The Higher Moments of Future Earnings*

Woo-Jin Chang
INSEAD†
woo-jin.chang@insead.edu
+33 (0)1 60 72 92 48

Steven J. Monahan
INSEAD
steven.monahan@insead.edu
+33 (0)1 60 72 92 14

Amine Ouazad
INSEAD
amine.ouazad@insead.edu
+33 (0)1 60 72 48 49

Florin Vasvari
London Business School‡
fvasvari@london.edu
+44 (0)20 7000 8117

This Draft: November 2014

* We benefited from conversations with Federico Gavazzoni, Lucie Tepla and Astrid Schornick, and the comments of Daniel Bens, Richard Frankel, Joseph Gerakos, Bjorn Jorgensen, Maria Ogneva, Stephen Penman, Stephen Ryan, Gil Sadka, Jerold Zimmerman (the editor), an anonymous reviewer, and seminar participants at INSEAD’s Finance Brownbag, the 2013 INSEAD Accounting Symposium, the 2013 London Business School Accounting Symposium, New York University, The Ohio State University, and WHU — Otto Beisheim School of Management.
† Europe Campus, Boulevard de Constance, 77305 Fontainebleau, France
‡ Regent's Park, London, NW1 4SA, United Kingdom
The Higher Moments of Future Earnings

Abstract

We use quantile regressions to evaluate the higher moments of future earnings. First, we evaluate the in-sample relations between current firm-level attributes and the moments of lead return on equity, ROE. We show that: (1) as current ROE 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 predictions of the standard deviation, skewness, and kurtosis of lead ROE. Moreover, when compared to predictions obtained via alternative approaches, our out-of-sample predictions: always contain incremental information content and are typically more reliable. Finally, we evaluate the relation between higher moments and market-based variables. These analyses demonstrate that equity prices are increasing in the variance and skewness of lead ROE but decreasing in the kurtosis of lead ROE. Credit spreads are increasing in the variance and kurtosis of lead return on assets, ROA, and decreasing in the skewness of lead ROA.

Keywords: Book-to-price ratio, bond ratings, bond yields, credit-default-swap spreads, earnings earnings-to-price ratio, implied cost of capital, forecasting, higher moments, quantile regressions, return on assets, return on equity, valuation multiples

JEL classification codes: C21, C53, G17, M41
1. **Introduction**

Earnings are a key economic variable and numerous studies in accounting and finance evaluate the relation between earnings and firm-level attributes. There is also a large literature that evaluates different approaches for forecasting earnings.\(^1\)

A limitation of these studies is that virtually all of them focus (implicitly or explicitly) on expected—i.e., the mean of—future earnings.\(^2\) However, as discussed in section two, extant analytical results imply that a number of interesting empirical questions can be asked about the economic relevance of the higher moments of future earnings. 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 earnings? Second, can these in-sample relations be used to develop reliable out-of-sample predictions of the higher moments of lead earnings?

We use a novel approach for estimating the higher moments of future earnings that is based on quantile regressions. As discussed in section three, quantile regressions are particularly appropriate in our setting for two reasons. First, the coefficients obtained from a sequence of quantile regressions can be used to infer the relation between firm-level attributes and the location and shape of the distribution of a firm’s lead earnings. Second, fitted values from quantile regressions can be used to generate consistent firm-level predictions of the moments of lead earnings.

---

\(^1\) A non-exhaustive list of studies that focus on earnings metrics includes: Freeman et al. [1982], Fama and French [2000, 2006], Fairfield and Yohn [2001], Nissim and Penman [2001, 2003], Banker and Chen [2006], Hou and Robinson [2006], Soliman [2008], Fairfield et al., [2009], Esplin et al. [2014], Hou et al. [2012], Li and Mohanram [2012], Penman [2012], and Gerakos and Gramercy [2013].

\(^2\) A contemporaneous study by Konstantinidi and Pope [2014] is a notable exception. We describe the differences between our study and the study by Konstantinidi and Pope [2014] in section two.
In our first set of analyses, which are described in section five, we document the in-sample relations between the moments of lead return on equity, 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, we find three noteworthy results. 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 that lead ROE is riskier. 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 are also 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]. We also show that dividend-paying firms have distributions of lead ROE that are less disperse and more positively skewed. This supports the dividend-smoothing hypothesis described in Lintner [1956] and the 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.

In our second set of analyses, which are discussed in section six, we validate our out-of-sample predictions of the higher moments of future earnings. We conduct our tests at the industry-year level. We do this because the realized moments of earnings are not observable at
the firm-year level. In particular, we use the law of total moments described in Brillinger [1969] to construct predictions of industry-year moments from contemporaneous predictions of firm-year moments. We then compare the industry-level predictions to the future realized industry-level moments.

The results of the industry-level tests lead to the conclusion that the quantile-based predictions are reliable. In particular, each of the predicted industry-level moments is positively associated with its future realized industry-level counterpart and explains a significant portion of the cross-sectional variation in the future realized industry-level moments. These associations remain after controlling for alternative predictions; and, the quantile-based predictions are typically more reliable and never less reliable than any of the alternative predictors 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 predictions, we also provide initial evidence about the economic relevance of the higher moments of lead earnings. Specifically, in our final set of analyses, which we discuss in section seven, we evaluate the relations between the higher moments of lead earnings and security prices.

We begin by evaluating the current, firm-level implied cost of equity capital, the book-to-price ratio and the earnings-to-price ratio. We show that each of these variables is negatively associated with the standard deviation and skewness of lead ROE and positively associated with the kurtosis of lead ROE. Hence, we show that, ceteris paribus, equity investors place higher prices on shares issued by firms with future earnings that are relatively volatile, positively-skewed, and thin-tailed, which is consistent with extant analytical and empirical results (e.g.,
Next, we evaluate credit-default-swap, CDS, spreads, bond yields in the secondary market and bond ratings. We show that each of these variables has a positive association with the standard deviation and kurtosis of lead return on assets, ROA, and a negative association with the skewness of lead ROA. This implies that, *ceteris paribus*, debt investors demand higher credit-risk premiums on debt securities issued by firms with future earnings that are relatively volatile, negatively-skewed and fat-tailed. These results are consistent with extant analytical results (e.g., Merton [1974]) as well as arguments made by practitioners (e.g., Dynkin et al. [2007]).

We make several contributions to the extant literature. First, we develop a general approach based on quantile regressions to estimate the higher moments of future earnings; and we show that our approach generates reliable out-of-sample predictions of the realized moments of lead earnings. These results are important given the difficulty involved in predicting higher moments in general; and, the fact that predicting the higher moments of earnings is even more challenging given that earnings represent a summary number that is reported fairly infrequently. Second, we provide unique evidence about the relations between the higher moments of lead earnings and important firm-level attributes. Finally, we provide initial evidence about the relevance of the higher moments of future earnings for the valuation of equity and debt securities.

---

3 Bondholders have a claim on the assets, which implies that, from their perspective, ROA is the relevant performance metric. Hence, we relate CDS spreads, bond yields and bond ratings to the moments of ROA not the moments of ROE. We discuss the model we use to predict the moments of ROA in section seven.
2. Motivation and Related Literature

A standard assumption in economics is that the value of a firm (firm’s equity) is determined by the properties of the firm’s future free cash flows (dividends). Moreover, results in the accounting literature (e.g., Christensen and Feltham [2009], CR hereafter) show that future operating earnings (earnings) and free cash flows (dividends) are interchangeable in the sense that a firm (firm’s equity) can be valued on the basis of the properties of either its future free cash flows (dividends) or its future operating earnings (earnings).

