$ is the year t-h—i.e., the current—value of the j^{th} independent variable ($j \in [0,k]$) for observation i, and $q \in (0,1)$ denotes the q^{th} quantile.

Equation (1) has a similar structure as the minimization problem underlying an ordinary least squares, OLS, regression. For example, the objective function involves choosing regression coefficients that minimize a function of the residuals. However, unlike the OLS minimization problem in which equal weight is put on each of the squared residuals, the estimation of a quantile regression involves assigning weights that depend on the *sign* of the residual. In particular, the weight put on a positive residual is q÷(1-q) orders of magnitude of the weight put on a negative residual. This implies that the coefficient vector B^q is chosen so that for each positive residual there are q÷(1-q) negative residuals; and, consequently, (1-q)×100 (q×100) percent of the residuals will lie above (below) the fitted value, which equals $\sum_{j=0}^{k} \beta_j^q x_{i,t-h,j}$. Hence, similar to OLS regressions, which yield fitted values that are consistent estimates of the conditional mean of $y_{i,t}$, quantile regressions yield fitted values that are consistent estimates of the qth conditional quantile of $y_{i,t}$, which we refer to as $QUANT_q(y_{i,t} \mid \cdot)$.[5]

The coefficients obtained from a quantile regression can also be interpreted in a similar manner as the coefficients obtained from an OLS regression. In particular, as discussed in Buchinsky [1998], β_j^q is a consistent estimate of $\frac{\partial QUANT_q(y_{i,t} \mid \cdot)}{\partial x_{i,t-h,j}}$. Hence, the coefficients obtained from a quantile regression reflect marginal effects. However, unlike the coefficients obtained from an OLS regression, which equal the change in the conditional mean of $y_{i,t}$ given a marginal change in $x_{i,t-h,j}$, coefficients obtained from a quantile regression equal the change in the qth conditional quantile of $y_{i,t}$ given a marginal change in $x_{i,t-h,j}$.

The facts described above have two important implications. First, the fact that the coefficients obtained from a quantile regression reflect marginal effects implies that we can use them to infer the association between $x_{i,t-h,j}$ and higher moments of $y_{i,t}$. To do this we begin by solving the minimization problem shown in equation (1) for a sequence of Q quantiles.[6] This yields Q estimates of β_j^q. Next, we evaluate how β_j^q varies with q and we infer the effect of $x_{i,t-h,j}$ on the moments of $y_{i,t}$. For example, suppose that $\beta_j^q = c > 0$ for all q. This implies that as $x_{i,t-h,j}$ increases all of the quantiles shift to the right by an equal amount. Hence, $x_{i,t-h,j}$ is associated with the location, but not the shape, of the distribution of $y_{i,t}$. On the other hand, suppose that for β_j^q is positive for all q and an increasing function of q. Hence, although increases in $x_{i,t-h,j}$ lead to increases in all of the quantiles of $y_{i,t}$, the upper quantiles increase by a larger amount. This implies that $x_{i,t-h,j}$ is positively associated with both the conditional mean and

---

[5] A formal discussion of this result is provided in Buchinsky [1998].
[6] Note that the solutions for the Q quantiles are estimated jointly.

conditional variance of $y_{i,t}$—i.e., as $x_{i,t-h,j}$ increases the conditional distribution of $y_{i,t}$ shifts to the right and becomes more disperse.

Second, the fact that quantile regressions yield fitted values that are consistent estimates of $QUANT_q(y_{i,t} \mid \cdot)$ implies that we can use the fitted values to develop out-of-sample estimates of the moments of $y_{i,t+h}$. To do this we begin by solving the minimization problem shown in equation (1) for a sequence of Q quantiles. This yields Q estimates of the coefficient vector B^q. These Q estimates reflect the *in-sample* relations between the conditional quantiles of the dependent variable measured at *year t* and the *year t-h* values of the independent variables. However, we want *out-of-sample* estimates of the *year t+h* quantiles. Hence, we combine the *year t* values of the independent variables with the estimated coefficient vector to predict the year t+h q^{th} conditional quantile of the dependent variable—i.e., $\sum_{j=1}^{k} \beta_j^q x_{i,t,j} = Q\hat{U}ANT_q(y_{i,t+h} \mid \cdot)$.

Next, we combine the Q predicted quantiles to arrive at a discrete estimate of the conditional cdf of $y_{i,t+h}$. Finally, we use standard statistical formulas to impute the conditional moments of $y_{i,t+h}$ from the cdf.

3.1 Modeling the Higher Moments of ROE

For each estimation year EY we assume the following linear relation between the q^{th} conditional quantile of ROE for year t and firm-level attributes measured at year t-h. We refer to year t as the *lead* year and we refer to year t-h as the *current* year.

$$\begin{aligned} QUANT_q(ROE_{i,t} \mid \cdot) = {} & \beta_{0,EY}^q + \beta_{1,EY}^q ROE_{i,t-h} + \beta_{2,EY}^q LOSS_{i,t-h} + \beta_{3,EY}^q (ROE_{i,t-h} \times LOSS_{i,t-h}) \\ & + \beta_{4,EY}^q ACC_{i,t-h} + \beta_{5,EY}^q LEV_{i,t-h} + \beta_{6,EY}^q PAYER_{i,t-h} + \beta_{7,EY}^q PAYOUT_{i,t-h} \end{aligned} \quad (2)$$

The variables in equation (2) are described in the table shown below.

Variable Name	Description
$ROE_{i,t}$	Earnings of firm i during year t divided by firm i's year t-h equity book value
$ROE_{i,t-h}$	Earnings of firm i during year t-h divided by firm i's year t-h equity book value
$LOSS_{i,t-h}$	An indicator variable that equals one (zero) if $ROE_{i,t-h} < 0$ ($ROE_{i,t-h} \geq 0$)
$ACC_{i,t-h}$	Accruals reported by firm i during year t-h divided by firm i's year t-h equity book value
$LEV_{i,t-h}$	Total assets of firm i for year t-h divided by firm i's year t-h equity book value
$PAYER_{i,t-h}$	An indicator variable that equals one (zero) if $PAYOUT_{i,t-h} > 0$ ($PAYOUT_{i,t-h} = 0$)
$PAYOUT_{i,t-h}$	Dividends paid by firm i during year t-h divided by firm i's year t-h equity book value

Our model is similar to the model used by Hou et al. [2012], who focus on forecasting the mean of ROE. However, there are two differences. First, Hou et al. [2012] do not deflate by equity book value. Second, Hou et al. [2012] do not include the interaction term $ROE_{i,t-h} \times LOSS_{i,t-h}$.

The motivation for the independent variables in equation (2) is straightforward. First, it is well-known (e.g., Freeman et al. [1982]) that ROE is persistent; hence, we include $ROE_{i,t-h}$ in our model. Second, there is ample evidence (e.g., Basu [1997]) that losses follow a different time-series process than profits; thus, we allow the coefficient on $ROE_{i,t-h}$ to vary with the sign of $ROE_{i,t-h}$. Third, evidence provided by Sloan [1996] implies that accruals are less persistent that cash flows. Consequently, we control for the portion of year t-h ROE that is attributable to year t-h accruals, $ACC_{i,t-h}$. Finally, well-known results in finance (e.g., Lintner [1956], Modigliani and Miller [1958], and Miller and Rock [1985]) show that firms' capital structure and payout policies are associated with the level and dispersion of ROE. Hence, we include $LEV_{i,t-h}$, $PAYER_{i,t-h}$, and $PAYOUT_{i,t-h}$ in our model.

In addition to being intuitively appealing and comparable to extant models such as that used by Hou et al. [2012], our model has two advantages. First, it is parsimonious and tractable. Second, it is superior to a number of more elaborate models. In particular, we evaluate models in which we add the following variables to equation (2): the log of sales, $SIZE_{i,t-h}$; an indicator for extreme ROE that equals one (zero) if $ROE_{i,t-h}$ is (is not) in the top or bottom tenth percentile of the annual distribution, $XTRM\_ROE_{i,t-h}$; the interaction between $XTRM\_ROE_{i,t-h}$ and $ROE_{i,t-h}$; the lagged change in earnings deflated by equity book value, $\Delta ROE_{i,t-h}$; an indicator for extreme changes in ROE that equals one (zero) if $\Delta ROE_{i,t-h}$ is (is not) in the top or bottom tenth percentile of the annual distribution, $XTRM\_\Delta ROE_{i,t-h}$; the interaction between the $XTRM\_\Delta ROE_{i,t-h}$ and $ROE_{i,t-h}$; the interaction between $XTRM\_\Delta ROE_{i,t-h}$ and $\Delta ROE_{i,t-h}$; the measure of unconditional conservatism described in Penman and Zhang [2002], $CONS_{i,t-h}$; the interaction between $CONS_{i,t-h}$ and $ROE_{i,t-h}$; and, various combinations of these aforementioned variables. In a set of untabulated results we show that none of these models generate better out-of-sample estimates than the estimates derived from equation (2). Moreover, adding these additional variables to equation (2) does not change the tenor of our results regarding the in-sample relations between the independent variables shown in equation (2) and the higher moments of lead ROE.

Our research design involves the following three steps. First, for each estimation year EY we obtain estimates of the coefficient vector $B^q_{EY} = \beta^q_{0,EY}, \ldots, \beta^q_{7,EY}$ for 150 different values of q $\in$ (0,1).[7] To obtain the coefficient vector for a particular value of q we solve the minimization problem shown in equation (1). When doing this we use a mix of time-series and cross-sectional data (i.e., panel data). We require that each panel contains at least five years of data; however, we never use more than ten years of data to construct a panel. For example, suppose the estimation year is 1990 (i.e., EY = 1990) and the forecast horizon is 3 (i.e., h = 3), we use values

[7] The 150 coefficient vectors are estimated jointly. The 150 values of q are in sequential order and all the pairs of consecutive values of q are equidistant. The number 150 is a function of our sample size and number of covariates. It represents the maximum number of quantile regressions that we can estimate while guaranteeing that the numerical estimates converge. As shown in Appendix A, as the sample size increases and the number of quantile regressions increases, the estimates of the predicted moments converge in probability to the moments of lead ROE.

of the dependent variable that fall between 1981 and 1990 and we use values of the independent variables that fall between 1978 and 1987. We include a firm in the panel if it has at least one valid observation during the relevant time span.

Second, we evaluate the in-sample relations between the higher moments of lead ROE and firm-level attributes. For each estimation year EY, value of q, and firm-level attribute j we obtain the relevant coefficient estimate (i.e., $\beta^{q}_{j,EY}$) and we compute the average of $\beta^{q}_{j,EY}$, which we refer to as $\beta^{q}_{j,AVG}$. Next, assuming a lag length of ten, we calculate the Newey-West adjusted standard error of $\beta^{q}_{j,AVG}$; and, we use the standard error to form a 95 percent confidence interval around $\beta^{q}_{j,AVG}$. We then graph $\beta^{q}_{j,AVG}$ and its confidence interval on q.

Finally, we develop our out-of-sample estimates. For a particular year t we obtain the contemporaneous (i.e., EY = t) estimated coefficient vector for each of the 150 values of q. We then predict the q^{th} conditional quantile of $ROE_{i,t+h}$ by calculating the inner product of the coefficient vector and a vector containing the contemporaneous (i.e., year t) values of the independent variables for firm i. Next, we combine the 150 predicted conditional quantiles to form our estimate of the conditional distribution of $ROE_{i,t+h}$, which we use to infer the conditional probability of a loss and the other conditional moments.[8]

4. Overview of Samples and Descriptive Statistics

In this section we briefly describe the estimation sample and prediction sample; and, we discuss descriptive statistics for the estimation sample. For additional details regarding our sample construction algorithm and variable definitions, please refer to Appendix B.

[8] We define the conditional probability of a loss as the largest value of q for which $\widehat{QUANT}_q\left(ROE_{i,t+h} \mid \cdot\right) < 0$.

For each forecast horizon $h \in [1,5]$, we form two samples: (1) the estimation sample and (2) the prediction sample. The estimation sample contains observations that are used to estimate the coefficients shown in equation (2). The prediction sample contains observations for which we develop out-of-sample, firm-level predictions of the cdf of lead ROE.

In Panel A of Table One we provide descriptive statistics for the estimation sample pertaining to a one-year forecast horizon (i.e., h = 1). Descriptive statistics for estimation samples pertaining to the other forecast horizons are available upon request. The mean (median) of $ROE_{i,t}$ is 0.031 (0.101). Twenty-four percent of the sample observations have negative $ROE_{i,t-1}$. The mean (median) of $ACC_{i,t-h}$ is -0.059 (-0.054). $LEV_{i,t-h}$ has a mean (median) of 2.408 (1.970), 43.8 percent of the observations pay dividends, and the average payout ratio is 0.024.

Panel B of Table One contains the correlation structure of the variables shown in equation (2). Pearson (Spearman) correlations are shown above (below) the diagonal. The correlations shown in the table equal the average of the annual correlations. The t-statistics equal the average correlation divided by its temporal standard error. When calculating the temporal standard error we make the Newey-West adjustment assuming a ten-year lag length. We tabulate results for the estimation sample that pertains to a one-year forecast horizon. Results for estimation samples pertaining to the other forecast horizons are available upon request.

Several correlations warrant discussion. First, the Pearson (Spearman) correlation between lead ROE (i.e., $ROE_{i,t}$) and current ROE (i.e., $ROE_{i,t-1}$) is 0.60 (0.70); hence, shocks to ROE have high persistence. Second, firms that are currently experiencing losses have lower lead ROE; in particular, the Pearson (Spearman) correlation between $ROE_{i,t}$ and $LOSS_{i,t-1}$ is -0.42 (-0.43). Third, the Pearson (Spearman) correlation between lead ROE and current accruals (i.e., $ACC_{i,t-1}$) is 0.11 (0.12), which implies that accruals are less persistent than cash flows. Current leverage (i.e., $LEV_{i,t-1}$) is uncorrelated with lead $ROE_{i,t}$. However, the Pearson (Spearman) correlation between lead ROE and $PAYER_{i,t-1}$ is 0.20 (0.23); and, the Pearson (Spearman) correlation between the current payout ratio (i.e., $PAYOUT_{i,t-1}$) and lead ROE is 0.22 (0.31).

5. Analyses of In-sample Relations

In this section we describe the relation between the independent variables shown in equation (2) and the moments of lead ROE. We use graphical evidence. In particular, for each estimation year EY, value of q, and firm-level attribute j we obtain the relevant coefficient estimate—i.e., $\beta^{q}_{j,EY}$. We then compute the average of $\beta^{q}_{j,EY}$ across estimation years, which we refer to as $\beta^{q}_{j,AVG}$, and the temporal standard error of $\beta^{q}_{j,AVG}$. When calculating the temporal standard error we make the Newey-West adjustment assuming a ten-year lag length. Next, we use the standard error to calculate a 95 percent confidence interval around the average; and, we graph $\beta^{q}_{j,AVG}$ and its confidence interval on q. For comparative purposes, we also graph the average coefficient, which we refer to as $\beta^{OLS}_{j,AVG}$, and the 95 percent confidence interval obtained from an OLS regression. We provide graphs of coefficients that relate to a one-year forecast horizon (i.e., h = 1). Graphs of coefficients that relate to other forecast horizons are available upon request.

The graphs presented in this section are based on regressions that are estimated on "de-medianed" independent variables. In particular, we set each of the independent variables equal to the difference between its raw value and its median value for the relevant panel. This de-medianing makes it easier to interpret the coefficient on the constant term. In particular, when we use de-medianed independent variables the estimated constant for quantile q—i.e., $\beta^{q}_{0,EY}$—equals the conditional q[th] quantile for the "typical" observation. That is, the observation for which each of the independent variables is equal to the median of that variable for the panel.

It is important to note that de-medianing only affects the estimate of the constant term and has no effect on the estimates of the slope coefficients. That is, the estimates of $\beta_1^q, \ldots, \beta_7^q$ obtained from estimating equation (2) on the de-medianed data are identical to the estimates of $\beta_1^q, \ldots, \beta_7^q$ obtained from estimating equation (2) on the original data. It is also important to note that we only use the de-medianed data to generate the graphs presented in this section: Our out-of-sample estimates are based on regressions estimated on the raw data.

We show the graph of the constant term, $\beta_{0,AVG}^q$, in Figure One. As discussed above, $\beta_{0,AVG}^q$ ($\beta_{0,AVG}^{OLS}$) equals the conditional q[th] quantile (conditional mean) of lead ROE for the "typical" observation.[9] As shown in Figure One, the typical observation has median (mean) lead ROE of 0.095 (0.069). Untabulated results show that the interquartile range of lead ROE for the typical observation is 0.126. Moreover, lead ROE for the typical observation is negative for all values of q that are less than 0.23—i.e., there is a 23 percent probability that the typical observation will experience a loss in year t+1.

Figure Two contains the graph of $\beta_{1,AVG}^q$, which is the coefficient on current ROE (i.e., $ROE_{i,t-1}$). We provide an in-depth discussion of this graph so that we can: (1) discuss the specific relation between current ROE and the moments of lead ROE and (2) make some general points about how to interpret the graphs of the remaining coefficients.

A natural starting point is to determine the relation between current ROE and the location of the distribution of lead ROE. To do this we evaluate the coefficient $\beta_{1,AVG}^{0.50}$ ($\beta_{1,AVG}^{OLS}$), which

[9] The graphs of $\beta_{0,AVG}^q$ and $\beta_{0,AVG}^{OLS}$ obtained from regressions estimated on the raw data are available upon request.

equals the change in the conditional median (mean) of lead ROE given a marginal change in current ROE. $\beta^{0.50}_{1,AVG}$ ($\beta^{OLS}_{1,AVG}$) equals 1.01 (0.85); hence, there is a positive association between current ROE and the median (mean) of lead ROE.

Second, we consider the relation between current ROE and the variance of lead ROE. As shown in Figure Two, $\beta^{q}_{1,AVG}$ is an increasing function of q (i.e., $\Delta\beta^{q}_{1,AVG}/\Delta q > 0$). This implies that as current ROE increases the higher quantiles of lead ROE increase by larger amounts than the lower quantiles—i.e., the distribution of lead ROE spreads out. Hence, there is a positive association between current ROE and the variance of lead ROE.

Finally, we note that: (1) for values of q < 0.80 the relation between $\beta^{q}_{1,AVG}$ and q is concave (i.e., $\Delta\Delta\beta^{q}_{1,AVG}/\Delta q < 0$) but (2) for values of q > 0.80 the relation between $\beta^{q}_{1,AVG}$ and q is convex (i.e., $\Delta\Delta\beta^{q}_{1,AVG}/\Delta q > 0$). Hence, firms with higher current ROE are more likely to have extreme values of lead ROE. That is, these firms have more leptokurtic (i.e., fat-tailed) distributions of lead ROE.

In light of the above, we conclude that firm's with high current ROE tend to have higher lead ROE that is more volatile and more extreme. Hence, although higher current profitability is associated with higher future profitability it also implies greater risk.

Figure Three contains the graph of $\beta^{q}_{2,AVG}$, which is the coefficient on the loss indicator (i.e., $LOSS_{i,t-1}$). The graph illustrates that, *ceteris paribus*, loss firms tend to have lower, more volatile lead ROE. In particular, $\beta^{0.50}_{2,AVG}$ ($\beta^{OLS}_{2,AVG}$) equals -0.01 (-0.07) and $\beta^{q}_{2,AVG}$ is increasing in q. Loss firms are also more likely to experience extreme poor performance. Specifically, the relation between $\beta^{q}_{2,AVG}$ and q is a concave for most values of q. Hence, loss firms have lead ROE that is more left-skewed.

Figure Four contains the graph of $-1\times\left(\beta^{q}_{1,AVG}+\beta^{q}_{3,AVG}\right)$.[10] We are interested in the total relation between current losses and lead ROE; hence, we evaluate the sum of the coefficient on $ROE_{i,t-1}$ (i.e., $\beta^{q}_{1,AVG}$) and the coefficient on the interaction of $ROE_{i,t-1}$ and $LOSS_{i,t-1}$ (i.e., $\beta^{q}_{3,AVG}$). We multiply the coefficients by negative one so that the graph shows the relation between larger losses (i.e., more negative ROE) and the quantiles of lead ROE. The graph illustrates that firms with higher current losses tend to have lower lead ROE. In particular, $-1\times\left(\beta^{0.50}_{1,AVG}+\beta^{0.50}_{3,AVG}\right)$ ($-1\times\left(\beta^{OLS}_{1,AVG}+\beta^{OLS}_{3,AVG}\right)$) equals -0.52 (-0.44). In addition, $-1\times\left(\beta^{q}_{1,AVG}+\beta^{q}_{3,AVG}\right)$ is increasing in q, which implies the magnitude of the current loss is positively associated with the variance of lead ROE.

In Figure Five we graph the relation between the coefficient on current accruals, $\beta^{q}_{4,AVG}$, and q. The results shown on the graph suggest that higher current accruals are associated with lower, riskier lead ROE. In particular, $\beta^{0.50}_{4,AVG}$ ($\beta^{OLS}_{4,AVG}$) equals -0.03 (-0.06): and, $\beta^{q}_{4,AVG}$ is increasing in q.

Figure Six contains the graph of $\beta^{q}_{5,AVG}$, which is the coefficient on current leverage (i.e., $LEV_{i,t-1}$). As the graph shows, current leverage is not associated with the median (mean) of lead ROE. $\beta^{q}_{5,AVG}$ is an increasing function of q, however; hence, current leverage is positively associated with the variance of lead ROE. Moreover, for the lower quantiles of q, the relation between $\beta^{q}_{5,AVG}$ and q is concave; however, for values of q > 0.80 the relation is convex. Thus, firms with high current leverage have more leptokurtic (i.e., fat-tailed) distributions of lead ROE.

---

[10] The confidence intervals shown in Figure Four relate to the standard error of the average of $-1\times\left(\beta^{q}_{1,EY}+\beta^{q}_{3,EY}\right)$ and the average of $-1\times\left(\beta^{OLS}_{1,EY}+\beta^{OLS}_{3,EY}\right)$. That is, we use the standard error of the average of the sum *not* the sum of the standard errors of the averages.

These results are consistent with fundamental theorems in classical finance (i.e., Modigliani and Miller [1958]) that show that equity becomes riskier as leverage increases.

In Figure Seven we show the graph of $\beta^{q}_{6,AVG}$, which is the coefficient on the dividend indicator (i.e., $PAYER_{i,t-1}$). First, $\beta^{OLS}_{6,AVG}$ equals 0.04 and $\beta^{0.50}_{6,AVG}$ equals 0.01. Hence, dividend-paying firms tend to have higher lead ROE; however, the effect primarily relates to the mean. Second, $\beta^{q}_{6,AVG}$ is a decreasing function of q, which implies that dividend-paying firms have less volatile lead ROE. Finally, for most values of q, the relation between $\beta^{q}_{6,AVG}$ and q is convex. This implies that dividend-paying firms are less likely to exhibit extreme poor performance—i.e., the distribution of lead ROE is less left-skewed.

Figure Eight contains the graph of $\beta^{q}_{7,AVG}$, which is the coefficient on $PAYOUT_{i,t-1}$. The graph illustrates that there is a complex relation between current payout ratios and the moments of lead ROE. First, regarding the location of lead ROE, higher current payout implies higher mean but lower median lead ROE. In particular, $\beta^{0.50}_{7,AVG}$ ($\beta^{OLS}_{7,AVG}$) equals -0.03 (0.09). Second, for values of q between 0.10 and 0.90, $\beta^{q}_{7,AVG}$ is a decreasing function of q; however, for values of $q \in \{(0,0.10) \cup (0.90,1.00)\}$, $\beta^{q}_{7,AVG}$ is an increasing function of q. Hence, as the current payout ratio increases the middle 80 percent of the distribution of lead ROE clusters together but the extreme quantiles become more spread out. This implies that firms with high payout ratios tend to exhibit either relatively small or relatively large deviations from the mean of lead ROE. However, these firms rarely exhibit moderate deviations from the mean of lead ROE.

Finally, in Figure Nine we show the pseudo r-squared from each of the quantile regressions and the r-squared from the OLS regression. The pseudo r-squared of a quantile regression measures the impact of the covariates on the ability of the quantile regression to explain the weighted sum of the absolute deviations.[11] (The weighted sum of the absolute deviations is the value of the objective function minimized in equation (1).) The pseudo r-squared is equal to zero if the model's explanatory variables do not explain more of the weighted absolute deviations than a model that contains only a constant term. On the other hand, the pseudo r-squared is equal to one if the model's predictions do not deviate from the realizations. The OLS r-squared is calculated in the usual way; and, it equals the fraction of the variance of ROE explained by the independent variables. The pseudo r-squared and the OLS r-squared are not directly comparable.

The results shown on the graph imply that the covariates significantly improve the model's fit. The lowest pseudo r-squared is approximately 27 percent. The model's fit is better for the smallest quantiles (i.e. for values of q below 0.50).

6. Analyses of Out-of-sample Estimates

In this section we evaluate our out-of-sample estimates of the moments of lead ROE. We begin by discussing descriptive statistics and correlations. Next, we evaluate the reliability of our firm-year estimates of the probability of a future loss and the standard deviation of lead ROE. Finally, we evaluate our predictions of the frequency of losses within an industry and the within-industry-year standard deviation, skewness, and kurtosis of lead ROE. All of the analyses described in this section are based on observations drawn from the prediction sample. Results reported in the tables relate to the out-of-sample forecasts made in year t of the moments of firm-level ROE in year t+1. Untabulated results for forecast horizons (i.e., values of h) between two and five are qualitatively similar and available upon request.

[11] The pseudo r-squared shown in Figure Nine is the standard statistic reported by Stata.

It is important to note that all of the estimates described in this section are out-of-sample. In particular, we develop a year t estimate of a variable in year t+h by combining regression coefficients with firm-specific values of the independent variables. The regression coefficients are obtained from regressions estimated on data that were available on or before the end of year t; and, the firm-specific values of the independent variables are measured at the end of year t.

6.1 Descriptive Statistics and Correlations

We provide descriptive statistics and correlations for the variables shown below.

Variable Name	Description
$Q\_MEAN_{i,t,t+h}$	Year t estimate of the mean of $ROE_{i,t+h}$
$Q\_PROB_{i,t,t+h}$	Year t estimate of the probability that $ROE_{i,t+h} < 0$
$Q\_STD_{i,t,t+h}$	Year t estimate of the standard deviation of $ROE_{i,t+h}$
$Q\_SKEW_{i,t,t+h}$	Year t estimate of the skewness of $ROE_{i,t+h}$
$Q\_KURT_{i,t,t+h}$	Year t estimate of the excess kurtosis of $ROE_{i,t+h}$

As discussed in section three, the variables shown above are inferred from our out-of-sample, firm-level estimates of $\hat{QUANT}_q(ROE_{i,t+h} \mid \cdot)$. Specifically, for firm i in year t we obtain the predicted values of $\hat{QUANT}_q(ROE_{i,t+h} \mid \cdot)$ for all 150 values of q. Next, we calculate the sample mean, standard deviation, skewness, and kurtosis for this "sample" of 150 values; and, we set $Q\_PROB_{i,t,t+h}$ equal to the largest value of q for which $\hat{QUANT}_q(ROE_{i,t+h} \mid \cdot) < 0$.

We define skewness as the ratio of the third central moment to the third power of $Q\_STD_{i,t,t+h}$; and, we calculate kurtosis by subtracting three from the ratio of the fourth central moment to the fourth power of $Q\_STD_{i,t,t+h}$. Hence, we evaluate standardized skewness and excess, standardized kurtosis. We do this for two reasons. First, by standardizing we eliminate the possibility that our measures of skewness and kurtosis are simply redundant measures of the variance. For example, if the standard deviation is high, non-standardized kurtosis will also be high even if the distribution is not leptokurtic. Second, the excess kurtosis of a normally distributed random variable is zero; hence, by providing descriptive statistics about excess kurtosis we can evaluate the extent to which our predicted cdf's differ from the normal distribution.

Panel A of Table Two contains descriptive statistics. Several comments are warranted. First, the average (typical) firm has positive $Q\_MEAN_{i,t,t+1}$; in particular, the mean (median) of $Q\_MEAN_{i,t,t+1}$ is 0.042 (0.097). However, $Q\_MEAN_{i,t,t+1}$ varies considerably across observations; for example, the standard deviation (interquartile range) of $Q\_MEAN_{i,t,t+1}$ is 0.237 (0.210). Moreover, untabulated results show that 27.8 percent of the observations have negative $Q\_MEAN_{i,t,t+1}$.

Second, there is also considerable within-sample variation in $Q\_PROB_{i,t,t+1}$. Although the sample average of $Q\_PROB_{i,t,t+1}$ is 25.1 percent, the predicted probability of a loss in year t+1 for the typical firm is only 9.3 percent. Moreover, untabulated results show that 29.9 percent of the sample observations are predicted to be profitable with probability one (i.e., $Q\_PROB_{i,t,t+1} = 0$). On the other hand, untabulated results show that 26.3 percent of the observations are more likely to suffer a loss than realize a profit (i.e., $Q\_PROB_{i,t,t+1} > 0.50$).

Third, $Q\_STD_{i,t,t+1}$ is large, which implies there is considerable uncertainty about future ROE. To understand the magnitude of $Q\_STD_{i,t,t+1}$ better we evaluate the coefficient of variation, $Q\_CV_{i,t,t+1}$, which equals the ratio of $Q\_STD_{i,t,t+1}$ to $|Q\_MEAN_{i,t,t+1}|$. The mean (median) of $Q\_CV_{i,t,t+1}$ is 2.366 (0.716). Hence, for the average (typical) observation, the standard deviation of lead ROE is more than twice as large as (70 percent of) the mean of lead ROE. Moreover, untabulated results show that the coefficient of variation for 62.4 percent of the observations exceeds 0.50. There is also considerable variation in the degree of uncertainty. For example, the interquartile range of $Q\_CV_{i,t,t+1}$ is 0.975.

Finally, the median of $Q\_SKEW_{i,t,t+1}$ ($Q\_KURT_{i,t,t+1}$) is -0.590 (1.748). Moreover, untabulated results show that 65.2 percent of the observations have distributions of lead ROE that are negatively skewed; and, 82.7 percent of the observations have distributions of lead ROE that are leptokurtic (i.e., fat-tailed). Hence, the typical observation in our sample has lead ROE that is drawn from a fat-tailed distribution with a long left tail. This implies that extreme deviations from the mean occur relatively often and that these deviations are more likely to be negative.

In Panel B of Table Two we show the correlation structure of the variables. The correlations shown in the table equal the average of the annual cross-sectional correlations. The t-statistics equal the ratio of the average correlation to its temporal standard error. When calculating the temporal standard error we make the Newey-West adjustment assuming a ten-year lag length. We discuss the Pearson correlations but the Spearman correlations lead to similar inferences.

Several comments are warranted. First, the Pearson correlations between $Q\_MEAN_{i,t,t+1}$ and $Q\_PROB_{i,t,t+1}$, $Q\_STD_{i,t,t+1}$, $Q\_SKEW_{i,t,t+1}$, and $Q\_KURT_{i,t,t+1}$ are -0.80, -0.62, 0.20, and 0.27, respectively. Hence, firms with high mean lead ROE are less likely to experience a loss, have less volatile earnings, are more likely to experience an extreme deviation from the mean (i.e., lead ROE is more leptokurtic), and positive extreme deviations are more likely than negative extreme deviations (i.e., lead ROE is more positively skewed).

Second, when the probability of experiencing a loss is high, lead ROE is more volatile, the likelihood of experiencing an extreme deviation from the mean is greater, and negative extreme deviations are more likely than positive extreme deviations. In particular, the Pearson correlations between $Q\_PROB_{i,t,t+1}$ and $Q\_STD_{i,t,t+1}$, $Q\_SKEW_{i,t,t+1}$, and $Q\_KURT_{i,t,t+1}$ are 0.69, -0.33, and -0.45, respectively.

Third, the Pearson correlation between $Q\_STD_{i,t,t+1}$ and $Q\_SKEW_{i,t,t+1}$ is -0.26; and, the correlation between $Q\_STD_{i,t,t+1}$ and $Q\_KURT_{i,t,t+1}$ is -0.38. Taken together, these two facts imply that as the variance of lead ROE increases the distribution becomes more (less) negatively (positively) skewed. Finally, the Pearson correlation between $Q\_SKEW_{i,t,t+1}$ and $Q\_KURT_{i,t,t+1}$ is 0.67. Hence, when extreme deviations from the mean are likely, extreme positive deviations are more likely than extreme negative deviations.

6.2 Reliability of Firm-year Estimates

In this section we evaluate the construct validity of our firm-level, out-of-sample estimates of the probability of a future loss and the standard deviation of future ROE. To do this we regress realized future losses, $LOSS_{i,t+h}$, on $Q\_PROB_{i,t,t+h}$ and realized unsigned future forecast errors $|FERR_{i,t+h}|$ on $Q\_STD_{i,t,t+h}$. We also compare each of the quantile-based estimates to one or more alternative estimates. These comparisons are based on results of Vuong [1989] tests and results obtained from multiple regressions.[12]

Regarding the probability of future losses, we consider an alternative estimate that equals the predicted value from a Logit regression of a loss indicator on the same set of independent variables that we include in our quantile regression. We estimate the Logit regressions and compute the out-of-sample Logit predictions using the same data and a similar algorithm as we

[12] All reported regression coefficients equal the average of the coefficients obtained from annual cross-sectional regressions. All reported t-statistics equal the average coefficient divided by its temporal standard error. Vuong statistics are obtained by computing the average of the annual slope coefficients obtained from the regression described on p. 318 of Vuong [1989], and then dividing the average by the temporal standard error of the coefficients. All temporal standard errors reflect the Newey-West adjustment assuming a ten-year lag length. We also estimate pooled cross-sectional regressions in which we include annual fixed effects and calculate clustered standard errors (annual clusters). The results of these regressions, which are available upon request, lead to similar inferences as the results shown in the tables.

use to obtain $Q\_PROB_{i,t,t+h}$. Logit regression results and descriptive statistics regarding the Logit estimates are untabulated but available upon request.

We consider two alternative approaches for estimating the standard deviation of lead ROE. In the first approach we set our year t estimate of the standard deviation of firm i's ROE in year t+h equal to the standard deviation of firm i's ROE during years t-9 to t. We refer to this as the historical firm-level approach. We refer to the second approach as the historical matched-sample approach. In this approach, which is based on the approach described in Larson and Resutek [2013], we match firm-year i,t to a sample of firms drawn from years t-4 to t. We match on three characteristics: size, return on equity, and the change in return on equity. Next, we calculate the standard deviation of ROE for the matched sample and use this as our year t estimate of the standard deviation of firm i's ROE in year t+h. We elaborate on the historical firm-year approach and the historical matched-sample approach in Appendix B.

The results shown in Table Three imply that $Q\_PROB_{i,t,t+1}$ is a reliable predictor of future losses. In columns (1) and (2) we show results of univariate regressions and Vuong tests. These results show that both the quantile-based estimate and the Logit-based estimate have a positive and statistically significant association with future losses. Moreover, per the r-squareds, both explain roughly 32 percent of the variation in $LOSS_{i,t+1}$; and, per the Vuong statistic, neither estimate is more reliable than the other. However, as shown in column (3) both estimates are incrementally informative about future losses.