Regarding equity value, it is clear that, ceteris paribus, it is increasing in the expected value—i.e., the mean or first moment—of a firm’s future earnings. Whether the higher moments of earnings are relevant is less obvious. As discussed in chapter one of Cochrane [2001] and CR, a standard assumption is that only higher co-moments matter and that the covariance between earnings and the stochastic discount factor is the only relevant co-moment. However, recent analytical results show that higher total moments are relevant. For example, Johnson [2004] shows that, for levered firms, equity value is increasing in the volatility of future earnings. On the other hand, Merton [1987] shows that, when incomplete information leads to market segmentation, equity prices are a decreasing function of volatility. The skewness of future earnings 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 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 earnings.\footnote{It is important to note that the analytical results in Merton [1987], Johnson [2004], Brunnermeier et al. [2007], Mitton and Vorkink [2007], and Barberis and Huang [2008] are based on two-period models. In period one agents invest and in period two they receive a terminal dividend and consume. Hence, although these papers often refer to the moments of returns, they also relate to the moments of either earnings or dividends. The reason for this is that, in a two-period model, the terminal price equals the terminal dividend, which, in turn, equals the terminal earnings.}

Higher moments are also potentially relevant in the context of debt valuation. As shown in Merton [1974], debt values are decreasing in asset 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 earnings that exhibit negative skewness or a higher likelihood of extreme outcomes (i.e., positive kurtosis).

More generally, the higher moments of earnings are potentially relevant in any situation where an agent’s wealth is either a direct or indirect function of earnings. This is especially true when, unlike capital-market participants, the agent cannot diversify away her exposure to risk. For example, are owners of private (or family) firms less willing to invest in projects with high downside risk? Do executives who receive earnings-based bonuses demand a higher salary when they face higher downside risk or does the firm simply limit the executives’ risk exposure? Do auditors charge higher fees when auditing a firm that has highly negatively-skewed earnings? Etc. Definitive answers to these questions require both an analytical model and valid empirical proxies. Although it is outside the scope of our study to provide the former, we do provide the latter.
The above implies there are a number of interesting empirical questions regarding the economic relevance of the higher moments of future earnings. Before these questions can be addressed, however, a reliable approach for estimating higher moments 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 implications of higher moments for the prices of equity and debt securities.

Regarding our research objective, it bears mentioning that forecasting higher moments, especially moments higher than the variance, is difficult in general. 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 \emph{ex ante} skewness (kurtosis) may have exhibited low \emph{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 earnings presents additional challenges given that earnings numbers are reported fairly infrequently. Consequently, whether we can develop reliable out-of-sample estimates of the moments of future firm-level earnings is an empirical question.

Finally, it is important that we compare our study to the contemporaneous study by Konstantinidi and Pope [2014] (KP hereafter), who also evaluate the quantiles of lead earnings. There are three key differences between our study and theirs. First, as discussed in section three and Appendix A, we develop a general approach that yields estimates of the moments of future earnings that are consistent and exhibit less finite-sample bias than the estimates developed by
KP. This is important because it allows us to draw clear-cut and in-depth inferences about the determinants and economic relevance of the higher moments of future earnings.

Second, we demonstrate that our out-of-sample predicted moments of lead earnings have construct validity. In particular, we evaluate their relation with the realized moments of lead earnings. To do this we develop a novel approach that is based on the law of total moments. This approach allows us to use industry-level tests to evaluate the reliability of our firm-level predictions. We show that there is a positive relation between each of our out-of-sample predictions and the relevant ex post realization. In addition, we also show that the quantile-based out-of-sample predictions always contain incremental explanatory power and are typically more reliable than predictions obtained from alternative approaches.

Finally, we evaluate the relation between the predicted moments of future earnings and the contemporaneous implied cost of capital, equity-valuation multiples and credit-default-swap spreads. Hence, we shed light on whether higher moments of earnings are relevant in the context of equity- and debt-valuation. Moreover, we show that different moments of future earnings matter differently and that the empirical relations we document are consistent with extant analytical results.

3. **Research Design**

   We begin by discussing how to estimate a single quantile and the properties of that estimate. Next, we provide an overview of how to use an estimate of the quantile function to develop consistent out-of-sample estimates of the moments of a random variable. Finally, we describe the empirical model we use to estimate the moments of earnings.
3.1 Estimating a Single Quantile

Quantile regression methods were introduced by Koenker and Bassett [1978] and developed by Buchinsky [1998]. These papers show that in order to estimate the \( q \)th quantile, we choose the coefficient vector \( b^q = b_0^q, \ldots, b_k^q \) that solves the minimization problem shown in equation (1).

\[
\arg\min_{b^q = b_0^q, \ldots, b_k^q} \frac{1}{N} \left\{ \sum_{i,y_{i,t} \leq \sum_{j \in 0,k} b_j^q x_{i,t-1,j}} q y_{i,t} - \sum_{j=0}^k b_j^q x_{i,t-1,j} \right\} + \sum_{i,y_{i,t} > \sum_{j \in 0,k} b_j^q x_{i,t-1,j}} (1-q) y_{i,t} - \sum_{j=0}^k b_j^q x_{i,t-1,j} \right\}
\]

In equation (1), \( y_{i,t} \) is the year \( t \)—i.e., the lead—value of the dependent variable for observation \( i \in [1,N] \), \( x_{i,t-1,j} \) is the year \( t-1 \)—i.e., the current—value of the \( j \)th independent variable \((j \in [0,k])\) for observation \( i \), and \( q \) denotes the \( q \)th quantile. Note that \( q \in Z = \{a, \ldots, q, \ldots, z\} \subset (0,1) \) and that the set \( Z \) is an ordered sequence of \( Q \) numbers. For example, if we estimate the 25th and 75th percentiles, \( Q = 2 \) and \( Z = \{0.25, 0.75\} \).

Equation (1) has a similar structure as the minimization problem underlying an ordinary least squares, OLS, regression. In particular, the objective function involves choosing 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, as shown in Koenker and Bassett [1978], the fitted

\[ \text{quant}_q \left( y_{i,t} \mid \cdot \right) = \sum_{j=0}^k b_j^q x_{i,t-1,j} \]

is an estimate of \( QUANT_q \left( y_{i,t} \mid \cdot \right) \) and \( b^q = b_0^q, \ldots, b_k^q \) is an estimate of \( B^q = B_0^q, \ldots, B_k^q \).

---

5 We use lower-case letters to refer to estimates and upper-case letters to refer to the unobservable, true values. Hence, \( \text{quant}_q \left( y_{i,t} \mid \cdot \right) = \sum_{j=0}^k b_j^q x_{i,t-1,j} \) is an estimate of \( QUANT_q \left( y_{i,t} \mid \cdot \right) \) and \( b^q = b_0^q, \ldots, b_k^q \) is an estimate of \( B^q = B_0^q, \ldots, B_k^q \).
value obtained from equation (1), \( \text{quant}_q(y_{i,t} | \cdot) = \sum_{j=0}^{k} b_j^q x_{i,t-1,j} \), is a consistent estimate of the \( q^{th} \) conditional quantile of \( y_{i,t} \), which we refer to as \( \text{QUANT}_q(y_{i,t} | \cdot) \).

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, \( b_j^q \) is a consistent estimate of \( B_j^q = \frac{\partial \text{QUANT}_q(y_{i,t} | \cdot)}{\partial x_{i,t-1,j}} \) (Buchinsky [1998]). Hence, the coefficients obtained from a quantile regression reflect marginal effects. However, unlike the coefficients obtained from an OLS regression, which equal the marginal effect of the \( x_{i,t-1,j} \) on the conditional mean of \( y_{i,t} \), coefficients obtained from a quantile regression equal the marginal effect of \( x_{i,t-1,j} \) on the \( q^{th} \) conditional quantile of \( y_{i,t} \).