In Table Four we report the results of tests of the reliability of $Q\_STD_{i,t,t+1}$. The results shown in the first (last) three columns relate to our comparison of $Q\_STD_{i,t,t+1}$ to the estimate obtained from the historical firm-year (matched-sample) approach. $Q\_STD_{i,t,t+1}$ is reliable both on an absolute and relative basis. As shown in columns (1) and (4) it has a positive and significant relation with realized unsigned forecast errors; and, as shown in columns (3) and (6) it contains incremental information content vis-à-vis both of the alternative estimates. Moreover, per the Vuong test results, $Q\_STD_{i,t,t+1}$ is a more reliable predictor of $|FERR_{i,t+1}|$ than either of the alternative estimates. In particular, $Q\_STD_{i,t,t+1}$ explains 21 percent of the variation in $|FERR_{i,t+1}|$ whereas the estimate obtained from the historical firm-year (matched-sample) approach explains none (nine percent) of the variation.

6.3 Reliability of Industry-year Estimates

In this section we develop and evaluate out-of-sample predictions of the frequency of losses within an industry-year and the within-industry-year standard deviation, skewness, and kurtosis of lead ROE. We assign all firms in the prediction sample to industry-years on the basis of their two-digit Standard Industrial Classification codes. We delete industry-years for which the number of industry members is less than ten.

Industry-level tests are relevant for two reasons. First, industry-level attributes are interesting *per se*. For example, a number of studies of industrial organization focus on the causes and consequences of industry differences; and, more generally, industry membership is a common way of characterizing firms, identifying peers, etc. Second, the realized distribution (moments) of ROE is (are) not observable at the firm-year level. Hence, direct tests of reliability cannot be conducted using firm-level data. However, the realized cross-sectional distribution is observable at the industry-year level. Hence, as discussed in section 6.3.2, results of industry-level tests provide evidence about the reliability of our firm-level estimates.

6.3.1 Reliability of Predictions of the Industry-year-level Risk of Future Losses

In Table Five we show results obtained from regressions of the realized frequency of losses for industry IND in year t+1, $LOSS_{IND,t+1}$, on the industry average of the predicted probability in year t of a loss in year t+h obtained from: (1) our quantile regressions and (2) the Logit regressions described on p. 26.[13] The quantile-based estimate, $Q\_PROB_{IND,t,t+h}$ is reliable. As shown in column (1) it has a positive, significant association with $LOSS_{IND,t+1}$; and, as shown in column (3), this association remains after controlling for the industry average of the predicted probabilities implied by the Logit model. Moreover, although both the quantile-based and Logit-based predictors explain more than 50 percent of the variation in industry-level loss frequency, the Vuong-test results imply that the quantile-based predictor explains the underlying data-generating process better.

6.3.2 Reliability of Predictions of within-industry-year Moments

We use the law of total moments to develop year t predictions of the within-industry-year standard deviation, skewness, and kurtosis of ROE in year t+h. Per the discussion in Section 6.1, we predict standardized skewness and excess, standardized kurtosis; however, for ease of discussion we refer to these two variables as simply skewness and kurtosis.

In this section we describe how we use the law of total variance to predict the standard deviation of lead ROE for a particular industry-year. The manner in which we predict skewness and kurtosis is similar; however, the underlying formulas are more complicated. Hence, we relegate the technical details related to how we arrive at these predictions to Appendix A.

[13] $LOSS_{IND,t+1}$ is the industry average of $LOSS_{i,t+1}$. When calculating $LOSS_{IND,t+1}$ we exclude firms for which we are unable to develop a prediction in year t of the probability of a loss in year t+1. For example, a firm that came into existence in year t+1 will not have observable attributes in year t; hence, we cannot estimate the probability of a loss for that firm. Consequently, we do not consider this firm when we compute $LOSS_{IND,t+1}$.

The law of total variance implies that the within-group standard deviation of lead ROE can be expressed in the following manner.

$$\sqrt{VAR_t\left(ROE_{i,t+h}\right)} \equiv \sqrt{VAR_t\left(E_t\left(ROE_{i,t+h} \mid \cdot\right)\right) + E_t\left(VAR_t\left(ROE_{i,t+h} \mid \cdot\right)\right)} \quad (3)$$

In the above equation, $VAR_t(\cdot)$ denotes the variance estimated in year t and $E_t(\cdot)$ is the year t expected value.

With equation (3) in mind, we do the following. First, for each firm i, year t, and forecast horizon h we obtain the estimates of the mean and variance of $ROE_{i,t+h}$, respectively. We obtain estimates from our quantile-based approach as well as the historical firm-level and historical matched-sample approaches. Next, for each industry-year and each estimation approach, we predict the within-industry standard deviation of $ROE_{i,t,t+h}$. We do this by taking the square root of the sum of: (1) the within-industry variance of the estimate of the mean of $ROE_{i,t+h}$ and (2) the industry mean of the estimate of the variance of $ROE_{i,t+h}$. Finally, we calculate the realized cross-sectional standard deviation of ROE in year t+h. We refer to the quantile-based year t prediction of the standard deviation of industry IND's ROE in year t+h as $Q\_STD_{IND,t,t+h}$; and, we refer to the realized standard deviation as $R\_SD_{IND,t+h}$. We use similar notation for predicted and realized values of skewness and kurtosis. [14]

In Table Six we show the results of regressing realized moments on predicted moments. Panels A, B, and C relate to regressions involving realized standard deviation, skewness, and kurtosis, respectively. Columns (1) through (3) show comparisons of our quantile-based predictions to predictions obtained from the historical firm-level approach; and, columns (4) through (6) show comparisons of our quantile-based predictions to predictions obtained from the

[14] When calculating realized moments for industry IND in year t+1 we exclude firms for which we are unable to develop a prediction in year t. For example, a firm that came into existence in year t+1 will not have observable attributes in year t; hence, we do not have estimates of the year t+1 moments for that firm. Consequently, we do not consider this firm when we compute the realized moments in year t+1.

historical matched-sample approach. As discussed above, these predictions are obtained via the law of total moments. We also evaluate an approach that does not rely on the law of total moments. We refer to it as the historical industry-level approach because it involves setting the year t prediction of a particular moment in year t+h equal to the historical industry-level moment for year t.[15] Comparisons of this approach to our quantile-based approach are shown in columns (7) through (9).

As shown in Panel A the $Q\_STD_{IND,t,t+1}$ is a reliable predictor on both an absolute and relative basis. As shown in columns (1), (4) and (7), it has a significantly positive association with $R\_SD_{IND,t+1}$ and it explains roughly 30 percent of the cross-sectional variation in $R\_SD_{IND,t+1}$. Per columns (3), (6) and (9), these associations remain after controlling for the predictions obtained from the historical firm-level, historical matched-sample, and historical industry-level approaches. Moreover, Vuong-test results show that the $Q\_STD_{IND,t,t+1}$ is a better predictor than each of the other estimates.

Results related to $Q\_SKEW_{IND,t,t+1}$ are similar to those for $Q\_STD_{IND,t,t+1}$. As shown in columns (1), (4) and (7) of Panel B, $Q\_SKEW_{IND,t,t+1}$ has a significant positive association with $R\_SKEW_{IND,t+1}$ and it explains approximately 14 percent of the cross-sectional variation in $R\_SKEW_{IND,t+1}$. Per columns (3), (6) and (9), these results remain after controlling for the predictions obtained from the historical firm-level, historical matched-sample, and historical industry-level approaches. Moreover, Vuong-test results show that $Q\_SKEW_{IND,t,t+1}$ is a better predictor than each of the other estimates.

Finally, as shown in columns (1), (4) and (7) of Panel C, $Q\_KURT_{IND,t,t+1}$ is positively associated with $R\_KURT_{IND,t,t+1}$ and it explains more than 14 percent of the variation in

[15] We also considered setting the year t prediction of a particular moment in year t+h equal to the historical industry-level moment for years t-4 through t as well as years t-9 through t. Untabulated results using these alternative proxies, which are available upon request, are similar to the results shown in Table Six.

$R\_KURT_{IND,t,t+1}$. Moreover, as shown in columns (3), (6) and (9), these results remain after controlling for the predictions obtained from the historical firm-level, historical matched-sample, and historical industry-level approaches. However, Vuong-test comparing $Q\_KURT_{IND,t,t+1}$ to the remaining estimates are mixed. $Q\_KURT_{IND,t,t+1}$ is neither more nor less reliable than the other three estimates.

6.4 Summary

The results described in sections 6.2 and 6.3 lead us to conclude that our methodology generates reliable *ex ante* estimates of the probability of future loss and the higher moments of future ROE. This fact is important for two reasons. First, given they relate to infrequent events, higher moments such as skewness and kurtosis are difficult to predict. Second, as discussed in section two, the higher moments of future ROE are potentially relevant in a number of different economic contexts. Although it is outside the scope of our study to provide in-depth evidence about these different contexts, we do provide some initial evidence in the next section.

7. Analyses of Valuation Multiples and Credit Ratings

In this section we evaluate the role that the higher moments of future ROE play in determining valuation multiples and credit ratings.

Table Seven contains results obtained from regressions of year t valuation multiples on $Q\_MEAN_{i,t,t+1}$, $Q\_STD_{i,t,t+1}$, $Q\_SKEW_{i,t,t+1}$ and $Q\_KURT_{i,t,t+1}$. The results in the first (last) two columns relate to regressions in which the earnings-to-price (book-to-price) ratio is the dependent variable. Earnings-to-price, $EP_{i,t}$, (book-to-price, $BP_{i,t}$) equals the ratio of firm i's year t earnings (equity book value) to its equity market value at the end year t. Columns (1) and (3) relate to results in which the out-of-sample estimates of the moments of lead ROE are the only independent variables. Columns (2) and (4) relate to results in which we include four control variables: (1) annual stock return for year t, $RET_{i,t}$; (2) the volatility of market-model residuals, $RET\_VOL_{i,t}$; (3) the skewness of market-model residuals, $RET\_SKEW_{i,t}$; and, the kurtosis of market-model residuals, $RET\_KURT_{i,t}$.[16] We remove observations for which the value of any of the variables in the regression falls in either the top or bottom percentile of its annual distribution.[17]

The results in Table Seven show that each of the out-of-sample estimates of the moments of future ROE is related to the earnings-to-price and book-to-price ratio. For example, firms with higher *ex ante* standard deviation and skewness of ROE have higher prices. These results are consistent with the analytical models developed by Johnson [2004], Brunnermeier et al. [2007], Mitton and Vorkink [2007], and Barberis and Huang [2008] as well as empirical results shown in Ang et al. [2006, 2009], Boyer et al. [2010], and Conrad et al. [2013]. On the other hand, equity prices are decreasing in *ex ante* firm-level kurtosis. This result is consistent with results shown in Dittmar [2004]; however, he considers co-kurtosis rather than raw kurtosis.

In Table Eight we show results obtained from regressions of year t long-term credit ratings on $Q\_MEAN_{i,t,t+1}$, $Q\_STD_{i,t,t+1}$, $Q\_SKEW_{i,t,t+1}$ and $Q\_KURT_{i,t,t+1}$. Credit ratings are obtained from Standard and Poors. They range between 2 and 23. Higher ratings reflect worse credit quality; for example, a rating of 2 (23) implies a letter rating of AAA (D). In column (1) we show results in which we include the out-of-sample estimates of the moments of ROE and five control variables: (1) the natural log of the ratio of firm i's year t equity market to the sum of

[16] $RET_{i,t}$ is measured over the 12-month period beginning on the fourth month of fiscal-year t-1 and ending on the third month of fiscal-year t. Market-model residuals are obtained from regressions of firm-level monthly returns on the contemporaneous return on the market portfolio. We use monthly returns drawn from a 12-month period ending three months after the last month of fiscal-year t. We then use the residuals to calculate $RET\_VOL_{i,t}$, $RET\_SKEW_{i,t}$, and $RET\_KURT_{i,t}$. This approach is similar to the approach described in Beaver et al. [2012] and Correia et al. [2012].

[17] We also estimate rank regressions using the entire sample. Results of these regressions, which are available upon request, are similar to the results shown in Table Seven.

all firm's contemporaneous equity market values, $LN\_SIZE_{i,t}$; (2) the ratio of firm i's year t liabilities to its year t assets, $LIAB\_ASST_{i,t}$; (3) the ratio of firm i's year t earnings before interest, taxes, depreciation, and amortization to its year t liabilities, $EBITDA\_LIAB_{i,t}$; (4) $RET_{i,t}$; and, (5) $RET\_VOL_{i,t}$. The control variables are based on the default prediction model described in Beaver et al. [2012].[18] In column (2) we show results in which we add the control variables $RET\_SKEW_{,t}$ and $RET\_KURT_{i,t}$.

The results in Table Eight show that each of the out-of-sample estimates of the moments of future ROE is related to contemporaneous credit ratings; and, the relations are intuitive. For instance, firms with higher *ex ante* standard deviation of ROE have worse credit ratings, which is consistent with analytical results in Merton [1974]. However, consistent with practitioner articles (e.g, Dynkin et al. [2007]) firm-level skewness and kurtosis are also relevant. In particular, bonds are considered riskier when the *ex ante* skewness (kurtosis) of ROE is negative (positive). This is not surprising given that bondholders face relatively high exposure to downside risk while benefitting little from positive shocks.

8. Conclusion

We develop an empirical approach that yields reliable out-of-sample, firm-level estimates of the higher moments of future ROE. This is a nontrivial contribution for two reasons. First, higher moments such as skewness and kurtosis are difficult to predict; hence, the fact that our out-of-sample estimates are reliable is significant in and of itself. Second, and perhaps more importantly, the higher moments of future ROE are potentially relevant in a number of economic

[18] Our results are not sensitive to the choice of controls. In untabulated results, which are available upon request, we consider a number of alternative control variables inspired by extant studies such as Kaplan and Urwitz [1979], Chava and Jarrow [2004], and Hann et al. [2007].

contexts; hence, by developing a reliable approach for estimating these moments, we lay the necessary groundwork for evaluating their relevance.

Our study suggests three paths for future research. First, the approach we develop can be used in other contexts such as the evaluation and prediction of the higher moments of return on invested capital, earnings growth, accruals, etc. Second, as discussed on pp. 12 and 13, our model is superior to a number of more elaborate models; nonetheless, the firm-level attributes we evaluate may not be exhaustive. Hence, future studies might focus on identifying other attributes that are related to the moments of lead ROE both in- and out-of-sample. Finally, we only provide initial evidence about the role that the higher moments of future ROE play in equity and credit valuation. Hence, further testing of these issues and others is clearly warranted.

Appendix A – Technical Details

This appendix summarizes the theory and the implementation of our estimates of conditional quantiles and conditional moments. First, we present the properties of the estimate of the q^{th} conditional quantile of the random variable $y_{i,t}$. Second, we present our estimator of the quantile function, which is defined over $q \in (0,1)$. Third, we present an estimator of the conditional moment of $y_{i,t}$ using the estimator of the quantile function. Finally, we describe how we use the law of total moments to predict within-industry-year skewness and kurtosis.

A.1 Prediction of the q^{th} *Conditional Quantile*

Observable covariates for observation i in year t-h are $x_{i,t\text{-}h,1}, \ldots, x_{i,t\text{-}h,k}$ and $x_{i,t\text{-}h,1} = 1\ \forall$ i. $F(y_{i,t})$ is the conditional distribution function of y for observation i in year t, and $QUANT_q(y_{i,t}|x_{i,t\text{-}h,1,\ldots,}x_{i,t\text{-}h,k}) = F_{i,t}^{-1}(q)$ is the q^{th} conditional quantile of $y_{i,t}$ for observation i in year t.

We are interested in estimating the conditional quantile function $QUANT_q(y_{i,t}|x_{i,t\text{-}h,1,\ldots,}x_{i,t\text{-}h,k})$ for observation i in year t given the observed covariates $x_{i,t\text{-}h,1,\ldots,}\ x_{i,t\text{-}h,k}$. We stack the observations of the k covariates into a single 1×k vector. When referring to this vector we use bold text (i.e., $\mathbf{x_{i,t\text{-}h}}$.is the value of the vector for observation i in year t-h). We also use bold text when referring to the k×1 vector of coefficients.

As is standard in the quantile regression literature, we assume that there is a linear relationship between each conditional quantile and the observable covariates.

Assumption 1: $QUANT_q(y_{i,t}|\mathbf{x_{i,t\text{-}h}}) = \mathbf{x_{i,t\text{-}h}}`\boldsymbol{\beta^q}$.

The estimator $\mathbf{b^q}$ of $\boldsymbol{\beta^q}$ minimizes the objective function:

$$\mathbf{b}^q = \text{argmin}_\beta \sum_{i=1}^{N} \rho_q(y_i - x_{i,t-h}'\beta^q) \qquad \text{(A.1)}$$

In equation (A.1), $\rho_q(u)$ is the check function and is defined as $\rho_q(u) = u(q-I(u<0))$ ($I(\cdot)$ is the indicator function).

Under the regularity conditions specified in Koenker and Bassett [1978], the estimator $\mathbf{b}^q$ is a $\sqrt{N}$ consistent and asymptotically normal estimator of $\boldsymbol{\beta}^q$ such that $\sqrt{N}(\mathbf{b}^q - \boldsymbol{\beta}^q) \longrightarrow N(0,V)$. The variance covariance matrix V is equal to $\left(\frac{q(1-q)}{[f_{u(q)}(0)]^2}\right)\left[E(x_{i,t-h}{}'x_{i,t-h})\right]^{-1}$, $u(q) = y_{i,t} - \beta^q$, and $f_{u(q)}(0)$ is the density of $u(q)$ at $u(q) = 0$. The estimate of the q^{th} quantile of $y_{i,t}$—i.e., $\hat{QUANT}_q(y_{i,t} \mid \mathbf{x_{i,t-h}}) = \mathbf{x_{i,t-h}}`\mathbf{b^q}$—is also a consistent and asymptotically normal estimator of $QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})$.

A.2 Prediction of the Quantile Function of $y_{i,t}$

The coefficients $\boldsymbol{\beta}^q$ and the estimates $\mathbf{x_{i,t-h}}`\boldsymbol{\beta}^q$ are computed for quantiles $0 < q_1 < q_2 < \ldots < q_Q < 0$. In practice, we form estimates for the largest value of Q such that the numerical estimates converge. We use these Q estimates of the coefficients to build an estimate of the quantile function $QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})$ for all values q in (0,1). This is defined as the following staircase function:

$$\begin{aligned}\hat{QUANT}_q(y_{i,t} \mid \mathbf{x_{i,t-h}}) &= \sum_{s=1}^{Q} I(q_{s-1} < q \le q_s) \times \hat{QUANT}_{q_s}(y_{i,t} \mid \mathbf{x_{i,t-h}}) \\ &+ I(q_Q < q \le 1) \times \hat{QUANT}_Q(y_{i,t} \mid \mathbf{x_{i,t-h}})\end{aligned} \tag{A.2}$$

The estimator of the quantile $QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})$ is not in general monotonic in q. Following Chernozhukov, Fernandez-Val, and Galichon [2010], we define the rearranged quantile function $QUANT_q^*(y_{i,t}|\cdot)$ and the rearranged estimator of the quantile function $\hat{QUANT}_q^*(y_{i,t} \mid \cdot)$.

Interestingly, as the number of quantiles Q and the number of observations goes to infinity, the estimate $\hat{QUANT}_q^*(y_{i,t} \mid \cdot)$ of the quantile function converges point by point and uniformly to the quantile function $QUANT_q^*(y_{i,t}|\cdot)$ for any compact subset of (0,1).

In particular, consider a sequence $0 < q_1^n < q_2^n < ... < q_{Q^n}^n < 1$ that satisfies $\max_s | q_s^n - q_{s+1}^n | \to 0$ as $n \to \infty$—i.e., the 'step' between quantile estimates goes to 0 as the number of quantile estimates goes to infinity. We assume that the lowest (highest) quantile goes to zero (one) as the sample size goes to infinity. For each quantile q_s^n the estimator $\mathbf{b}^{q_s}$ is defined as in (A.1).

Proposition 1. For any compact subset $\Theta \subset (0,1)$, the rearranged estimator $\hat{QUANT}_q^*(y_{i,t} \mid \cdot)$ of the quantile function converges uniformly in probability to the rearranged quantile function $QUANT_q^*(y_{i,t}|\cdot)$ on Θ.

Proof. This follows from Corollary 3 of Proposition 5 of Chernozhukov, Fernandez-Val, and Galichon (2010).□

A.3 Prediction of the Conditional Moments of $y_{i,t}$

Expected variance, expected skewness, and expected kurtosis are entirely determined by the expected second, third, and fourth moments of $y_{i,t}$. The j^{th} raw moment of $y_{i,t}$ is the expected mean of the j^{th} power of $y_{i,t}$.

$$E\left(y_{i,t}^j\right) = \int_{-\infty}^{+\infty} y^j f_{i,t-h}(y) dy \qquad \text{(A.3)}$$

The unknown term here is the distribution function $f_{i,t}(y)$. Although equation (A.3) is the standard textbook formula, an equivalent formula involves the quantile function $QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})$. In particular, note that the q^{th} quantile is the cumulative distribution function of $y_{i,t}$ at q—i.e., $QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}}) = F^{-1}(q)$. Consequently, the change in variables implies $F^{-1}(q) = y$, which, in turn, implies $QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})^j = y^j$ and $dq = f_{i,t+h}(y)dy$. Finally the j^{th} moment is:

$$E\left(y_{i,t}^j\right) = \int_0^1 QUANT_q\left(y_{i,t}|\mathbf{x_{i,t-h}}\right)^j dq \qquad \text{(A.4)}$$

Using the analogy principle of Manski [1988], an estimator of the j^{th} moment of $y_{i,t}$ is provided by equation (A.4), substituting the estimator of the quantile function for the quantile function:

$$\hat{E}(y_{i,t}^j) = \int_{q_1^n}^{q_{Q^n}^n} \hat{QUANT}_q(y_{i,t}|\mathbf{x_{i,t-h}})^j dq \qquad \text{(A.5)}$$

The bounds of the integral are set to q_1^n and $q_{Q^n}^n$ instead of 0 and 1.

Proposition 2. The estimator $\hat{E}(y_{i,t}^j)$ is a consistent estimator of $E(y_{i,t}^j)$ as the sample size (i.e., N) goes to infinity and the number of steps (i.e., n) goes to infinity.

Proof. To prove consistency, choose two arbitrary numbers $\varepsilon > 0$ and $\nu > 0$.

Take N, a number such that, if $n > N$, then:

$$\left| \int_{q_1^n}^{q_{Q^n}^n} QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})^j dq - \int_0^1 QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})^j dq \right| < \varphi$$

Using Proposition 1, the estimate of the quantile function converges uniformly in probability to the quantile function. Hence, there exists N^* (depending on n) such that, if the number of observations is greater than N^*, then:

$$P\left(\left|\int_{q_1^n}^{q_{Q^n}^n} \hat{QUANT}_q(y_{i,t}|\mathbf{x_{i,t-h}})^j dq - \int_{q_1^n}^{q_{Q^n}^n} QUANT_q(y_{i,t}|\mathbf{x_{i,t-h}})^j dq\right| > \varepsilon\right) < \nu$$

The consistency of the estimate of the moment follows from these two inequalities□.

A.4 Estimating the Firm-year Conditional Mean, Standard Deviation, Skewness, and Kurtosis

The conditional moments of $y_{i,t+h}$ are estimated in the following manner. First, one quantile regression at each quantile q_s provides the estimator $\mathbf{b}^{q_s}$. We set q_s equal to s÷(1+Q) for s = 1, 2, …, Q and we use the Stata software package *qreg*. Second, we store all coefficient vectors $\mathbf{b}^{q_s}$ in a single data set that contains one observation for each quantile and one column for each covariate; hence, the dataset contains Q rows and k columns: Third, given the covariates $\mathbf{x_{i,t}}$ the data set is augmented by an additional variable that we refer to as PREDICTION. PREDICTION equals the cross product of the covariates with the coefficients for the q^{th} quantile; and, it represents the estimate of the q^{th} quantile conditional on $\mathbf{x_{i,t}}$. Finally, the estimate the j^{th} moment of $y_{i,t+h}$ is the j^{th} moment of PREDITION. The estimated conditional variance, skewness, and kurtosis are obtained by typing the Stata commands: *summarize PREDICTION detail*.

A.5 Predicting Within-industry-year Skewness and Kurtosis

As shown in Klugman, Panjer, and Willmot [1998], the law of total moments implies that the within-group third central moment of lead ROE can be expressed in the following manner.

$$M_t^3\left(ROE_{i,t+h}\right) = E_t\left(M_t^3\left(ROE_{i,t+h} \mid \cdot\right)\right) + 3 \times COV_t\left(E_t\left(ROE_{i,t+h} \mid \cdot\right), VAR_t\left(ROE_{i,t+h} \mid \cdot\right)\right) + M_t^3\left(E_t\left(ROE_{i,t+h} \mid \cdot\right)\right) \quad \text{(A.6)}$$

In the above equation, $M_t^3(\cdot)$ denotes the third central moment estimated in year t and $COV_t(\cdot,\cdot)$ is the covariance estimated in year t.

With equation (A.6) in mind, we do the following. First, for each firm i, year t, forecast horizon h, and estimation approach—i.e, quantile, historical firm-level, and historical matched-sample—we form estimates of the mean, variance, and third central moment of $ROE_{i,t+h}$. Second, for a particular industry and year t we identify all firms with non-missing values of the aforementioned estimates. Third, we predict the within-industry third moment of $ROE_{i,t,t+h}$ by summing: (1) the within-industry mean of the firm-level estimates of the third central moment; (2) three times the within-industry covariance of the firm-level estimates of the variance and the mean; and, (3) the within-industry third moment of the firm-level estimates of the mean. Finally, we compute predicted, industry-level skewness by dividing the aforementioned sum by the third power of the predicted industry-level standard deviation.

We base our prediction of the fourth central moment on the equation shown below (Klugman, Panjer, and Willmot [1998]).

$$M_t^4\left(ROE_{i,t+h}\right) = E_t\left(M_t^4\left(ROE_{i,t+h} \mid \cdot\right)\right) + 4 \times COV_t\left(M_t^3\left(ROE_{i,t+h} \mid \cdot\right), E_t\left(ROE_{i,t+h} \mid \cdot\right)\right) + 6 \times COSKEW_t\left(VAR_t\left(ROE_{i,t+h} \mid \cdot\right), E_t\left(ROE_{i,t+h} \mid \cdot\right), E_t\left(ROE_{i,t+h} \mid \cdot\right)\right) + 6 \times E_t\left(VAR_t\left(ROE_{i,t+h} \mid \cdot\right)\right) \times VAR_t\left(E_t\left(ROE_{i,t+h} \mid \cdot\right)\right) + M_t^4\left(E_t\left(ROE_{i,t+h} \mid \cdot\right)\right) \quad \text{(A.7)}$$

In the above equation, $M_t^4(\cdot)$ denotes the fourth central moment estimated in year t and $COSKEW_t(\cdot,\cdot,\cdot)$ is the non-standardized coskewness estimated in year t.

With equation (A.7) in mind, we do the following. First, for each firm i, year t, forecast horizon h, and estimation approach—i.e, quantile, historical firm-level, and historical matched-sample—we form estimates of the mean, variance, third central moment, and fourth central moment of $ROE_{i,t+h}$. Second, for a particular industry and year t we identify all firms with non-missing values of the aforementioned estimates. Third, we predict the within-industry fourth moment of $ROE_{i,t,t+h}$ by summing: (1) the within-industry mean of the firm-level estimates of the fourth central moment; (2) four times the within-industry covariance of the firm-level estimates of the third central moment and the mean; (3) six times the within-industry coskewness of the firm-level estimates of the variance, mean, and mean; (4) six times the product of the within-industry mean of the firm-level estimates of the variance and the mean; and, (5) the within-industry fourth central moment of the firm-level estimates of the mean. Finally, we compute the predicted, conditional skewness by dividing the aforementioned sum by the fourth power of the predicted industry-level standard deviation.

Appendix B – Sample Construction and Related Issues

In this appendix we provide a detailed description of how we construct the various samples underlying our in-sample estimates, out-of-sample predictions, tests of reliability, and analyses of valuation multiples and credit ratings.

B.1 Construction of Estimation and Prediction Samples

For each forecast horizon $h \in [1,5]$, we form two samples: (1) the estimation sample and (2) the prediction sample. The estimation sample contains observations that are used to estimate the coefficients shown in equation (2). The prediction sample contains observations for which we develop out-of-sample, firm-level predictions of the cdf of lead ROE.

We obtain our data from the Compustat North America Annual file. We use the Compustat variable IB, Income Before Extraordinary Items, as our measure of earnings. The Compustat variable CEQ, Common/Ordinary Equity - Total, is our measure of equity book value. We use the balance sheet approach described in Sloan [1996] to estimate accruals. Total assets equals Compustat variable AT, Assets - Total; and, dividends equals Compustat variable DVPSX_F, Dividends per Share - Ex-Date - Fiscal.

To form the estimation sample we identify all observations that have non-missing values of the variables shown in equation (2) and positive values of equity book value in year t-h. Next, we delete extreme observations, which we define as observations for which: $|ROE_{i,t}| > 2$, $|ROE_{i,t-h}| > 2$, $|ACC_{i,t-h}| > 2$, $LEV_{i,t-h} \notin [1,20]$, and $PAYOUT_{i,t-h} \notin [0,1]$. When h equals 1 (i.e., a one-year forecast horizon), the estimation sample contains 174,215 firm-years with independent (dependent) variables drawn from the time-period spanning 1963 to 2010 (1964 to 2011). The sample size decreases as h increases; for example, when h equals 5, the estimation sample contains 122,935 firm-years with independent (dependent) variables drawn from the time-period spanning 1963 to 2006 (1968 to 2011).

To form our prediction sample we identify all firm-years with positive equity book value in year t and non-missing values of $ROE_{i,t} = IB_{i,t} \div CEQ_{i,t}$, $LOSS_{i,t}$, $ROE_{i,t} \times LOSS_{i,t}$, $ACC_{i,t}$, $LEV_{i,t}$, $PAYER_{i,t}$, and $PAYOUT_{i,t}$. We do not remove extreme observations nor do we remove observations with missing values of lead ROE. We limit our prediction sample to firm-years drawn from 1973 to 2011. The prediction sample contains 170,522 firm-years. However, because some of our tests involve comparing *ex ante* predictions to *ex post* realizations, the number of observations underlying the results shown in Tables Three through Eight is lower. In particular, as shown in Table Three (Five) there are 156,973 (2,056) firm-years (industry-years) for which we have both an out-of-sample estimate of $Q\_PROB_{i,t,t+1}$ ($Q\_PROB_{IND,t,t+1}$) and a non-missing value of $LOSS_{i,t+1}$ ($LOSS_{IND,t+1}$). Finally, as discussed below, sample sizes underlying our tests also vary because some of our tests involve: (1) comparing our quantile-based estimates to alternative estimates that cannot be calculated for all the observations in the prediction sample or (2) market multiples or credit ratings that are missing for some of the observations in the prediction sample.

B.2 Alternative Out-of-sample Estimates of Higher Moments

We compare our quantile-based estimates of the standard deviation of lead ROE to three alternatives: (1) historical firm-level estimates; (2) historical matched-sample estimates; and (3) historical industry-level estimates.

To calculate the historical firm-level estimate for firm-year i,t we use firm i's ROE in years t-9 through t. If firm i,t has missing ROE for any year between year t-9 and t, we set the estimated moment equal to the median of the historical firm-level estimates for the remaining firms in the industry. If there are no firms from the industry with non-missing ROE in all the years between years t-9 and t, we code the observation as missing. As shown in Table Four, 134,552 of the firm-years in the prediction sample have a non-missing value of $|FERR_{i,t+1}|$ and a non-missing historical firm-level estimate. As shown in Table Six, there are 1,960 industry-years for which: (1) the realized industry-level moment in year t+1 is non-missing and (2) there are ten observations in the prediction sample with non-missing historical firm-level estimates that can be used to construct industry-level estimates via the law of total moments.

To calculate the historical matched-sample estimate for firm-year i,t we use the algorithm described in Larson and Resutek [2013]. First, we select all firms with non-missing ROE and ΔROE during years t-4 through t. Next, we eliminate from this group the firms that are not in the same NYSE size decile (size refers to total assets) as firm-year i,t. Finally, within this set of firms we select those firms with similar ROE and ΔROE. We first define similar as ROE (ΔROE) that is within plus or minus 0.50 percentage points of firm i's ROE (ΔROE) in year t. However, if the number of firms meeting this criterion is less than five, we define similar as ROE (ΔROE) that is between 80 percent and 120 percent of firm i's ROE (ΔROE) in year t. As shown in Table Four, 85,142 of the firm-years in the prediction sample have a non-missing value of $|FERR_{i,t+1}|$ and a non-missing historical matched-sample estimate. As shown in Table Six, there are 1,579 industry-years for which: (1) the realized industry-level moment in year t+1 is non-missing and (2) there are ten observations in the prediction sample with non-missing historical matched-sample estimates that can be used to construct industry-level estimates via the law of total moments.

To calculate the historical industry-level estimate for industry-year IND,t we identify all observations from industry IND with non-missing ROE in year t. If there are less than ten observations meeting this criterion, we code the historical industry-level estimate as missing. If there are ten or more non-missing values, we set the historical industry-level estimate of a particular moment equal to the sample moment for this group of firms. As shown in Table Six, there are there are 2,056 industry-years for which: (1) the realized industry-level moment in year t+1 is non-missing; (2) there are ten observations in the prediction sample; and, (3) there are ten non-missing values of ROE that can be used to calculate the historical industry-year estimate.

B.2 Analyses of Market Multiples and Credit Ratings

In our regressions that use market multiples as the dependent variable we use all observations from the prediction sample with non-missing values of the variables used in the regression. However, we remove extreme observations, which we define as those for which any of the variables used in the regression is in either the top or bottom one percent of its annual distribution. After applying these criteria there are 155,025 (125,468) observations available for estimating the regression that excludes (includes) the independent variables $RET_{i,t}$, $RET\_VOL_{i,t}$, $RET\_SKEW_{i,t}$, and $RET\_KURT_{i,t}$.

In our regressions that use credit ratings as the dependent variable we use all observations from the prediction sample with non-missing values of the variables in the regression. However, we remove extreme observations, which we define as those for which any of the independent variables in the regression is in either the top or bottom one percent of its annual distribution. After applying these criteria there are 20,937 (20,148) observations available for estimating the regression that excludes (includes) the independent variables $RET\_SKEW_{i,t}$ and $RET\_KURT_{i,t}$. The sample size for these tests is relatively small because credit ratings are not available prior to 1985 and some firms do not have debt that is rated by Standard and Poors.

References

ANG, A.; R. HODRICK; Y. XING; AND X. ZHANG. "The Cross-Section of Volatility and Expected Return." *The Journal of Finance* (February 2006): 259-299.

ANG, A.; R. HODRICK; Y. XING; AND X. ZHANG. "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence." *Journal of Financial Economics* (January 2009): 1-23.

BANKER, R.D.; AND L. CHEN. "Predicting Earnings Using a Model Based on Cost Variability and Cost Stickiness." *The Accounting Review* (March 2006): 285-307.

BARBERIS, N.; M. HUANG. "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices." *The American Economic Review* (December 2008): 2066-2100.

BASU, S. "The Conservatism Principle and the Asymmetric Timeliness of Earnings." *Journal of Accounting and Economics* (December 1997): 3-37.