3.2 Using the Quantile Function to Estimate Moments

Equation (1) relates to a single quantile. However, the moments of \( y_{i,t} \) depend on \textit{all} of the quantiles. For example, in equation (A.2) of Appendix A we show how to express the variance of \( y_{i,t} \) as a function of the quantiles of \( y_{i,t} \). Hence, in order to estimate the moments of \( y_{i,t} \), we need to first estimate the its quantile function, which is defined as the set of all the true quantiles of \( y_{i,t} \).

In light of the discussion in section 3.1, it is tempting to use the following naïve, three-step approach to estimate the moments of \( y_{i,t} \). First, estimate equation (1) for a large number of quantiles—i.e., use a large value of \( Q \). Second, convert the \( Q \) coefficient vectors obtained from step one into \( Q \) fitted values. Finally, use the estimated quantiles to calculate the moments of \( y_{i,t} \).

Unfortunately, there are two problems associated with the naïve approach. Following Basset and Koenker [1982], we refer to the first problem as the “quantile-crossing” problem. It refers to the fact that the estimated quantiles obtained from equation (1) are not monotonically
increasing in \( q \)—i.e., estimates of lower quantiles may exceed estimates of higher quantiles. The likelihood of crossing increases as the quantiles estimated become closer together. For example, the likelihood that \( \text{quant}_{0.25}(y_{i,t} | \cdot) \) will cross \( \text{quant}_{0.30}(y_{i,t} | \cdot) \) exceeds the likelihood that \( \text{quant}_{0.25}(y_{i,t} | \cdot) \) will cross \( \text{quant}_{0.75}(y_{i,t} | \cdot) \). Nonetheless, even for quantiles that are far apart, the likelihood that the estimated quantiles will cross is nontrivial.\(^6\)

The second problem with the naïve approach is that it generates biased and inconsistent estimates of the moments of \( y_{i,t} \). This may seem counterintuitive given that, as discussed in section 3.1, \( \text{quant}_q(y_{i,t} | \cdot) \) is a consistent estimate of \( \text{QUANT}_q(y_{i,t} | \cdot) \) for all \( q \). However, as both the number of quantiles (i.e., \( Q \)) estimated and the sample size (i.e., \( N \)) approaches infinity, the set \( \{\text{quant}_a(y_{i,t} | \cdot), \ldots, \text{quant}_q(y_{i,t} | \cdot), \ldots, \text{quant}_z(y_{i,t} | \cdot)\} \) does not converge to the true quantile function. The reason for this is that the estimated quantiles do not converge at a uniform rate—i.e., for any \( i \neq j \), \( \text{quant}_i(y_{i,t} | \cdot) \) will converge at a different rate than \( \text{quant}_j(y_{i,t} | \cdot) \).

To circumvent the two problems described above we use the method of rearranged quantiles described in Chernozhukov et al. [2010]. In particular, we begin by solving the minimization problem shown in equation (1) for a sequence of \( Q \) quantiles. Next, we rearrange these estimated quantiles in the manner described in Chernozhukov et al. [2010]. (We elaborate on the rearrangement in Appendix A.) This leads to an estimated quantile function that: (1) is increasing in \( q \); (2) exhibits less finite-sample bias than the estimate obtained from the naïve approach; and, (3) converges to the true quantile function of \( y_{i,t} \). Finally, we use the estimated

---

\(^6\) Some studies “avoid” the quantile-crossing problem by using heuristic estimates that are based on a small number of quantiles that are fairly far apart. For example, the estimated inter-quartile range might be used to estimate the variance of \( y_{i,t} \). There are two problems with this approach. First, even if the estimated quantiles do not cross, they are still biased. Hence, this approach simply hides the problem. Second, for any arbitrary distribution, these types of heuristic estimates do not equal the true moments. Hence, using them implies sacrificing an important advantage of the quantile regression approach: It does not require an assumption about the underlying distribution.
rearranged quantile function and the formulas described in Appendix A to arrive at consistent out-of-sample estimates of the variance, skewness and kurtosis of $y_{i,t-1}$.

### 3.3 Modeling the Higher Moments of Earnings

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 $t-1$. We refer to year $t$ as the *lead* year and we refer to year $t-1$ as the *current* year.$^7$

\[
QUANT_q \left( ROE_{i,t} | \cdot \right) = B_{0,EY}^q + B_{1,EY}^q \cdot ROE_{i,t-1} + B_{2,EY}^q \cdot LOSS_{i,t-1} + B_{3,EY}^q \left( ROE_{i,t-1} \times LOSS_{i,t-1} \right) + B_{4,EY}^q \cdot ACC_{i,t-1} + B_{5,EY}^q \cdot LEV_{i,t-1} + B_{6,EY}^q \cdot PAYER_{i,t-1} + B_{7,EY}^q \cdot PAYOUT_{i,t-1}
\]

(2)

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

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROE_{i,t}</td>
<td>Earnings of firm $i$ during year $t$ divided by firm $i$’s year $t$-1 equity book value</td>
</tr>
<tr>
<td>ROE_{i,t-1}</td>
<td>Earnings of firm $i$ during year $t$-1 divided by firm $i$’s year $t$-1 equity book value</td>
</tr>
<tr>
<td>LOSS_{i,t-1}</td>
<td>An indicator variable that equals one (zero) if $ROE_{i,t-1} &lt; 0$ ($ROE_{i,t-1} \geq 0$)</td>
</tr>
<tr>
<td>ACC_{i,t-1}</td>
<td>Accruals reported by firm $i$ during year $t$-1 divided by firm $i$’s year $t$-1 equity book value</td>
</tr>
<tr>
<td>LEV_{i,t-1}</td>
<td>Total assets of firm $i$ for year $t$-1 divided by firm $i$’s year $t$-1 equity book value</td>
</tr>
<tr>
<td>PAYER_{i,t-1}</td>
<td>An indicator variable that equals one (zero) if $PAYOUT_{i,t-1} &gt; 0$ ($PAYOUT_{i,t-1} = 0$)</td>
</tr>
<tr>
<td>PAYOUT_{i,t-1}</td>
<td>Dividends paid by firm $i$ during year $t$-1 divided by firm $i$’s year $t$-1 equity book value</td>
</tr>
</tbody>
</table>

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-1} 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-1} to vary with the sign of ROE_{i,t-1}. Third, evidence provided by Sloan [1996] implies that accruals are less persistent that cash flows. Consequently, we control for the portion of year t-1 ROE that is attributable to year t-1 accruals, ACC_{i,t-1}. Finally, well-known results in finance (e.g., Lintner [1956], Modigliani

---

$^7$ 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 $ROE_{i,t-1} \times LOSS_{i,t-1}$.
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 LEVi,t-1, PAYERi,t-1, and PAYOUTi,t-1 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, SALEi,t-1; an indicator for extreme ROE that equals one (zero) if ROEi,t-1 is (is not) in the top or bottom tenth percentile of the annual distribution, XTRM_ROEi,t-1; the interaction between XTRM_ROEi,t-1 and ROEi,t-1; the lagged change in earnings deflated by equity book value, ΔROEi,t-1; an indicator for extreme changes in ROE that equals one (zero) if ΔROEi,t-1 is (is not) in the top or bottom tenth percentile of the annual distribution, XTRM_ΔROEi,t-1; the interaction between the XTRM_ΔROEi,t-1 and ROEi,t-1; the interaction between XTRM_ΔROEi,t-1 and ΔROEi,t-1; the measure of unconditional conservatism described in Penman and Zhang [2002], CONSi,t-1; the interaction between CONSi,t-1 and ROEi,t-1; and, various combinations of the aforementioned variables. In a set of untabulated results we show that none of these models generate better out-of-sample predictions than the predictions that are based on 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 $B_{EY} = [B_{0,EY}, \ldots, B_{7,EY}]$ for 150 different values of
q ∈ (0,1). 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 never use more than ten years of data to construct a panel. For example, suppose the estimation year is 1990 (i.e., EY = 1990), we use values of the dependent variable that fall between 1981 and 1990 and we use values of the independent variables that fall between 1980 and 1989. 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., $b_{j,EY}^q$) and we compute the average of $b_{j,EY}^q$, which we refer to as $b_{j,AVG}^q$. Next, assuming a lag length of ten, we calculate the Newey-West adjusted standard error of $b_{j,AVG}^q$; and, we use the standard error to form a 95 percent confidence interval around $b_{j,AVG}^q$. We then graph $b_{j,AVG}^q$ and its confidence interval on q. We elaborate on how to interpret these graphs in section five.