BEAVER, W.H.; M. CORREIA; AND M.F. MCNICHOLS. "Do Differences in Financial Reporting Attributes Impair the Predictive Ability of Financial Ratios for Bankruptcy?" *Review of Accounting Studies* (December 2012): 969-1010.

BOYER, B.H.; T. MITTON; AND K. VORKINK. "Expected Idiosyncratic Skewness." *Review of Financial Studies* (January 2010): 169-202.

BRAV, A.; J.R. GRAHAM; C.R. HARVEY; AND R. MICHAELY. "Payout Policy in the 21st Century." *Journal of Financial Economics* (September 2005): 483-527.

BRILLINGER, D. "The Calculation of Cumulants Via Conditioning." *Annals of The Institute of Statistical Mathematics* (December 1969): 215-218.

BRUNNERMEIER, M.K.; C. GOLLIER; AND J.A. PARKER. "Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns." *The American Economic Review* (May 2007): 159-165.

BUCHINSKY, M. "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research." *The Journal of Human Resources* (Winter 1998): 88-126.

CHAVA, S.; AND R. JARROW. "Bankruptcy Prediction with Industry Effects." *Review of Finance* (2004): 537-569.

CHERNOZHUKOV, V.; I. FERNANDEZ-VAL; AND A. GALICHON. "Quantile and Probability Curves without Crossing." *Econometrica* (May 2010): 1093-1125.

CONRAD, J.; R.F. DITTMAR; AND E. GHYSELS. "Ex Ante Skewness and Expected Stock Returns." *The Journal of Finance* (February 2013): 85-124.

CORREIA, M.; S. RICHARDSON; AND I. TUNA. "Value Investing in Credit Markets." *Review of Accounting Studies* (September 2012): 572-609.

DYNKIN, L.; A. GOULD; J. HYMAN; V. KONSTANTINOVSKY; AND B. PHELPS. *Quantitative Management of Bond Portfolios (Advances in Financial Engineering)*. Princeton University Press, 2007.

ESPLIN, A.; M. HEWITT; M. PLUMLEE; AND T. YOHN. "Disaggregating Operating and Financing Activities: Implications for Forecasts of Future Profitability." Working Paper, Indiana University and University of Utah, 2011.

FAIRFIELD, P.M.; R. RAMNATH; AND T.L. YOHN. "Do Industry-level Analyses Improve Forecasts of Financial Performance?" *Journal of Accounting Research* (March 2009): 147-178.

FAIRFIELD, P.M.; AND T.L. YOHN. "Using Asset Turnover and Profit Margin to Forecast Change in Profitability" *Review of Accounting Studies* (December 2001): 372-386.

FAMA, E. F.; AND K. FRENCH. "Forecasting Profitability and Earnings." *The Journal of Business* (April 2000): 161-175.

FAMA, E. F.; AND K. FRENCH. "Profitability, Investment and Average Returns." *Journal of Financial Economics* (December 2006): 491-518.

FREEMAN, R.; J. A. OHLSON; AND S. H. PENMAN. "Book Rate-of-Return and Prediction of Earnings Changes: An Empirical Investigation." *Journal of Accounting Research* (Autumn 1982): 639-653.

GERAKOS J.; AND R. B. GRAMACY. “Regression-based Earnings Forecasts.” Working Paper, University of Chicago, 2013.

GIVOLY, D.; AND C. HAYN. “The Changing Time-series Properties of Earnings, Cash Flows and Accruals: Has Financial Reporting Become More Conservative?” *Journal of Accounting and Economics* (June 2000): 287-320.

HANN, R. N.; F. HEFLIN; AND K. R. SUBRAMANAYAM. “Fair-Value Pension Accounting.” *Journal of Accounting and Economics* (December 2007): 328-358.

HOU, K.: AND D.ROBINSON. “Industry Concentration and Average Stock Returns: *Journal of Finance* (August 2006): 1927-1956.

HOU, K.; M. A. VAN DIJK; AND Y. ZHANG. “The Implied Cost of Capital: A New Approach.” *Journal of Accounting and Economics* (June 2012): 504-526.

JOHNSON, T.C. “Forecast Dispersion and the Cross-Section of Expected Returns.” *The Journal of Finance* (October 2004): 1957-1978,

KAPLAN, R. S.; AND G. URWITZ. “Statistical Models of Bond Ratings: A methodological Inquiry.” *The Journal of Business* (1979): 231-261.

KLEPPER, S. “Entry, Exit, Growth, and Innovation over the Product Life Cycle.” *The American Economic Review* (June 1996): 562-583.

KLUGMAN, S.; H. PANJER; AND G. WILLMOT. *Loss Models: From Data to Decisions*. New York: John Wiley & Sons, 1998.

KOENKER, R.; AND G. BASSETT JR. “Regression Quantiles.” *Econometrica* (January 1978): 33-50.

KONSTANTINIDI, T.; AND P. F. POPE. “Forecasting Risk in Future Earnings.” Working Paper. Cass Business School, 2012.

LARSON, C.; AND R.J. RESUTEK. “Sources of Investor Uncertainty and Expected Stock Returns.” Working Paper, Washington University in St. Louis and Dartmouth College, 2013.

LI, K.; AND P. MOHANRAM. “Implied Cost of Capital for Firms without Earnings forecasts: A Fundamental Approach, Working Paper, University of Toronto, 2012.

LINTNER, J. “Distribution of Incomes of Corporations among Dividends, Retained Earnings, and Taxes.” *The American Economic Review* (May 1956): 97-113.

MANSKI, C. F. *Analog Estimation Methods in Econometrics*. Chapman & Hall, 1988.

MERTON, R.C. “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates.” *The Journal of Finance* (May 1974): 449-470.

MERTON, R.C. “A Simple Model of Capital Market Equilibrium with Incomplete Information.” *The Journal of Finance* (July 1987): 483-510.

MILLER, M.; AND K. ROCK. “Dividend Policy under Asymmetric Information.” *Journal of Finance* (September 1985): 1031-1051.

MITTON, T.; AND K. VORKINK. “Equibrium Underdiverfication and the Preference for Skewness.” *Review of Financial Studies* (January 2007): 1255-1288.

MODIGLIANI, F.; AND M. MILLER. “The Cost of Capital, Corporation Finance and the Theory of Investment.” *The American Economic Review* (June 1958): 261-297.

NISSIM, D.; AND S. H. PENMAN. “Ratio Analysis and Equity Valuation: From Research to Practice.” *Review of Accounting Studies* (March 2001): 109-154.

NISSIM, D.; AND S. H. PENMAN. “Financial Statement Analysis of Leverage and How It Informs about Profitability and Price-to-Book Ratios.” *Review of Accounting Studies* (December 2003): 531-560.

PENMAN, S.H. *Financial Statement Analysis and Security Valuation*, McGraw Hill, 2012.

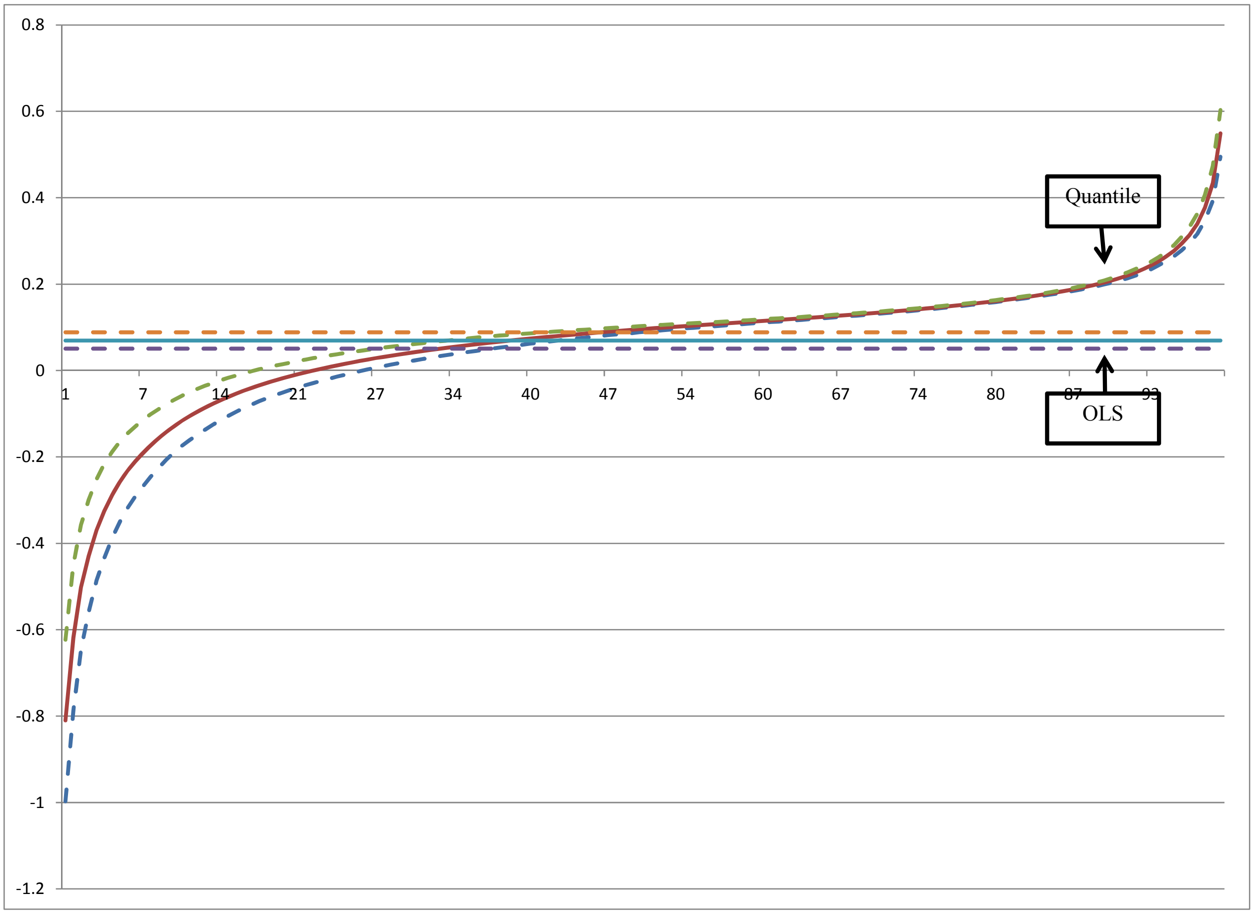
PENMAN, S. H.; AND X. ZHANG. “Accounting Conservatism, the Quality of Earnings, and Stock Returns.” *The Accounting Review* (April 2002): 237-264.

SLOAN, R. “Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?” *The Accounting Review* (July 1996): 289-315.

SOLIMAN, M. "The Use of DuPont Analysis by Market Participants." *The Accounting Review* (May 2008): 823-853.

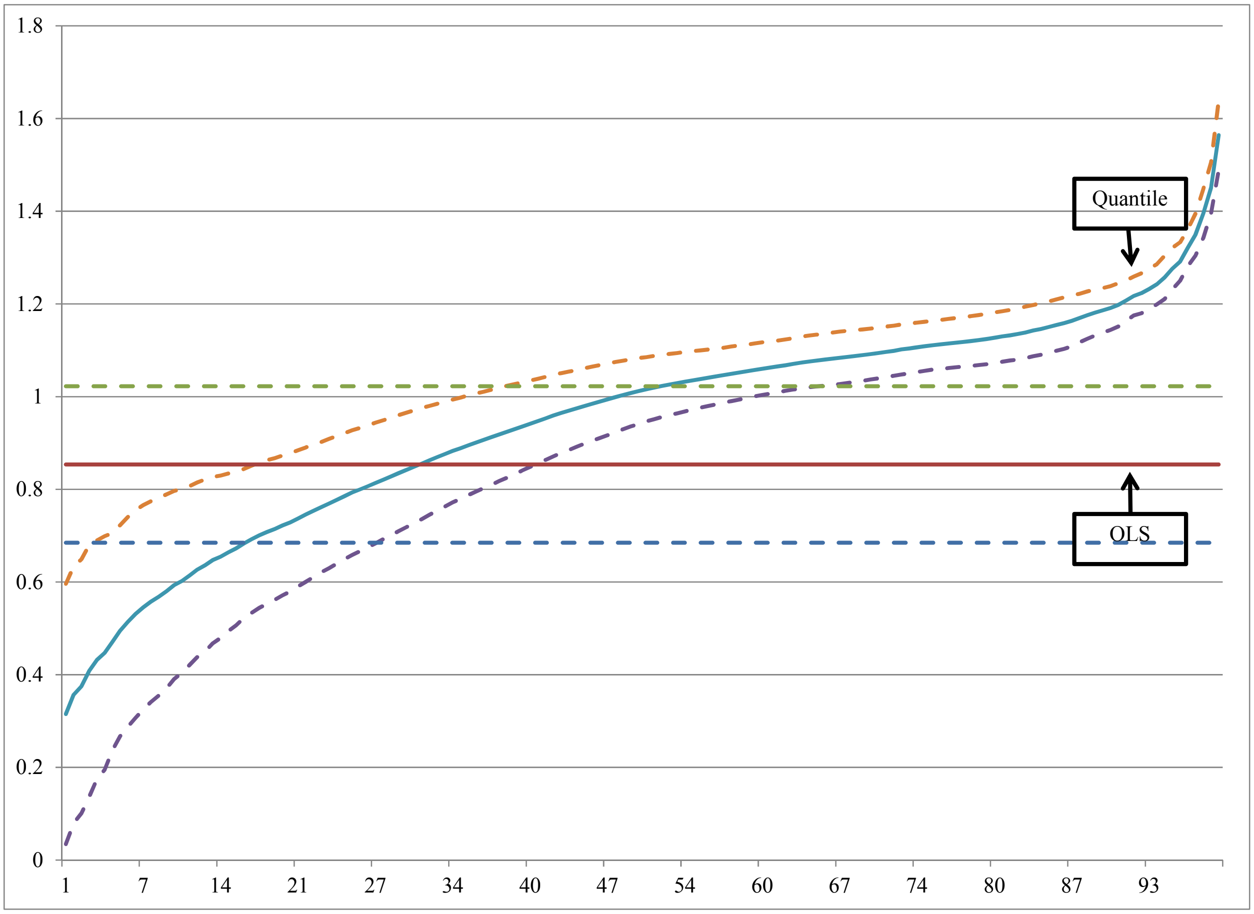
VUONG, Q. H. "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses." *Econometrica* (March 1989): 307-333.

Figure One – Graph of $\beta_{0,AVG}^{q}$ on q



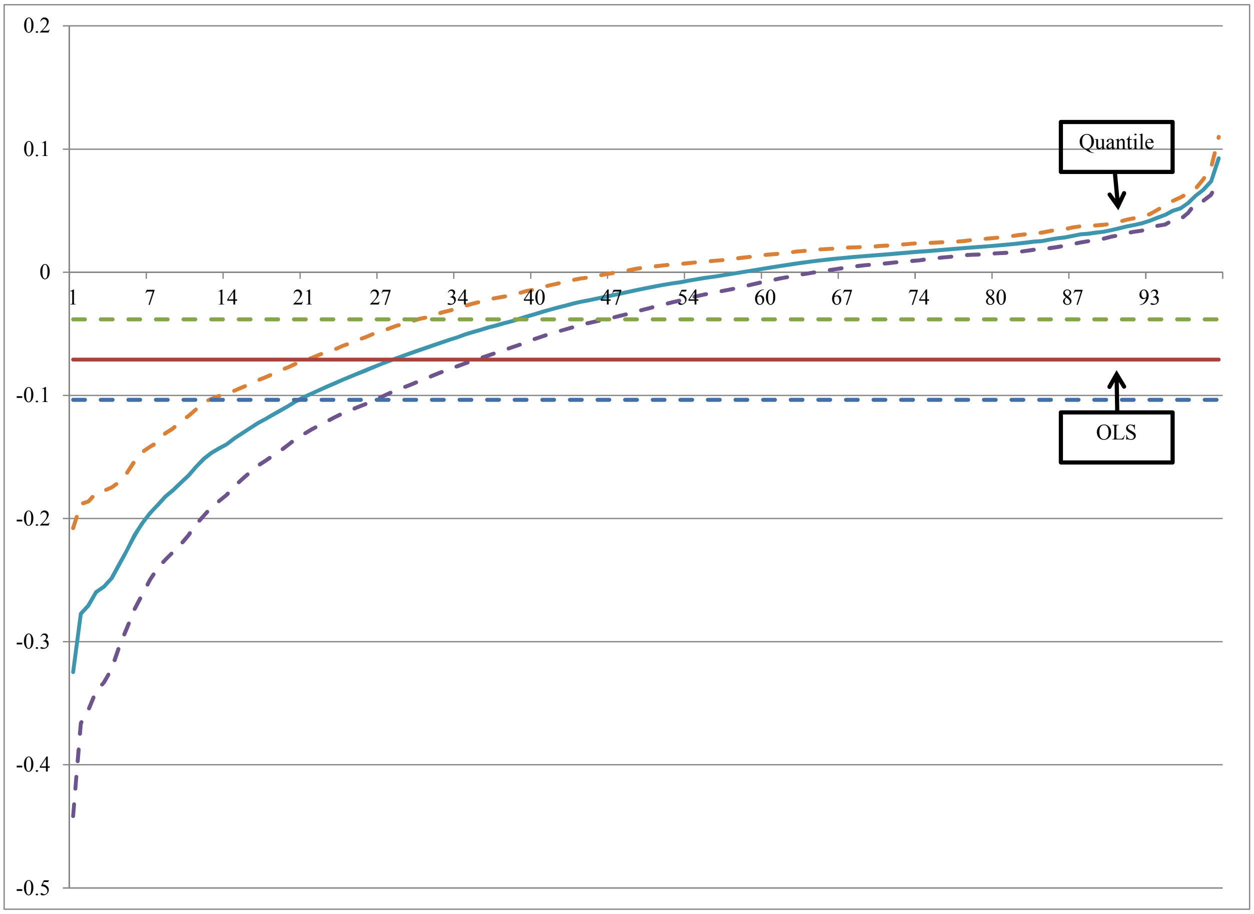
$\beta_{0,AVG}^{q}$ is the average coefficient on the constant term. Quantiles are shown on the x-axis and values of the constant are shown on the y-axis. The solid line is the average across estimation years of the estimates of the constant. Dashed lines equal the average ± 1.96 multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Two – Graph of $\beta_{1,AVG}^{q}$ on q



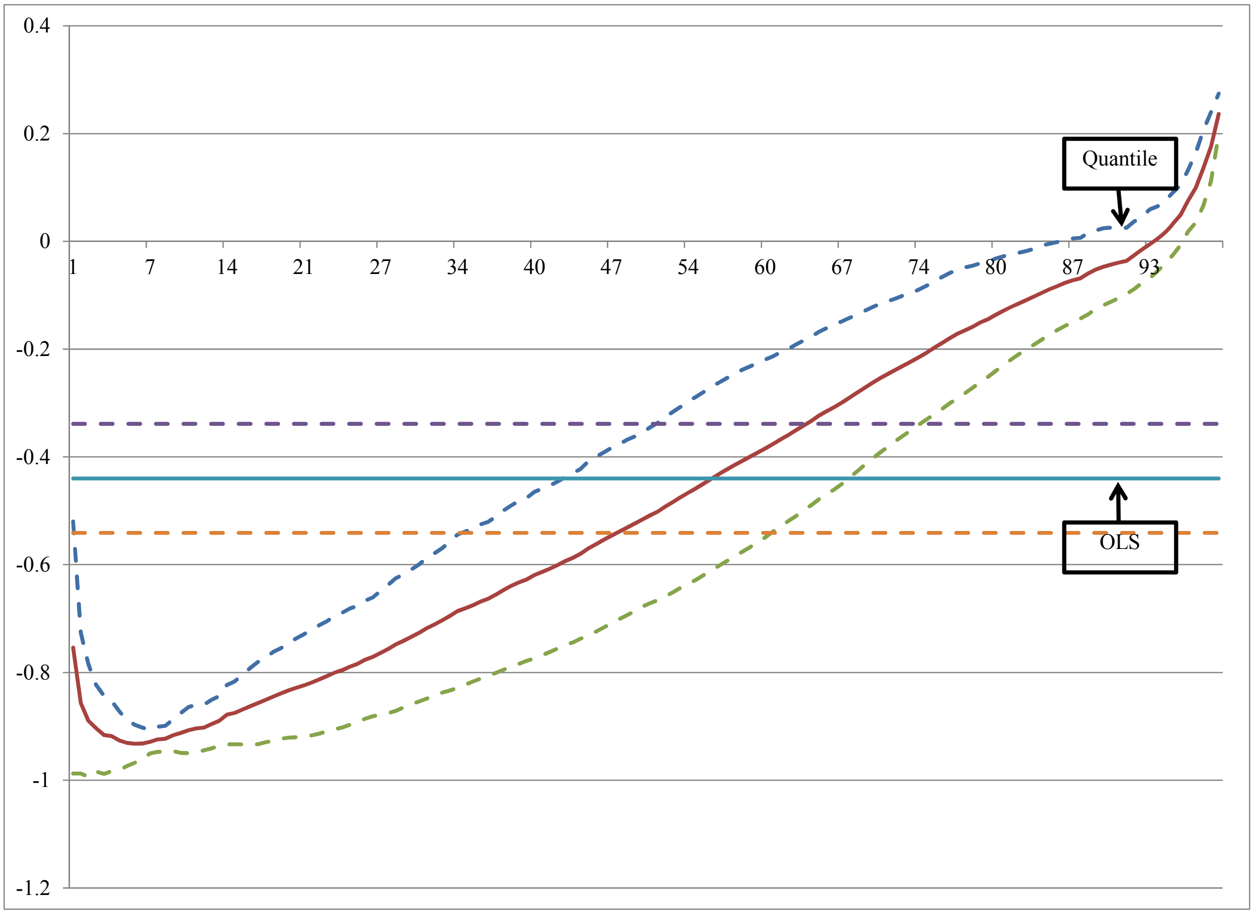
$\beta_{1,AVG}^{q}$ is the average coefficient on $ROE_{i,t-1}$. Quantiles are shown on the x-axis and values of the coefficient are shown on the y-axis. The solid line is the average across estimation years of the estimates of the coefficient. Dashed lines equal the average ± 1.96 multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Three – Graph of $\beta^{q}_{2,AVG}$ on q



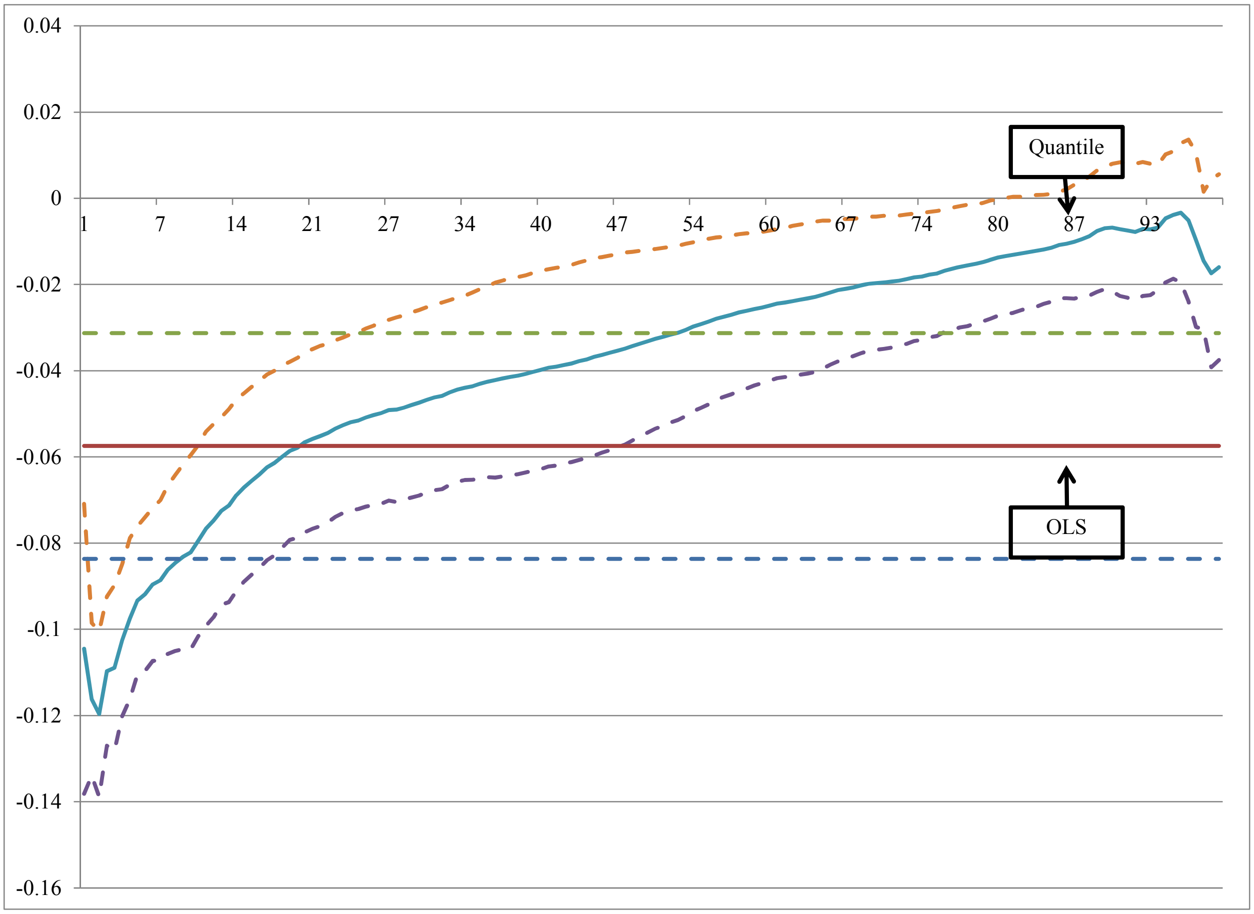
$\beta^{q}_{2,AVG}$ is the average coefficient on $LOSS_{i,t-1}$. Quantiles are shown on the x-axis and values of the coefficient are shown on the y-axis. The solid line is the average across estimation years of the estimates of the coefficient. Dashed lines equal the average ± 1.96 multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Four – Graph of $-\left(\beta_{1,AVG}^{q}+\beta_{3,AVG}^{q}\right)$ on q



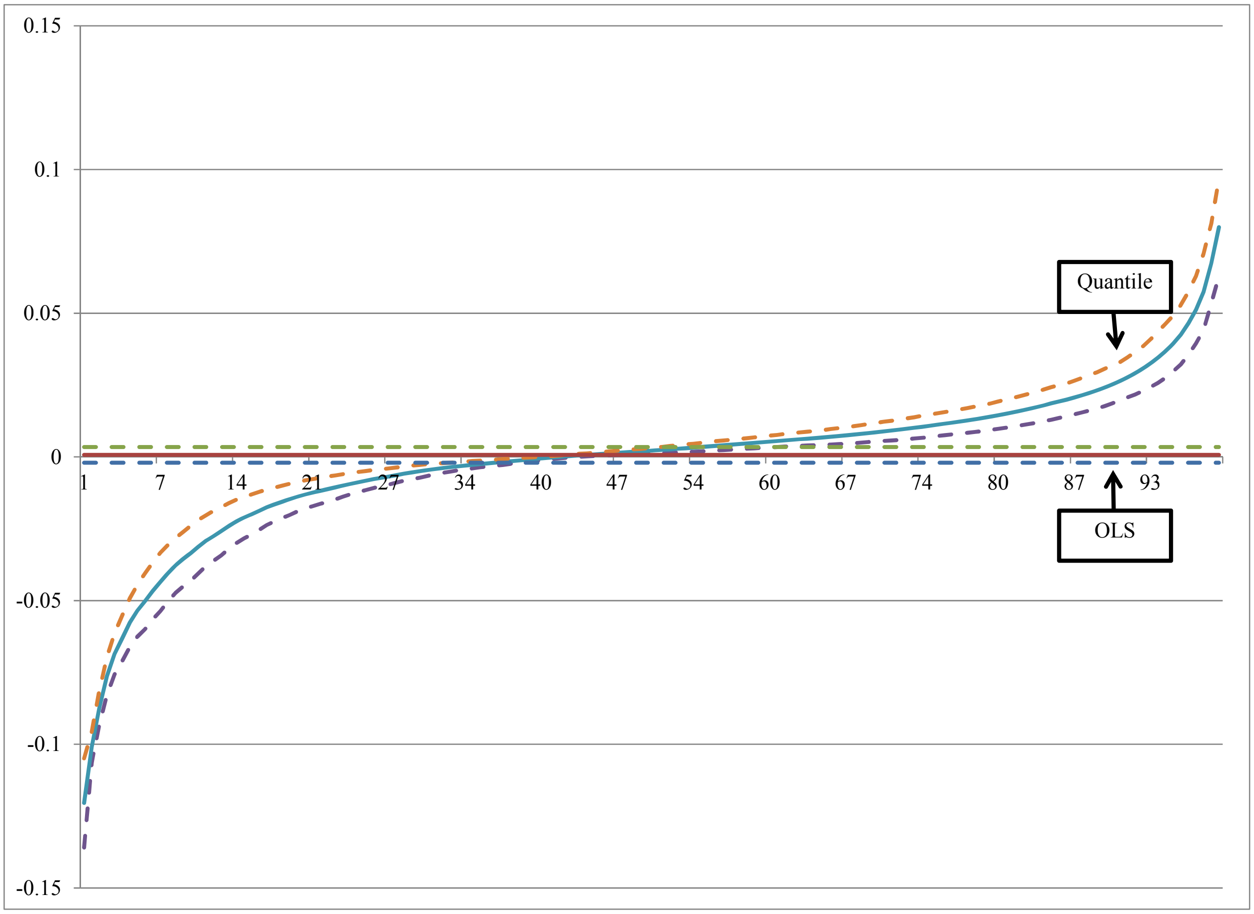
$\beta_{1,AVG}^{q}$ is the average coefficient on $ROE_{i,t-1}$ and $\beta_{3,AVG}^{q}$ is the average coefficient on $LOSS_{i,t-1}\times ROE_{i,t-1}$. Quantiles are shown on the x-axis and values of the coefficient are shown on the y-axis. The solid line is the average across estimation years of the estimates of the coefficient. Dashed lines equal the average ± 1.96 multiplied by the standard error of the temporal averages of $-1\times\left(\beta_{1,EY}^{q}+\beta_{3,EY}^{q}\right)$ and $-1\times\left(\beta_{1,EY}^{OLS}+\beta_{3,EY}^{OLS}\right)$—i.e., we use the standard error of the average of the sum *not* the sum of the standard errors of the averages. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Five - Graph of $\beta^{q}_{4,AVG}$ on q



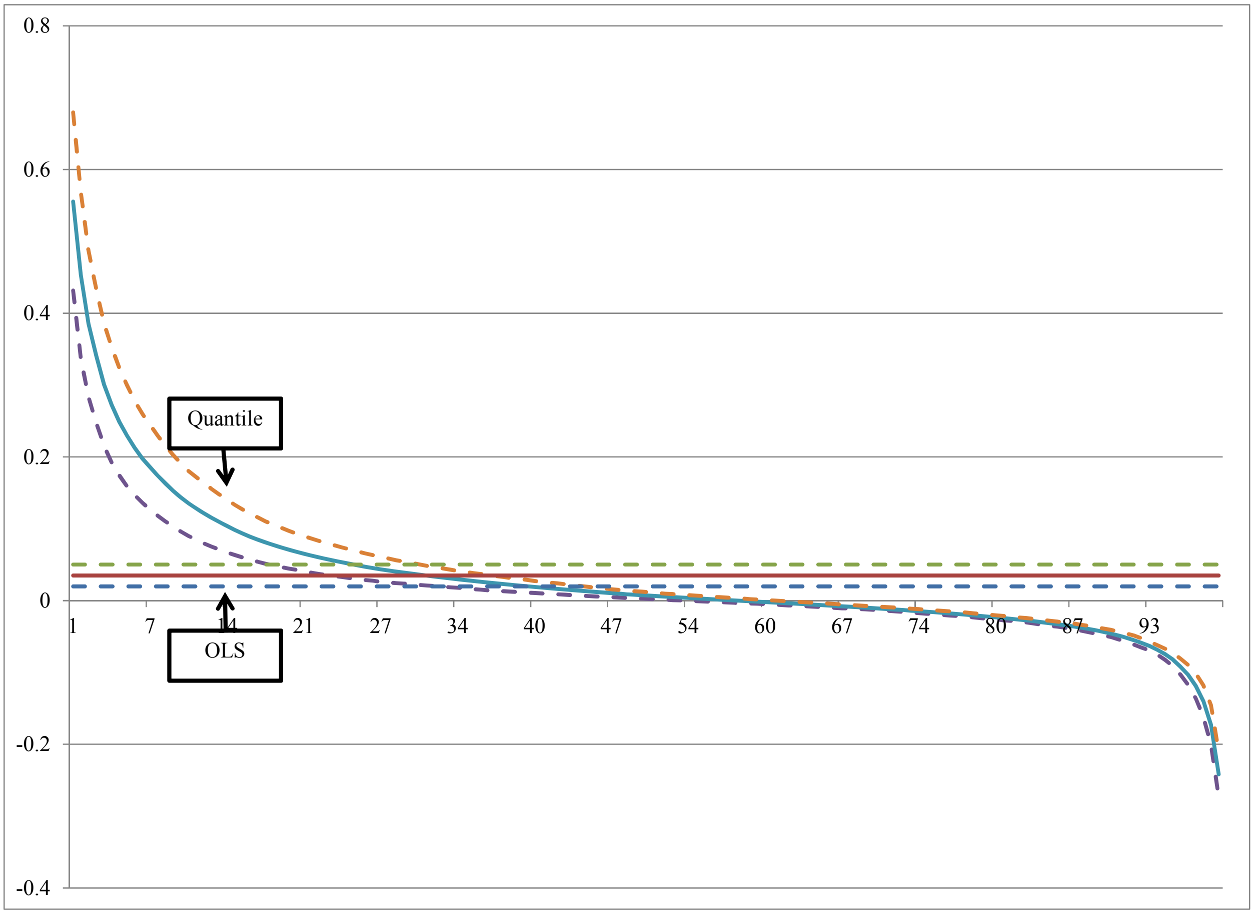
$\beta^{q}_{4,AVG}$ is the average coefficient on $ACC_{i,t-1}$. Quantiles are shown on the x-axis and values of the coefficient are shown on the y-axis. The solid line is the average across estimation years of the estimates of the coefficient. Dashed lines equal the average ± 1.96 multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Six - Graph of $\beta^{q}_{5,AVG}$ on q