Finally, we develop our out-of-sample predictions. 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$_{i,t+1}$ 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, per the discussion in section 3.2 and Appendix A, we

---

8 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 discussed in Appendix A, as the sample size increases and the number of quantile regressions increases, the estimates of the predicted moments converge to the moments of lead ROE.

9 We estimate three alternative sets of regressions in which we use a panel that contains: (1) only one year of data (i.e., a single cross-section); (2) no more than three years of data; and, (3) no more than five years of data. Untabulated results based on these regressions are similar to the results shown in the figures and tables.
rearrange the estimated quantiles. Finally, we use the rearranged quantiles and the equations described in Appendix A to calculate our estimates of the conditional moments.

4. Overview of Samples and Descriptive Statistics

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 moments of lead ROE.\(^{10}\)

In Panel A of Table 1 we provide descriptive statistics for the estimation sample. The mean (median) of \(\text{ROE}_{i,t}\) is 0.031 (0.101). Twenty-four percent of the sample observations have negative \(\text{ROE}_{i,t-1}\). The mean (median) of \(\text{ACC}_{i,t-1}\) is -0.059 (-0.054). \(\text{LEV}_{i,t-1}\) 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 1 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.

Several correlations warrant discussion. First, the Pearson (Spearman) correlation between lead ROE (i.e., \(\text{ROE}_{i,t}\)) and current ROE (i.e., \(\text{ROE}_{i,t-1}\)) 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 \(\text{ROE}_{i,t}\) and \(\text{LOSS}_{i,t-1}\) is -0.42

\(^{10}\) For additional details regarding our sample construction algorithm and variable definitions, please refer to Appendix B.
(-0.43). Third, the Pearson (Spearman) correlation between lead ROE and current accruals (i.e., $\text{ACC}_{i,t-1}$) is 0.11 (0.12), which implies that accruals are less persistent than cash flows. Current leverage (i.e., $\text{LEV}_{i,t-1}$) is uncorrelated with lead ROE$_{i,t}$. However, the Pearson (Spearman) correlation between lead ROE and $\text{PAYER}_{i,t-1}$ is 0.20 (0.23); and, the Pearson (Spearman) correlation between the current payout ratio (i.e., $\text{PAYOUT}_{i,t-1}$) 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 $\text{EY}$, value of $q$, and firm-level attribute $j$ we obtain the relevant coefficient estimate—i.e., $b_{j,EY}^q$. We then compute the average of $b_{j,EY}^q$ across estimation years, which we refer to as $b_{j,\text{AVG}}^q$, and the temporal standard error of $b_{j,\text{AVG}}^q$. 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 $b_{j,\text{AVG}}^q$ and its confidence interval on $q$. For comparative purposes, we also graph the average coefficient, which we refer to as $b_{j,\text{AVG}}^{\text{OLS}}$, and the 95 percent confidence interval obtained from an OLS regression.

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., $b_{0,EY}^q$—equals
the conditional q\textsuperscript{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 coefficient vector $b_1^q, \ldots, b_j^q$ obtained from estimating equation (2) on the de-medianed data is identical to the coefficient vector $b_1^q, \ldots, b_j^q$ 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.

Before discussing our results it is useful to provide an overview of how to interpret the graphs of the slope coefficients. To do this we discuss three hypothetical examples of graphs of $Q$ estimated slope coefficients on q. As discussed in section 3.1, $b_j^q$ is a consistent estimate of the marginal effect of $x_{i,t-1,j}$ on the qth conditional quantile of $y_{i,t}$. First, suppose that $b_j^q = c$ for all q. This implies that as $x_{i,t-1,j}$ increases all of the quantiles shift to the right by an equal amount. Hence, $x_{i,t-1,j}$ is associated with the location, but not the shape, of the distribution of $y_{i,t}$. Second, suppose that $b_j^q > 0$ for all q and it an increasing function of q (i.e., $\Delta b_j^q / \Delta q > 0$). Hence, increases in $x_{i,t-1,j}$ lead to increases in all of the quantiles of $y_{i,t}$ and the upper quantiles increase by a larger amount. This implies that $x_{i,t-1,j}$ is positively associated with both the conditional mean and conditional variance of $y_{i,t}$—i.e., as $x_{i,t-1,j}$ increases the conditional distribution of $y_{i,t}$ shifts to the right and becomes more disperse.

Finally, suppose that for all q $b_j^q > 0$, $\Delta b_j^q / \Delta q > 0$, and $\Delta \Delta b_j^q / \Delta q < 0$—i.e., $b_j^q$ is always positive and it is an increasing, concave function of q. As discussed above, this implies that $x_{i,t-1,j}$ is positively associated with the conditional mean and variance of $y_{i,t}$. It also implies
that $x_{i,t-1,j}$ is negatively associated with the conditional skewness of $y_{i,t}$. The reason for this is that as $q$ approaches zero, the rate of change in the dispersion in $y_{i,t}$ increases. Hence, higher values of $x_{i,t-1,j}$ are associated with distributions with extremely long left tails.

We show the graph of the constant term, $b^{q}_{0, AVG}$, in Figure 1. As discussed above, $b^{q}_{0, AVG}$ ($b^{OLS}_{0, AVG}$) equals the conditional $q^{th}$ quantile (conditional mean) of lead ROE for the “typical” observation. As shown in Figure 1, 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 2 contains the graph of $b^{q}_{1, AVG}$, which is the coefficient on current ROE (i.e., ROE$_{i,t-1}$). As shown in Figure 2, $b^{0.50}_{1, AVG}$ ($b^{OLS}_{1, AVG}$) equals 1.01 (0.85); hence, there is a positive association between current ROE and the median (mean) of lead ROE. In addition, $b^{q}_{1, AVG}$ is an increasing function of $q$ (i.e., $\Delta b^{q}_{1, AVG}/\Delta q > 0$), which implies there is a positive association between current ROE and the variance of lead ROE. Finally, for values of $q < 0.80$ the relation between $b^{q}_{1, AVG}$ and $q$ is concave (i.e., $\Delta \Delta b^{q}_{1, AVG}/\Delta q < 0$) but for values of $q > 0.80$ the relation between $b^{q}_{1, AVG}$ and $q$ is convex (i.e., $\Delta \Delta b^{q}_{1, AVG}/\Delta q > 0$). Hence, firms with higher current ROE are more likely to have extreme values of lead ROE. That is, these firms have ROE that is more leptokurtic (i.e., fat-tailed).