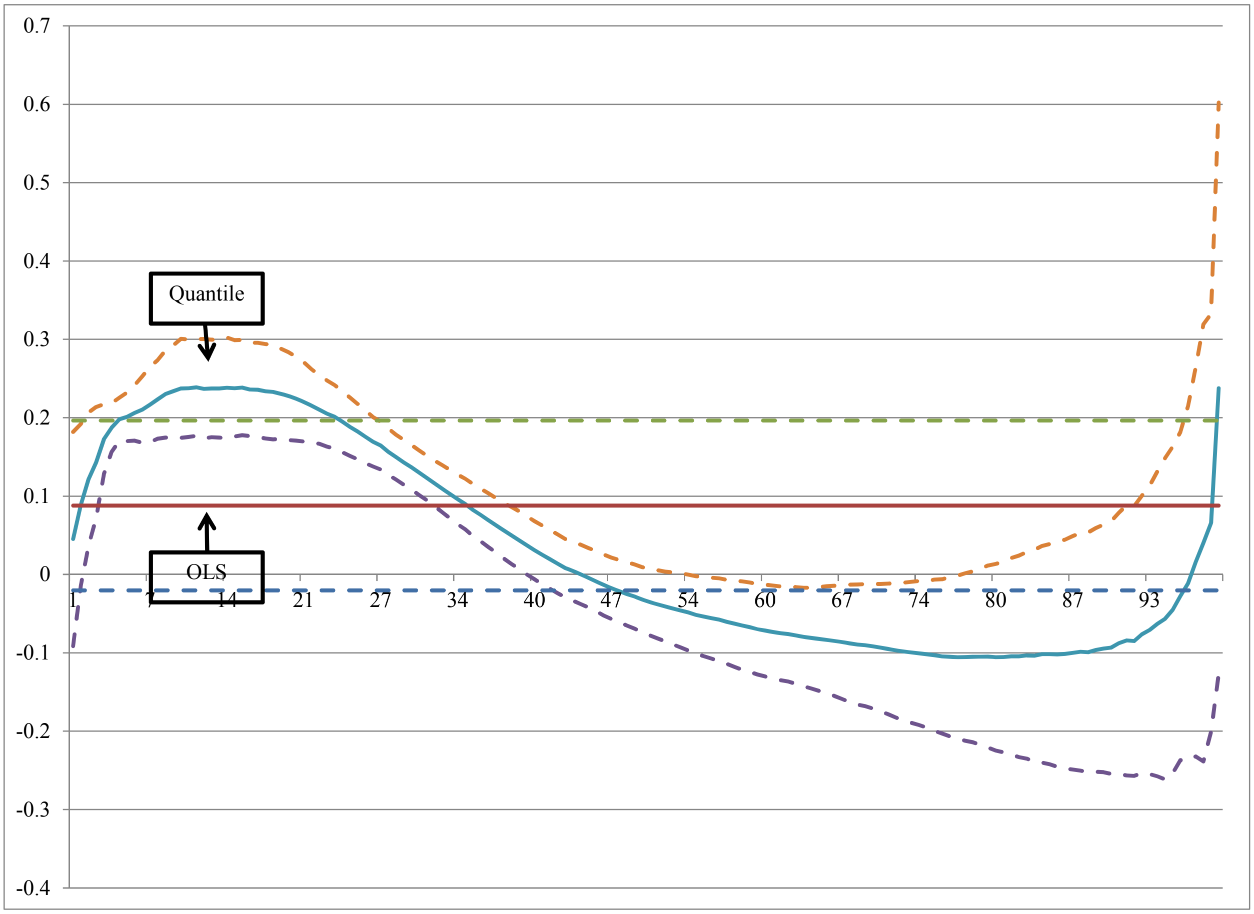
$\beta^{q}_{5,AVG}$ is the average coefficient on $LEV_{i,t-1}$. Quantiles are shown on the x-axis and values of the coefficient are shown on the y-axis. The solid line is the average across estimation years of the estimates of the coefficient. Dashed lines equal the average ± 1.96 multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Seven - Graph of $\beta_{6,AVG}^{q}$ on q



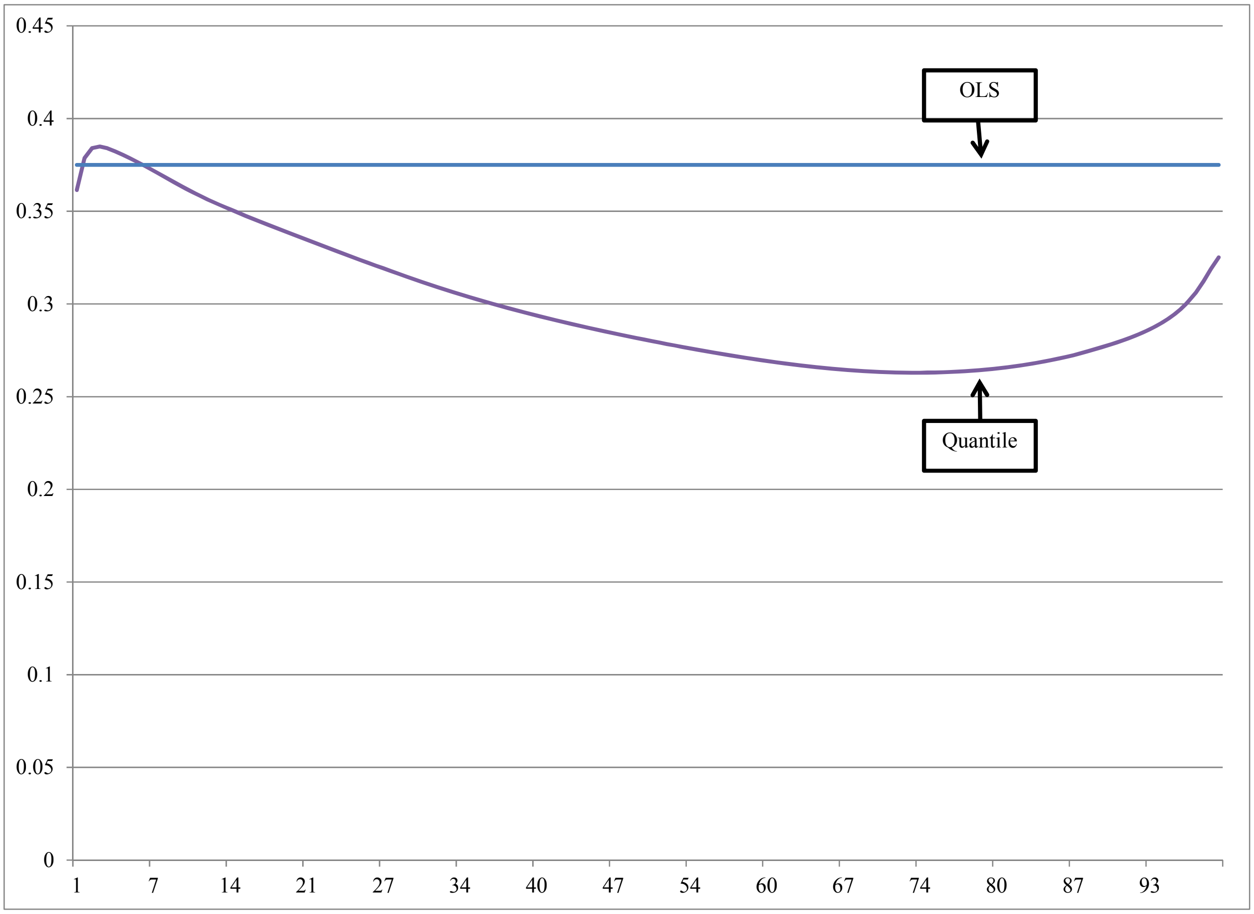
$\beta_{6,AVG}^{q}$ is the average coefficient on $PAYER_{i,t-1}$. Quantiles are shown on the x-axis and values of the coefficient are shown on the y-axis. The solid line is the average across estimation years of the estimates of the coefficient. Dashed lines equal the average ± 1.96 multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Eight - Graph of $\beta^{q}_{7,AVG}$ on q



$\beta^{q}_{7,AVG}$ is the average coefficient on $PAYOUT_{i,t-1}$. Quantiles are shown on the x-axis and values of the coefficient are shown on the y-axis. The solid line is the average across estimation years of the estimates of the coefficient. Dashed lines equal the average ± 1.96 multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Figure Nine - Quantile Regression Pseudo R-squared and OLS R-squared



Quantiles are shown on the x-axis and values of the r-squareds are shown on the y-axis.

Table One – Descriptive Statistics for the Estimation Sample

Panel A – Descriptive Statistics for the Estimation Sample Pertaining to a One-year Forecast Horizon

	Mean	Standard Deviation	Minimum	p1	p10	p25	p50	p75	p90	p99	Maximum	N
$ROE_{i,t}$	0.031	0.331	-2.000	-1.308	-0.328	-0.015	0.101	0.178	0.273	0.698	1.994	174,215
$ROE_{i,t-1}$	0.024	0.301	-2.000	-1.310	-0.265	0.006	0.096	0.155	0.222	0.511	1.977	174,215
$LOSS_{i,t-1}$	0.240	0.427	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	174,215
$ROE_{i,t-1}\times LOSS_{i,t-1}$	-0.084	0.244	-2.000	-1.310	-0.265	0.000	0.000	0.000	0.000	0.000	0.000	174,215
$ACC_{i,t-1}$	-0.059	0.312	-1.997	-1.122	-0.345	-0.165	-0.054	0.060	0.234	0.837	1.984	174,215
$LEV_{i,t-1}$	2.408	1.598	1.000	1.053	1.224	1.471	1.970	2.783	3.876	9.242	19.996	174,215
$PAYER_{i,t-1}$	0.438	0.496	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	174,215
$PAYOUT_{i,t-1}$	0.024	0.049	0.000	0.000	0.000	0.000	0.000	0.037	0.072	0.192	0.994	174,215

$ROE_{i,t}$ is earnings of firm i during year t divided by firm i's year t-1 equity book value. $ROE_{i,t-1}$ is earnings of firm i during year t-1 divided by firm i's year t-1 equity book value. $LOSS_{i,t-1}$ is an indicator variable that equals one (zero) if $ROE_{i,t-1} < 0$ ($ROE_{i,t-1} \geq 0$). $ACC_{i,t-1}$ is accruals reported by firm i during year t-1 divided by firm i's year t-1 equity book value. $LEV_{i,t-1}$ is total assets of firm i for year t-1 divided by firm i's year t-1 equity book value. $PAYER_{i,t-1}$ is an indicator variable that equals one (zero) if $PAYOUT_{i,t-1} > 0$ ($PAYOUT_{i,t-1} = 0$). $PAYOUT_{i,t-1}$ is dividends paid by firm i during year t-1 divided by firm i's year t-1 equity book value.

Panel B – Cross-sectional Correlations

	$ROE_{i,t}$	$ROE_{i,t-1}$	$LOSS_{i,t-1}$	$ROE_{i,t-1} \times LOSS_{i,t-1}$	$ACC_{i,t-1}$	$LEV_{i,t-1}$	$PAYER_{i,t-1}$	$PAYOUT_{i,t-1}$
$ROE_{i,t}$		0.60	-0.42	0.42	0.11	0.01	0.20	0.22
		(30.41)	(-10.00)	(6.70)	(5.83)	(0.43)	(5.92)	(20.48)
$ROE_{i,t-1}$	0.70		-0.65	0.85	0.28	-0.06	0.25	0.27
	(35.42)		(-23.04)	(15.6)	(10.55)	(-1.89)	(13.08)	(9.75)
$LOSS_{i,t-1}$	-0.43	-0.63		-0.61	-0.20	0.08	-0.30	-0.19
	(-7.79)	(-7.58)		(-68.54)	(-10.31)	(2.83)	(-19.52)	(-16.46)
$ROE_{i,t-1} \times LOSS_{i,t-1}$	0.44	0.65	-0.99		0.26	-0.15	0.22	0.13
	(7.47)	(7.38)	(-199.46)		(11.92)	(-5.64)	(14.11)	(13.84)
$ACC_{i,t-1}$	0.12	0.22	-0.20	0.21		-0.16	-0.03	-0.06
	(5.25)	(9.93)	(-10.08)	(10.33)		(-3.59)	(-2.16)	(-4.80)
$LEV_{i,t-1}$	0.08	0.05	0.01	-0.02	-0.21		-0.01	0.08
	(2.65)	(1.81)	(0.34)	(-0.64)	(-8.29)		(-0.22)	(1.79)
$PAYER_{i,t-1}$	0.23	0.28	-0.30	0.31	-0.06	0.07		0.58
	(6.52)	(9.29)	(-19.52)	(18.53)	(-3.27)	(1.02)		(27.91)
$PAYOUT_{i,t-1}$	0.31	0.36	-0.29	0.29	-0.09	0.11	0.88	
	(21.23)	(23.72)	(-13.2)	(12.85)	(-5.41)	(2.00)	(18.21)	

Pearson product moment (Spearman rank order) correlations are shown above (below) the diagonal. Correlations are calculated as the means of annual the correlations. t-statistics are shown in parentheses. A particular t-statistic equals the mean of the annual correlation divided by its standard error. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Table Two – Descriptive Statistics for Quantile-based Out-of-sample Forecasts

Panel A – Descriptive Statistics for One-year Forecasts Horizon

	Mean	**Standard Deviation**	**Minimum**	**p1**	**p10**	**p25**	**p50**	**p75**	**p90**	**p99**	**Maximum**	**N**
$Q\_MEAN_{i,t,t+1}$	0.042	0.237	-1.334	-0.845	-0.231	-0.043	0.097	0.166	0.233	0.520	2.047	170,522
$Q\_PROB_{i,t,t+1}$	0.251	0.317	0.000	0.000	0.000	0.000	0.093	0.553	0.807	0.927	1.000	170,522
$Q\_STD_{i,t,t+1}$	0.123	0.131	0.002	0.006	0.023	0.042	0.076	0.159	0.279	0.655	1.231	170,522
$Q\_CV_{i,t,t+1}$	2.366	221.964	0.012	0.039	0.138	0.292	0.716	1.267	2.240	13.000	90,342.540	170,522
$Q\_SKEW_{i,t,t+1}$	0.192	2.032	-8.423	-4.147	-1.376	-0.881	-0.590	0.586	3.780	5.226	8.103	170,522
$Q\_KURT_{i,t,t+1}$	6.243	9.215	-1.572	-1.036	-0.599	0.662	1.748	8.993	20.703	36.374	85.842	170,522

$Q\_MEAN_{i,t,t+1}$ is the year t estimate of the mean of $ROE_{i,t+1}$. $Q\_PROB_{i,t,t+1}$ is the year t estimate of the probability that $ROE_{i,t+1} < 0$. $Q\_STD_{i,t,t+1}$ is the year t estimate of the standard deviation of $ROE_{i,t+1}$. $Q\_CV_{i,t,t+1} = Q\_STD_{i,t,t+h} \div |Q\_MEAN_{i,t,t+h}|$. $Q\_SKEW_{i,t,t+1}$ is the year t estimate of the skewness of $ROE_{i,t+1}$. $Q\_KURT_{i,t,t+1}$ is the year t estimate of the excess kurtosis of $ROE_{i,t+1}$.

Panel B – Cross-sectional Correlations

	$Q\_MEAN_{i,t,t+1}$	$Q\_PROB_{i,t,t+1}$	$Q\_STD_{i,t,t+1}$	$Q\_SKEW_{i,t,t+1}$	$Q\_KURT_{i,t,t+1}$
$Q\_MEAN_{i,t,t+1}$		-0.80	-0.62	0.20	0.27
		(-30.58)	(-9.97)	(6.56)	(3.96)
$Q\_PROB_{i,t,t+1}$	-0.81		0.69	-0.33	-0.45
	(-13.57)		(67.26)	(-6.12)	(-9.08)
$Q\_STD_{i,t,t+1}$	-0.47	0.78		-0.26	-0.38
	(-6.90)	(58.01)		(-9.50)	(-14.06)
$Q\_SKEW_{i,t,t+1}$	0.21	-0.41	-0.34		0.67
	(3.70)	(-7.79)	(-9.03)		(4.21)
$Q\_KURT_{i,t,t+1}$	0.38	-0.57	-0.63	0.14	
	(2.83)	(-4.60)	(-7.93)	(1.24)	

Pearson product moment (Spearman rank order) correlations are shown above (below) the diagonal. Correlations are calculated as the means of annual the correlations. t-statistics are shown in parentheses. A particular t-statistic equals the mean of the annual correlation divided by its standard error. We use Newey-West adjusted standard errors assuming a lag-length of ten years.

Table Three – Regressions of $LOSS_{i,t+1}$ on $Q\_PROB_{i,t,t+1}$ and the Logit-based Estimate

	A = Logit					
	(1)		(2)		(3)	
$Q\_PROB_{i,t,t+1}$	0.88	***			0.43	***
	(39.30)				(9.91)	
$A\_PROB_{i,t,t+1}$			1.03	***	0.54	***
			(54.70)		(8.59)	
Intercept	0.08	***	0.01		0.04	***
	(8.14)		(0.99)		(8.79)	
R-squared	0.32		0.32		0.33	
Vuong	-0.99					
Firm-years	156,973					
Years	38					

$LOSS_{i,t+1}$ is the dependent variable. It is an indicator variable that equals one (zero) if $ROE_{i,t+1} < 0$ ($ROE_{i,t+1} \geq 0$). $Q\_PROB_{i,t,t+1}$ is the quantile-based year t estimate of the probability that $ROE_{i,t+1} < 0$. $A\_PROB_{i,t,t+1}$ is Logit-based year t estimate of the probability that $ROE_{i,t+1} < 0$.

Reported regression coefficients equal the average of the coefficients obtained from annual cross-sectional regressions. t-statistics are shown in parentheses. t-statistics equal the average coefficient divided by its temporal standard error. Vuong statistics are obtained by computing the average of the annual slope coefficients obtained from the regression described on p. 318 of Vuong [1989], and then dividing the average by the temporal standard error of the coefficients. All temporal standard errors reflect the Newey-West adjustment assuming a ten-year lag length. *,** and *** represent rejection of a two-sided test at the five percent level, one percent level and 0.50 percent level, respectively.