In light of the above, we conclude that firms 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 3 contains the graph of $b_{2,AVG}^q$, 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 and more volatile lead ROE. In particular, $b_{2,AVG}^{0.50}$ ($b_{2,AVG}^{OLS}$) equals -0.01 (-0.07) and $b_{2,AVG}^q$ is increasing in q. Loss firms are also more likely to experience extreme poor performance. Specifically, the relation between $b_{2,AVG}^q$ and q is a concave for most values of q. Hence, loss firms have lead ROE that is more left-skewed.

Figure 4 contains the graph of $-1 \times (b_{1,AVG}^q + b_{3,AVG}^q)$. 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., $b_{1,AVG}^q$) and the coefficient on the interaction of ROE$_{i,t-1}$ and LOSS$_{i,t-1}$ (i.e., $b_{3,AVG}^q$). 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, $-1 \times (b_{1,AVG}^{OLS} + b_{3,AVG}^{OLS})$ equals -0.52 (-0.44). In addition, $-1 \times (b_{1,AVG}^q + b_{3,AVG}^q)$ is increasing in q, which implies the magnitude of the current loss is positively associated with the variance of lead ROE.

In Figure 5 we graph the relation between the coefficient on current accruals, $b_{4,AVG}^q$, and q. The results shown on the graph suggest that higher current accruals are associated with lower and riskier lead ROE. In particular, $b_{4,AVG}^{0.50}$ ($b_{4,AVG}^{OLS}$) equals -0.03 (-0.06) and $b_{4,AVG}^q$ is increasing in q.

---

11 The confidence intervals shown in Figure 4 relate to the standard error of the average of $-1 \times (b_{1,LY}^q + b_{3,LY}^q)$ and the average of $-1 \times (b_{1,LY}^{OLS} + b_{3,LY}^{OLS})$. That is, we use the standard error of the average of the sum not the sum of the standard errors of the averages.
Figure 6 contains the graph of $b_{5,AVG}^q$, 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. $b_{5,AVG}^q$ 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 $b_{5,AVG}^q$ 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. 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 7 we show the graph of $b_{6,AVG}^q$, which is the coefficient on the dividend indicator (i.e., PAYER$_{i,t-1}$). First, $b_{6,AVG}^{OLS}$ equals 0.04 and $b_{6,AVG}^{0.50}$ equals 0.01. Hence, dividend-paying firms tend to have higher lead ROE; however, the effect primarily relates to the mean. Second, $b_{6,AVG}^q$ 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 $b_{6,AVG}^q$ 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 exhibits higher positive skewness.

Figure 8 contains the graph of $b_{7,AVG}^q$, 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, $b_{7,AVG}^{0.50}$ ($b_{7,AVG}^{OLS}$) equals -0.03 (0.09). Second, for values of q between 0.10 and 0.90, $b_{7,AVG}^q$ is a decreasing function of q. However, for values of q $\in \{(0,0.10)\cup(0.90,1.00)\}$, $b_{7,AVG}^q$ 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 9 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.\(^{12}\) (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 Predictions**

In this section we evaluate the out-of-sample predictions of the moments of lead ROE. All of the analyses described in this section are based on observations drawn from the prediction sample. It is important to note that all of the predictions described in this section are out-of-sample. In particular, we develop a year t prediction of a variable in year t+1 by combining

\(^{12}\) The pseudo r-squared shown in Figure 9 is the standard statistic reported by Stata.
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.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_MEAN_{i,t+1}</td>
<td>Year t estimate of the mean of ROE_{i,t+1}</td>
</tr>
<tr>
<td>Q_STD_{i,t+1}</td>
<td>Year t estimate of the standard deviation of ROE_{i,t+1}</td>
</tr>
<tr>
<td>Q_SKEW_{i,t+1}</td>
<td>Year t estimate of the skewness of ROE_{i,t+1}</td>
</tr>
<tr>
<td>Q_KURT_{i,t+1}</td>
<td>Year t estimate of the excess kurtosis of ROE_{i,t+1}</td>
</tr>
</tbody>
</table>

As discussed in section three, the variables shown above are inferred from our out-of-sample estimates of the rearranged quantiles. Specifically, for firm i in year t we obtain and rearrange the predicted quantiles for all 150 values of q. Next, we calculate the mean, standard deviation, skewness, and kurtosis for this “sample” of 150 values.\(^{13}\)

Panel A of Table 2 contains descriptive statistics. Several comments are warranted. First, the average (typical) firm has positive Q_MEAN_{i,t+1}; in particular, the mean (median) of Q_MEAN_{i,t+1} is 0.042 (0.097). However, Q_MEAN_{i,t+1} varies considerably across observations. For example, the standard deviation (interquartile range) of Q_MEAN_{i,t+1} is 0.237 (0.210). Moreover, untabulated results show that 27.8 percent of the observations have negative Q_MEAN_{i,t+1}.

\(^{13}\) We define skewness as the ratio of the third central moment to the third power of Q_STD_{i,t+1}; and, we calculate kurtosis by subtracting three from the ratio of the fourth central moment to the fourth power of Q_STD_{i,t+1}. 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 the distribution of lead ROE differs from the normal distribution.
Second, $Q_{\text{STD}}_{i,t,t+1}$ is large, which implies there is considerable uncertainty about future ROE. To understand the magnitude of $Q_{\text{STD}}_{i,t,t+1}$ better we evaluate the coefficient of variation, $Q_{\text{CV}}_{i,t,t+1}$, which equals the ratio of $Q_{\text{STD}}_{i,t,t+1}$ to $|Q_{\text{MEAN}}_{i,t,t+1}|$. The mean (median) of $Q_{\text{CV}}_{i,t,t+1}$ 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 $Q_{\text{CV}}_{i,t,t+1}$ is 0.975.

Finally, the median of $Q_{\text{SKEW}}_{i,t,t+1}$ ($Q_{\text{KURT}}_{i,t,t+1}$) 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 2 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.

We note that the Pearson correlations between $Q_{\text{MEAN}}_{i,t,t+1}$ and $Q_{\text{STD}}_{i,t,t+1}$, $Q_{\text{SKEW}}_{i,t,t+1}$, and $Q_{\text{KURT}}_{i,t,t+1}$ are -0.62, 0.20, and 0.27, respectively. Hence, firms with high mean lead ROE also tend to have less volatile, more positively skewed and more extreme ROE.
We also note that the Pearson correlation between $Q_{\text{STD},i,t,t+1}$ and $Q_{\text{SKEW},i,t,t+1}$ is -0.26; and, the correlation between $Q_{\text{STD},i,t,t+1}$ and $Q_{\text{KURT},i,t,t+1}$ is -0.38. Taken together, these two facts imply that as the variance of lead ROE increases, the distribution of lead ROE becomes more (less) negatively (positively) skewed. Finally, the Pearson correlation between $Q_{\text{SKEW},i,t,t+1}$ and $Q_{\text{KURT},i,t,t+1}$ 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 Out-of-sample Predictions

In this section we evaluate the out-of-sample predictions of 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. We use industry-level tests for two reasons. First, the realized moments of ROE are not observable at the firm-year level. Hence, tests of reliability cannot be conducted using firm-level data. However, the realized cross-sectional distribution is observable at the industry-year level. Consequently, results of industry-level tests provide evidence about the reliability of our firm-level estimates. Second, 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.

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 $t+1$. Per the discussion in footnote 13, 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 the space below 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 describe them in Appendix A.

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

\[
\sqrt{VAR(ROE_{i,t+1})} = \sqrt{VAR(E(ROE_{i,t+1} | i)) + E(VAR(ROE_{i,t+1} | i))}
\]

(3)

In the above equation, \(VAR(\cdot)\) denotes the variance and \(E(\cdot)\) is the expected value.

With equation (3) in mind, we do the following. First, for each firm \(i\) and year \(t\) we obtain estimates of the mean and the variance of \(ROE_{i,t+1}\). 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 \(ROE_{i,t+1}\). We do this by taking the square root of the sum of: (1) the within-industry variance of the predictions of the mean of \(ROE_{i,t+1}\) and (2) the industry mean of the predictions of the variance of \(ROE_{i,t+1}\). Finally, we calculate the realized cross-sectional standard deviation of ROE in year \(t+1\). We refer to the quantile-based year \(t\) prediction of the standard deviation of industry IND’s ROE in year \(t+1\) as \(Q_{STD_{IND,t+1}}\); and, we refer to the realized standard deviation as \(R_{STD_{IND,t+1}}\). We use similar notation for predicted and realized values of skewness and kurtosis.\(^{14}\)

In Tables 3, 4 and 5 we show the results of regressing realized moments on predicted moments. Each table contains 12 columns. In columns (1) through (3), we compare our quantile-

\(^{14}\) 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 \(t+1\).
based predictions to predictions obtained from the historical firm-year approach whereas in columns (4) through (6) we compare our prediction to predictions obtained from the historical matched-sample approach. As discussed above, both of these approaches as well as our quantile-based approach rely on the law of total moments.

We also compare our quantile-based predictions to two approaches that do not rely on the law of total moments. Results in columns (7) through (9) relate to predictions obtained by setting the year t prediction of a particular moment in year t+1 equal to the industry-level moment of analysts’ forecasts made in year t. Results in columns (10) through (12) relate to predictions obtained from the historical industry-level approach, which involves setting the year t prediction of a particular moment in year t+1 equal to the industry-level moment of realized ROE in year t.

As shown in in columns (1), (4) and (10) of Table 3 \( Q_{\text{STD}_{\text{IND},t,t+1}} \) has a significant positive association with \( R_{\text{STD}_{\text{IND},t+1}} \) and it explains roughly 30 percent of the cross-sectional variation in \( R_{\text{STD}_{\text{IND},t+1}} \). Hence, it is a reliable predictor on an absolute basis. Per columns (3), (6), (9) and (12), \( Q_{\text{STD}_{\text{IND},t+1}} \) is incrementally informative vis-à-vis each of the alternative predictors. Moreover, Vuong-test results show that the \( Q_{\text{STD}_{\text{IND},t+1}} \) is a better predictor than three of the estimates and that is has the same predictive power as the estimate obtained from analysts’ forecasts.

\[ \text{Untabulated results using these alternative predictions are similar to the results shown in Tables 3 through 5.} \]

\[ \text{We also considered setting the year t prediction of a particular moment in year t+1 equal to the industry-level moment of realized ROE (regression residuals) for years t-4 through t as well as years t-9 through t. Untabulated results using these alternative proxies are similar to the results shown in Tables 3 through 5.} \]
Regarding Table 3, the comparison between the quantile-based predictions and the predictions that are based on analysts’ forecasts warrants additional discussion. As shown in column (7), for this subsample, the quantile-based predictions do not have a positive association with R_STD\textsubscript{IND,t+1}. Moreover, per the Vuong test, they are not better than the predictions obtained from analysts’ forecasts. To understand these results better we evaluate each annual regression. We find that the coefficient on Q_STD\textsubscript{IND,t+1} for the year 2000 is 107.27 and that this is attributable to one observation that has a DF beta of 449.\textsuperscript{17} We remove this observation, re-estimate the regression for the year 2000, and combine the estimates from that regression with the original estimates from the other annual regressions. Untabulated results show that, after making this adjustment, the average of the annual coefficients on Q_STD\textsubscript{IND,t+1} is 1.64 (t-statistic of 11.55) and the Vuong statistic is 3.42, which implies that the quantile-based predictions are reliable on an absolute basis and are better than the predictions obtained from analysts’ forecasts.

Results related to Q_SKEW\textsubscript{IND,t+1} are similar to those for Q_STD\textsubscript{IND,t+1}. As shown in columns (1), (4), (7) and (10) of Table 4, Q_SKEW\textsubscript{IND,t+1} has a significant positive association with R_SKEW\textsubscript{IND,t+1} and it explains approximately 14 percent of the cross-sectional variation in R_SKEW\textsubscript{IND,t+1}. Per columns (3), (6), (9) and (12), these results remain after controlling for the predictions obtained from the alternative approaches. Moreover, Vuong-test results show that Q_SKEW\textsubscript{IND,t+1} is a better predictor than three of the alternative estimates and that it as good as the predictor based on analysts’ forecasts.

\textsuperscript{17} The observation relates to industries that have a two-digit SIC code of 39: 3910 Jewelry, Silverware & Plated Ware; 3911 Jewelry, Precious Metal; 3931 Musical Instruments; 3942 Dolls & Stuffed Toys; 3944 Games, Toys & Children's Vehicles (No Dolls & Bicycles); 3949 Sporting & Athletic Goods, NEC; 3950 Pens, Pencils & Other Artists' Materials; and, 3960 Costume Jewelry & Novelties.
Finally, as shown in columns (1), (4), (7) and (10) of Table 5, $Q_{KURT}^{IND,t,t+1}$ is positively associated with $R_{KURT}^{IND,t,t+1}$ and it explains more than 14 percent of the variation in $R_{KURT}^{IND,t,t+1}$. Moreover, as shown in columns (3), (6), (9) and (12) these results remain after controlling for the predictions obtained from the each of the alternative approaches. Finally, show $Q_{KURT}^{IND,t,t+1}$ is as good of a predictor as each of the alternative predictors.

The results described above imply that our methodology generates reliable _ex ante_ estimates of the higher moments of future ROE. This fact is important for three reasons. First, given they relate to infrequent events, higher moments such as skewness and kurtosis are difficult to predict. Second, our prediction sample relates to 170,522 firm-years. This is large. In contrast, as discussed in Appendix B, there are only 134,522, 85,142 and 36,462 firm-years with sufficient data to compute moments per the historical-firm-level, historical-matched-sample, and analysts’-forecasts approach, respectively.\(^{18}\) Hence, in addition to generating predictions that are typically superior to the predictions generated by the other approaches, the quantile-based approach generates many more predictions.

Finally, as discussed in section two, the higher moments of future earnings 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 each of these different contexts, we do provide evidence on the pricing of equity and debt securities in the next section.

7. **Analyses of Equity- and Credit-market Variables**

7.1 **Analyses of Equity-market Variables**

We analyze three equity-market variables: (1) the implied cost of capital, $ICC_{i,t}$; (2) the ratio of forward-earnings to price, $EP_{i,t}$; and, (3) the ratio of equity book value to equity market

\(^{18}\) The historical industry-level approach does not generate firm-level predictions.
value, BP_{i,t}. Results in Hou et al. [2012] and Penman and Reggiani [2013] show that these variables are related to risk.¹⁹