Table Four – Regressions of $|FERR_{i,t+1}|$ on $Q\_STD_{i,t,t+1}$ and Alternative Estimates

	A = Historical Firm-level			A = Historical Matched-sample		
	(1)	(2)	(3)	(4)	(5)	(6)
$Q\_STD_{i,t,t+1}$	1.27 ***		1.27 ***	1.26 ***		1.13 ***
	(22.21)		(21.97)	(19.71)		(20.73)
$A\_STD_{i,t,t+1}$		0.09	0.02 *		2.90 ***	0.85 ***
		(1.74)	(2.28)		(16.22)	(11.16)
Intercept	0.01 ***	0.15 ***	0.01 ***	0.01	0.10 ***	0.01 *
	(4.57)	(6.92)	(4.41)	(1.29)	(6.39)	(2.09)
R-squared	0.21	0.00	0.21	0.21	0.09	0.21
Vuong	16.73 ***			16.40 ***		
Firm-years	134,552			85,142		
Years	38			38		

$|FERR_{i,t+1}|$ is the dependent variable. It equals the absolute value of the difference between $ROE_{i,t+1}$ and $Q\_MEAN_{i,t+1}$ ($Q\_MEAN_{i,t,t+1}$ is the year t estimate of the mean of $ROE_{i,t+1}$). $Q\_STD_{i,t,t+1}$ is the quantile-based year t estimate of the standard deviation of $ROE_{i,t+1}$. In columns (2) and (3) $A\_STD_{i,t,t+1}$ is the estimate of the standard deviation of $ROE_{i,t+1}$ obtained from the historical firm-level approach. In columns (5) and (6) $A\_STD_{i,t,t+1}$ is the estimate of the standard deviation of $ROE_{i,t+1}$ obtained from the historical matched-sample approach. The historical firm-level and historical matched-sample approaches are described on p. 26 and pp. 43-44 of Appendix B.

Reported regression coefficients equal the average of the coefficients obtained from annual cross-sectional regressions. t-statistics are shown in parentheses. t-statistics equal the average coefficient divided by its temporal standard error. Vuong statistics are obtained by computing the average of the annual slope coefficients obtained from the regression described on p. 318 of Vuong [1989], and then dividing the average by the temporal standard error of the coefficients. All temporal standard errors reflect the Newey-West adjustment assuming a ten-year lag length. *,** and *** represent rejection of a two-sided test at the five percent level, one percent level and 0.50 percent level, respectively.

Table Five – Regressions of $LOSS_{IND,t+1}$ on $Q\_PROB_{IND,t,t+1}$ and Logit-based Estimates

	A = Logit					
	(1)		(2)		(3)	
$Q\_PROB_{IND,t,t+1}$	1.09	***			0.92	***
	(18.96)				(7.71)	
$A\_PROB_{IND,t,t+1}$			1.20	***	0.18	
			(19.60)		(1.02)	
Intercept	0.03	***	-0.04	*	0.02	
	(3.93)		(-2.18)		(1.19)	
R-squared	0.54		0.53		0.54	
Vuong	3.70			***		
Industry-years	2,056					
Years	38					

$LOSS_{IND,t+1}$ is the dependent variable. It equals the industry average of $LOSS_{i,t+1}$, which is an indicator variable that equals one (zero) if $ROE_{i,t+1} < 0$ ($ROE_{i,t+1} \geq 0$). $Q\_PROB_{IND,t,t+1}$ is the industry average of $Q\_PROB_{i,t,t+1}$, which is the quantile-based year t estimate of the probability that $ROE_{i,t+1} < 0$. $A\_PROB_{IND,t,t+1}$ is the industry average of $A\_PROB_{i,t,t+1}$, which is the Logit-based year t estimate of the probability that $ROE_{i,t+1} < 0$.

Reported regression coefficients equal the average of the coefficients obtained from annual cross-sectional regressions. t-statistics are shown in parentheses. t-statistics equal the average coefficient divided by its temporal standard error. Vuong statistics are obtained by computing the average of the annual slope coefficients obtained from the regression described on p. 318 of Vuong [1989], and then dividing the average by the temporal standard error of the coefficients. All temporal standard errors reflect the Newey-West adjustment assuming a ten-year lag length. *,** and *** represent rejection of a two-sided test at the five percent level, one percent level and 0.50 percent level, respectively.

Table Six – Regressions of Realized Industry-level Year t+1 Moments on the Year t Out-of Sample Estimates

Panel A – Regressions of $R\_STD_{IND,t+1}$ on $Q\_STD_{IND,t,t+1}$ and Alternative Estimates

	A = Historical Firm-level			A = Historical Matched-sample			A = Historical Industry-level		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Q\_STD_{IND,t,t+1}$	1.41 ***		1.39 ***	1.49 ***		1.33 ***	2.20 ***		2.15 ***
	(16.39)		(14.56)	(18.46)		(11.53)	(3.99)		(3.67)
$A\_STD_{IND,t,t+1}$		0.13	-0.01		0.62 ***	0.14		0.08 *	0.03 **
		(1.87)	(-0.23)		(6.56)	(1.76)		(2.15)	(2.72)
Intercept	0.01	0.31 ***	0.01	-0.01	0.17 ***	-0.01	-0.14	0.41 ***	-0.14
	(1.04)	(7.63)	(1.50)	(-0.52)	(4.47)	(-1.17)	(-1.27)	(4.39)	(-1.23)
R-squared	0.29	0.04	0.31	0.33	0.22	0.37	0.25	0.06	0.27
Vuong	5.68 ***			4.27 ***			6.13 ***		
Industry-years	1,960			1,579			2,056		
Years	38			38			38		

$R\_STD_{IND,t+1}$ is the dependent variable. It equals the realized within-industry standard deviation of $ROE_{i,t+1}$. $Q\_STD_{IND,t+1}$ is calculated by applying the law of total variance to the year t quantile-based estimates of the mean and variance of $ROE_{i,t+1}$. In columns (2) and (3) $A\_STD_{i,t,t+1}$ is calculated by applying the law of total variance to the year t historical firm-level estimates of the mean and variance of $ROE_{i,t+1}$. In columns (5) and (6) $A\_STD_{i,t,t+1}$ is calculated by applying the law of total variance to the year t historical matched-sample estimates of the mean and variance of $ROE_{i,t+1}$. In columns (8) and (9) $A\_STD_{i,t,t+1}$ equals the within-industry standard deviation of $ROE_{i,t}$. The historical firm-level and historical matched-sample approaches are described on p. 26 and pp. 43-44 of Appendix B. The historical industry-level approach is described on p. 30 and pp. 44-45 of Appendix B. The law of total variance is described on p. 29.

Panel B – Regressions of $R\_SKEW_{IND,t+1}$ on $Q\_SKEW_{IND,t,t+1}$ and Alternative Estimates

	A = Historical Firm-level			A = Historical Matched-sample			A = Historical Industry-level		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Q\_SKEW_{IND,t,t+1}$	0.56 ***		0.56 ***	0.56 ***		0.53 ***	0.51 ***		0.49 ***
	(14.70)		(14.88)	(7.19)		(7.47)	(17.06)		(13.40)
$A\_SKEW_{IND,t,t+1}$		0.01	0.00		0.28 ***	0.08 ***		0.11 **	0.05 ***
		(1.61)	(-0.10)		(13.18)	(3.62)		(3.28)	(4.41)
Intercept	-0.51 ***	-1.43 ***	-0.50 ***	-0.30 ***	-1.14 ***	-0.27 ***	-0.68 ***	-1.50 ***	-0.69 ***
	(-4.42)	(-7.19)	(-4.42)	(-5.47)	(-9.02)	(-6.41)	(-4.51)	(-8.56)	(-4.98)
R-squared	0.14	0.00	0.15	0.16	0.08	0.18	0.13	0.03	0.14
Vuong	10.77 ***			2.86 **			7.48 ***		
Industry-years	1,960			1,579			2,056		
Years	38			38			38		

$R\_SKEW_{IND,t+1}$ is the dependent variable. It equals the realized within-industry skewness of $ROE_{i,t+1}$. $Q\_SKEW_{IND,t+1}$ is calculated by applying the law of total moments to the year t quantile-based estimates of the mean, variance and third central moment of $ROE_{i,t+1}$. In columns (2) and (3) $A\_SKEW_{i,t,t+1}$ is calculated by applying the law of total moments to the year t historical firm-level estimates of the mean, variance and third central moment of $ROE_{i,t+1}$. In columns (5) and (6) $A\_SKEW_{i,t,t+1}$ is calculated by applying the law of total moments to the year t historical matched-sample estimates of the mean, variance and third central moment of $ROE_{i,t+1}$. In columns (8) and (9) $A\_SKEW_{i,t,t+1}$ equals the within-industry skewness of $ROE_{i,t}$. The historical firm-level and historical matched-sample approaches are described on p. 26 and pp. 43-44 of Appendix B. The historical industry-level approach is described on p. 30 and pp. 44-45 of Appendix B. The law of total moments is described on pp. 39-41 of Appendix A.

Panel C – Regressions of $R\_KURT_{IND,t+1}$ on $Q\_KURT_{IND,t,t+1}$ and Alternative Estimates

	A = Historical Firm-level			A = Historical Matched-sample			A = Historical Industry-level		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Q\_KURT_{IND,t,t+1}$	0.67 ***		0.64 ***	0.41 ***		0.25 **	0.54 **		0.49 ***
	(4.43)		(4.54)	(4.64)		(3.09)	(3.07)		(3.63)
$A\_KURT_{IND,t,t+1}$		0.03 *	0.03 *		0.46 ***	0.40 ***		0.22 ***	0.19 ***
		(2.06)	(2.03)		(6.35)	(5.34)		(4.93)	(4.91)
Intercept	3.77 ***	10.92 ***	2.80 **	4.91 ***	5.94 ***	3.40 ***	6.52 *	8.83 ***	2.55 *
	(3.66)	(17.97)	(2.82)	(5.41)	(10.18)	(4.13)	(2.32)	(20.22)	(2.20)
R-squared	0.17	0.08	0.24	0.17	0.18	0.25	0.14	0.18	0.28
Vuong	1.17			0.17			-0.61		
Industry-years		1,960			1,579			2,056	
Years		38			38			38	

$R\_KURT_{IND,t+1}$ is the dependent variable. It equals the realized within-industry kurtosis of $ROE_{i,t+1}$. $Q\_KURT_{IND,t+1}$ is calculated by applying the law of total moments to the year t quantile-based estimates of the mean, variance, third central moment and fourth central moment of $ROE_{i,t+1}$. In columns (2) and (3) $A\_KURT_{i,t,t+1}$ is calculated by applying the law of total moments to the year t historical firm-level estimates of the mean, variance, third central moment and fourth central moment of $ROE_{i,t+1}$. In columns (5) and (6) $A\_KURT_{i,t,t+1}$ is calculated by applying the law of total moments to the year t historical matched-sample estimates of the mean, variance, third central moment and fourth central moment of $ROE_{i,t+1}$. In columns (8) and (9) $A\_SKEW_{i,t,t+1}$ equals the within-industry kurtosis of $ROE_{i,t}$. The historical firm-level and historical matched-sample approaches are described on p. 26 and pp. 43-44 of Appendix B. The historical industry-level approach is described on p. 29-30 and pp. 44-45 of Appendix B. The law of total moments is described on pp. 39-41 of Appendix A.

Reported regression coefficients equal the average of the coefficients obtained from annual cross-sectional regressions. t-statistics are shown in parentheses. t-statistics equal the average coefficient divided by its temporal standard error. Vuong statistics are obtained by computing the average of the annual slope coefficients obtained from the regression described on p. 318 of Vuong [1989], and then dividing the average by the temporal standard error of the coefficients. All temporal standard errors reflect the Newey-West adjustment assuming a ten-year lag length. *,** and *** represent rejection of a two-sided test at the five percent level, one percent level and 0.50 percent level, respectively.

Table Seven – Regressions of Valuation Multiples on Quantile-based Out-of Sample Estimates

	Earnings-to-price				Book-to-price			
	(1)		(2)		(1)		(2)	
$Q\_MEAN_{i,t,t+1}$	0.57	***	0.53	***	-2.11	**	-2.09	**
	(10.46)		(9.96)		(-2.85)		(-2.77)	
$Q\_STD_{i,t,t+1}$	-0.51	***	-0.50	***	-1.91	***	-1.93	***
	(-4.78)		(-4.93)		(-6.89)		(-10.18)	
$Q\_SKEW_{i,t,t+1}$	-0.03	***	-0.03	***	-0.07	***	-0.06	***
	(-5.1)		(-5.1)		(-4.71)		(-6.16)	
$Q\_KURT_{i,t,t+1}$	0.01	***	0.01	***	0.02	*	0.01	*
	(4.14)		(4.12)		(2.33)		(2.47)	
$RET_{i,t}$			0.01	***			-0.16	***
			(3.85)				(-4.5)	
$RET\_VOL_{i,t}$			-0.07				-0.18	
			(-1.83)				(-0.55)	
$RET\_SKEW_{i,t}$			0.01	***			0.04	***
			(4.33)				(5.19)	
$RET\_KURT_{i,t}$			0.00				0.02	***
			(-0.47)				(3.77)	
Intercept	-0.02		0.00		1.13	***	1.10	***
	(-1.43)		(-0.33)		(4.94)		(4.81)	
R-squared	0.58		0.58		0.14		0.21	
Firm-years	155,025		125,468		155,025		125,468	
Years	39				39			

Earnings-to-price (book-to-price) is the dependent variable. Earnings-to-price (book-to-price) equals the ratio of firm i's year t earnings (equity book value) to its equity market value at the end year t. $Q\_MEAN_{i,t,t+1}$ is the year t quantile-based estimate of the mean of $ROE_{i,t+1}$. $Q\_STD_{i,t,t+1}$ is the year t quantile-based estimate of the standard deviation of $ROE_{i,t+1}$. $Q\_SKEW_{i,t,t+1}$ is the year t quantile-based estimate of the skewness of $ROE_{i,t+1}$. $Q\_KURT_{i,t,t+1}$ is the year t quantile-based estimate of the excess kurtosis of $ROE_{i,t+1}$. $RET_{i,t}$ is the annual stock return of firm i in year t. $RET\_VOL_{i,t}$ the volatility of historical market-model residuals for firm i. $RET\_SKEW_{i,t}$ is the skewness of historical market-model residuals for firm i. $RET\_KURT_{i,t}$.is the kurtosis of historical market-model residuals for firm i.

Reported regression coefficients equal the average of the coefficients obtained from annual cross-sectional regressions. t-statistics are shown in parentheses. t-statistics equal the average coefficient divided by its temporal standard error. Vuong statistics are obtained by computing the average of the annual slope coefficients obtained from the regression described on p. 318 of Vuong [1989], and then dividing the average by the temporal standard error of the coefficients. All temporal standard errors reflect the Newey-West adjustment assuming a ten-year lag length. *,** and *** represent rejection of a two-sided test at the five percent level, one percent level and 0.50 percent level, respectively.

Table Eight – Regressions of Credit Ratings on Quantile-based Out-of Sample Estimates

	(1)		(2)	
$Q\_MEAN_{i,t,t+1}$	-1.40	***	-1.48	***
	(-5.94)		(-7.98)	
$Q\_STD_{i,t,t+1}$	1.17	*	1.09	*
	(2.61)		(2.29)	
$Q\_SKEW_{i,t,t+1}$	-1.11	***	-1.10	***
	(-11.97)		(-11.72)	
$Q\_KURT_{i,t,t+1}$	0.18	***	0.18	***
	(4.82)		(4.78)	
$LN\_SIZE_{i,t}$	-0.74	***	-0.74	***
	(-11.31)		(-11.59)	
$LIAB\_ASST_{i,t}$	1.08	*	1.16	*
	(2.65)		(2.52)	
$EBITDA\_LIAB_{i,t}$	-2.95	***	-2.93	***
	(-11.39)		(-13.29)	
$RET_{i,t}$	0.60	***	0.64	***
	(14.35)		(12.66)	
$RET\_VOL_{i,t}$	21.47	***	22.80	***
	(16.94)		(15.06)	
$RET\_SKEW_{i,t}$			-0.10	***
			(-5.15)	
$RET\_KURT_{i,t}$			-0.09	***
			(-3.7)	
Intercept	2.10	*	2.23	**
	(2.75)		(2.95)	
R-squared	0.64		0.64	
Firm-years	20,937		20,148	
Years	27			

Credit ratings are the dependent variable. They are obtained from Standard and Poors. They range between 2 and 23. Higher ratings reflect worse credit quality; for example, a rating of 2 (23) implies a letter rating of AAA (D). $Q\_MEAN_{i,t,t+1}$ is the year t quantile-based estimate of the mean of $ROE_{i,t+1}$. $Q\_STD_{i,t,t+1}$ is the year t quantile-based estimate of the standard deviation of $ROE_{i,t+1}$. $Q\_SKEW_{i,t,t+1}$ is the year t quantile-based estimate of the skewness of $ROE_{i,t+1}$. $Q\_KURT_{i,t,t+1}$ is the year t quantile-based estimate of the excess kurtosis of $ROE_{i,t+1}$. $LN\_SIZE_{i,t}$ is the natural log of the ratio of firm i's year t equity market to the sum of all firm's contemporaneous equity market values. $LIAB\_ASST_{i,t}$ is the ratio of firm i's year t liabilities to its year t assets. $EBITDA\_LIAB_{i,t}$ is the ratio of firm i's year t earnings before interest, taxes, depreciation, and amortization to its year t liabilities. $RET_{i,t}$ is the annual stock return of firm i in year t. $RET\_VOL_{i,t}$ the volatility of historical market-model residuals for firm i. $RET\_SKEW_{i,t}$ is the skewness of historical market-model residuals for firm i. $RET\_KURT_{i,t}$.is the kurtosis of historical market-model residuals for firm i.

Reported regression coefficients equal the average of the coefficients obtained from annual cross-sectional regressions. t-statistics are shown in parentheses. t-statistics equal the average coefficient divided by its temporal standard error. Vuong statistics are obtained by computing the average of the annual slope coefficients obtained from the regression described on p. 318 of Vuong [1989], and then dividing the average by the temporal standard error of the coefficients. All temporal standard errors reflect the Newey-West adjustment assuming a ten-year lag length. *,** and *** represent rejection of a two-sided test at the five percent level, one percent level and 0.50 percent level, respectively.