Table 6 contains results obtained from regressions of equity-market-based variables on Q_{MEAN,i,t+1}, Q_{STD,i,t+1}, Q_{SKEW,i,t+1} and Q_{KURT,i,t+1}. The results in the first pair of columns relates to regressions in which ICC_{i,t} is the dependent variable. ICC_{i,t} is based on the model described in Gebhardt et al. [2001]. However, following Hou et al. [2012], our forecasts of ROE for years t+1 through t+3 are obtained from the OLS version of equation (2).²⁰ The second (third) pair of columns relates to regressions in which EP_{i,t} (book-to-price, BP_{i,t}) is the dependent variable. EP_{i,t} (BP_{i,t}) equals the ratio of firm i’s expected year t+1 earnings (equity book value at the end of year t) to its equity market value at the end year t. Columns (1), (3) and (5) relate to results in which the out-of-sample estimates of the moments of lead ROE are the only independent variables. Columns (2), (4) and (6) relate to results in which we include six control variables: (1) equity market value at the end of year t, SIZE_{i,t}; (2) equity beta for year t, BETA_{i,t}; (3) annual stock return for year t, RET_{i,t}; (4) the volatility of market-model residuals, RET_STD_{i,t}; (5) the skewness of market-model residuals, RET_SKEW_{i,t}; and, (6) the kurtosis of market-model residuals, RET_KURT_{i,t}.²¹ We remove observations for which the value of any of

---

¹⁹ Hou et al. [2012] provide a detailed discussion of the implied cost of equity capital whereas Penman and Reggiani [2013] focus on EP_{i,t} and BP_{i,t}. Penman and Reggiani [2013] show that, if there is no growth in expected residual income, EP_{i,t} is a comprehensive measure of risk—i.e., a higher EP_{i,t} implies higher risk. However, when there is growth in expected residual income, BP_{i,t} is also relevant because it captures the risk of the growth—i.e., higher values of BP_{i,t} imply higher risk. Similar results can be found in Penman and Zhu [2014].

²⁰ Please refer to p. 524 of Hou et al. [2012] for further details regarding the calculation of ICC_{i,t}.

²¹ 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]. BETA_{i,t} is estimated in three steps. First, for each industry based on Fama-French 48 classifications, we estimate the slope coefficient obtained from an OLS weighted regressions of industry-level monthly returns on the contemporaneous return on the market portfolio. The weights equal the contemporaneous equity market values. We use monthly returns drawn from a 60-month period ending three months after the last month of fiscal-year t. Second,
the variables in the regression falls in either the top or bottom percentile of its annual distribution.

The results in Table 6 show that firms with higher standard deviation of lead ROE or more positively skewed lead ROE have lower values of ICC_{i,t}, EP_{i,t} and BP_{i,t}. 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, ICC_{i,t}, EP_{i,t} and BP_{i,t} are each increasing in *ex ante* firm-level kurtosis. This result is consistent with results shown in Dittmar [2002].

7.2 *Analyses of Credit-market Variables*

We analyze three credit-market variables: (1) credit-default-swap, CDS, spreads in year t; (2) bond yields in year t; and, (3) credit ratings in year t. We evaluate the relation between each of these variables and year t predictions of the moments of lead ROA. To estimate the moments of lead ROA we make three modifications to the equation (2). First, we replace ROE with return on assets, ROA. We compute ROA_{i,t} (ROA_{i,t-1}) as the ratio of the sum of year t (t-1) earnings and net interest expense to total assets in year t-1. Second, we omit the variable LEV_{i,t-1} from the model to avoid mechanical associations between the predicted moments and the credit-market variables. Finally, with the exception of the indicator variables, we deflate the remaining independent variables by total assets in year t-1 not equity book value in year t-1.\(^{22}\) We refer to the year t prediction of the mean, standard deviation, skewness and kurtosis of firm i’s ROA in

---

\(^{22}\) Graphs of the coefficients and r-squareds are similar to the corresponding graphs shown in Figures 1 through 9. Using the approach described in section six we show that there is a positive association between the year t predicted moments of ROA and the year t+1 realized moments of ROA.
year $t+1$ as $\text{MEAN\_ROA}_{i,t,t+1}$, $\text{STD\_ROA}_{i,t,t+1}$, $\text{SKEW\_ROA}_{i,t,t+1}$ and $\text{KURT\_ROA}_{i,t,t+1}$, respectively.

We retrieve CDS spreads from the MarkIt Database and use the spreads on the 5-year CDS contracts with modified restructuring clauses. We obtain bond yields from a combined dataset which uses secondary market bond trades from both Trace or Mergent FISD. Finally, we retrieve credit ratings on senior debt from Standard and Poors and code them numerically from 1 (lowest credit risk) to 24 (highest credit risk). All three credit market variables are measured four months after the end of fiscal year $t$. Higher CDS spreads, bond yields and credit ratings correspond to higher credit risk.

Table 7 contains results obtained from regressions of the credit-market variables on $\text{MEAN\_ROA}_{i,t,t+1}$, $\text{STD\_ROA}_{i,t,t+1}$, $\text{SKEW\_ROA}_{i,t,t+1}$ and $\text{KURT\_ROA}_{i,t,t+1}$. The results in the first pair of columns relates to regressions in which the CDS spread is the dependent variable. The second pair of columns relates to regressions in which the bond yield is the dependent variable whereas the final pair of columns shows regressions in which the bond ratings is the dependent variable. Columns (1), (3) and (5) relate to results in which the out-of-sample estimates of the moments of lead ROA are the only independent variables. Columns (2), (4) and (6) relate to results in which we include control variables. Eight of these control variables are common to all three regression specifications: (1) $\text{BP}_{i,t}$; (2) 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, $\text{LN\_SIZE}_{i,t}$; (3) the ratio of firm $i$’s year $t$ liabilities to its year $t$ assets, $\text{LIAB\_ASST}_{i,t}$; (4) the ratio of firm $i$’s year $t$ earnings before interest, taxes, depreciation, and amortization to its year $t$ liabilities, $\text{EBITDA\_LIAB}_{i,t}$; (5) $\text{RET}_{i,t}$; (6) $\text{RET\_VOL}_{i,t}$; (7) $\text{RET\_SKEW}_{i,t}$; and, (8) $\text{RET\_KURT}_{i,t}$. When evaluating bond yields we also control for the remaining number of years to the date on which

---

23 For further details about the MarkIt data please refer to Shivakumar et al. [2011].
the bond matures, $T_{2\text{MAT},it}$, and the log of the aggregate par value of the bond at the date of issuance, $B_{\text{OND\_SIZE},it}$. The control variables are based on the default prediction model described in Beaver et al. [2012]. We also follow Beaver et al. [2012] and winzorize each variable to its first and 99\textsuperscript{th} percentiles. Each regression includes industry fixed effects based on the Fama-French 12 classifications.

The results in Table 7 show that firms with higher \textit{ex ante} standard deviation of ROA have higher CDS spreads, bond yields and credit ratings, consistent with the analytical results in Merton [1974]. However, consistent with practitioner articles (e.g, Dynkin et al. [2007]) skewness and kurtosis are also relevant. In particular, credit prices reflect higher risk premiums when the \textit{ex ante} skewness (kurtosis) of ROA is more negative (positive). This is not surprising given that debt holders 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 predictions of the higher moments of future earnings; and, we show that these predictions are related to equity- and credit-market variables. These two contributions are nontrivial for several 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, equity and credit markets play a central role in the intermediation of funds from investors to firms and in allocating those funds across firms. We shed additional light on the role that earnings and earnings attributes play in this process.

\footnote{Our results are not sensitive to the choice of controls. In untabulated results 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].}
Finally, the fact that our approach is rigorous and that it generates reliable predictions implies that it can be used in future research projects. This is important because our equity- and credit-market results suggest that the higher moments of earnings are potentially relevant in many economic contexts. The reason for this is that many economic agents are, unlike capital market participants, unable to diversify away their exposure to idiosyncratic shocks. Hence, when exposed to high downside risk (or upside potential), these agents are especially likely to alter either their behavior or the institutional arrangements in which they participate.

In addition to further evaluating the economic relevance of the higher moments of future earnings, there are two other 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 we discuss above, our model is superior to a number of more elaborate models. Nonetheless, the firm-level attributes we evaluate are not exhaustive. Hence, future studies might focus on identifying other attributes that are related to the moments of future earnings both in- and out-of-sample.
Appendix A – Relationship between Quantiles and Higher Moments

In this appendix we describe the relation between the moments of a random variable \( y_{i,t} \) and the quantile function of \( y_{i,t} \). Next, we show that, using the rearranged quantile function developed by Chernozhukov et al. [2010], these formulas yield consistent estimates. Finally, we show how to use the law of total moments to convert firm-level moments into predicted industry-level skewness and kurtosis.

A.1 Formulas Relating the Quantile Function and the Higher Moments

The relation between the variance of \( y_{i,t} \), \( \text{VAR}(y_{i,t} \mid \cdot) \), the cumulative distribution function of \( y_{i,t} \), \( F(y_{i,t} \mid \cdot) \), and the expected value of \( y_{i,t} \), \( E(y_{i,t} \mid \cdot) \), is shown in equation (A.1).

\[
\text{VAR}(y_{i,t} \mid \cdot) = \int_{-\infty}^{+\infty} y_{i,t}^2 dF(y_{i,t} \mid \cdot) - (E(y_{i,t} \mid \cdot))^2
\]  

(A.1)

(A.1) be re-expressed as a functions of the true quantiles of \( y_{i,t} \), \( \text{QUANT}_q(y_{i,t} \mid \cdot) \).

\[
\text{VAR}(y_{i,t} \mid \cdot) = \int_0^1 \text{QUANT}_q(y_{i,t} \mid \cdot) dq - (E(y_{i,t} \mid \cdot))^2
\]  

(A.2)

An estimate of the variance of \( y_{i,t} \), \( \text{var}(y_{i,t} \mid \cdot) \), is obtained by taking the average of the squared values of the Q estimated quantiles, \( \text{quant}_q(y_{i,t} \mid \cdot) \), and then subtracting the square of the estimate of the expected value, \( e(y_{i,t} \mid \cdot) \).

\[
\text{var}(y_{i,t} \mid \cdot) = \frac{1}{Q} \sum_{q \in \mathbb{Z}} (\text{quant}_q(y_{i,t} \mid \cdot))^2 - (e(y_{i,t} \mid \cdot))^2
\]  

(A.3)

Recall from section three that \( Z = \{a, \ldots, q, \ldots, z\} \subset (0,1) \) is an ordered sequence of Q numbers. \( e(y_{i,t} \mid \cdot) \) is the average of the Q estimated quantiles. Equations that are similar to (A.3) can be used to obtain estimates of the centered third and fourth moments of \( y_{i,t} \), which, via standard formulas, can be converted into estimates of the skewness and excess kurtosis of \( y_{i,t} \).
A.2 Consistency of the Predicted Moments

Under the regularity conditions described in Koenker and Bassett [1978], the vector of estimates $b^q$ is a $\sqrt{N}$ consistent and asymptotically normal estimator of the true coefficients $B^q$. Formally, $\sqrt{N}(b^q - B^q) \rightarrow \text{NORMAL}(0, V)$. The $k \times k$ variance-covariance matrix, $V$, is equal to:

$$V = \left[ E(x_{i,t-1}', x_{i,t-1}) \right]^{-1}$$

$u(q) = y_{i,t} - x_{i,t-1} B^q$, and $f_{u(q)}(0)$ is the density of $u(q)$ at $u(q) = 0$.

The estimate of the $q^{th}$ quantile of $y_{i,t}$, $\text{quant}_q(y_{i,t} | \cdot)$, is then a consistent and asymptotically normal estimator of $\text{QUANT}_q(y_{i,t} | \cdot)$ for each quantile $q$.

As described in section 3.2, in general a sequence of $q$ estimated quantiles is not increasing in $q$—i.e., the “quantile-crossing” problem. Following Chernozhukov et al. [2010], we rearrange the quantile estimates so that: (1) the rearranged quantiles, $\text{quant}_q^*(y_{i,t} | \cdot)$, are increasing in $q$ and (2) they converge at a uniform rate. These two facts imply that the moments based on the rearranged quantiles have less finite-sample bias and are consistent.

An illustration of how the rearrangement works is shown in Figure A.1, which contains graphs of the estimated quantile function without rearrangement (dashed line) and the rearranged quantile function (solid line). These graphs pertain to one of our sample observations: Agnico Eagle Mines Ltd. in 2003. The non-rearranged quantile function exhibits strong non-monotonicity or “crossing.” In particular, for values of $q$ between zero and the 0.10, the non-rearranged function is decreasing in $q$. The function then begins to steadily increase; however, once $q$ becomes greater than 0.90, the function begins to rapidly decrease. In fact, the estimated quantiles between the 65th percentile and the 90th percentile are higher than the estimate of the
95\textsuperscript{th} percentile. Hence, the dashed line is not a uniformly consistent estimator of the true quantile function and any predicted moment based on it will exhibit substantial bias and be inconsistent.

For our sample the lack of monotonicity of the non-rearranged quantile function is not limited to Agnico Eagle Mines Ltd in 2003. In particular, we evaluate the number of consecutive “crossed” quantile pairs. That is pairs \{q,q’\}, q’ > q for which q’ immediately follows q in the sequence of estimates and the non-rearranged quantile function is lower at q’ than at q. For the average (median) observation in our sample 21 (three) percent of the non-rearranged quantiles are non-increasing. 69 percent of the observations have at least one pair of consecutive non-rearranged quantiles that “cross,” and, for 28 percent of the observations, a third of the non-rearranged quantiles are non-increasing. This implies that rearranging is beneficial but it is not so pervasive as to call into question the overall estimation approach.

Figure A.1 also shows the rearranged quantile function (solid line) for the company Agnico Eagle Mines Ltd. By design the rearranged quantile function is increasing in q. Furthermore, Chernozhukov et al. [2010] show that as Q (i.e., the number of estimated quantiles) and N (i.e., the number of observations) go to infinity, the sequence of Q rearranged estimates, 
\[
\left\{\text{quant}_q^*(y_{ij} | \cdot), \ldots, \text{quant}_q^*(y_{ij} | \cdot), \ldots, \text{quant}_q^*(y_{ij} | \cdot)\right\},
\]
converges uniformly to the true quantile function (Corollary 3 of Proposition 5 of Chernozhukov et al. [2010]). That is, the rearranged quantiles converge at a uniform rate. Hence, estimates of the moments that are based on the rearranged quantiles are consistent estimates of the moments of lead earnings.\(^{25}\)

\(^{25}\) The proof of this result is available upon request. In addition, the Stata code we use to estimate the in-sample relations, rearranged quantile functions, and out-of-sample predictions is also available upon request.
A.3 Predicting Within-industry-year Skewness and Kurtosis

As shown in Klugman et al. [1998], the law of total moments implies that the within-group third central moment, $M^3(\text{ROE}_{i,t+1})$, and fourth central moment, $M^4(\text{ROE}_{i,t+1})$, of lead ROE can be expressed as shown in equations (A.4) and (A.5), respectively.

$$M^3(\text{ROE}_{i,t+1}) = E[M^3(\text{ROE}_{i,t+1} \mid \cdot)] + 3 \times \text{COV}(E(\text{ROE}_{i,t+1} \mid \cdot), \text{VAR}(\text{ROE}_{i,t+1} \mid \cdot)) + M^3(E(\text{ROE}_{i,t+1} \mid \cdot))$$

$$M^4(\text{ROE}_{i,t+1}) = E[M^4(\text{ROE}_{i,t+1} \mid \cdot)] + 4 \times \text{COV}(E[M^3(\text{ROE}_{i,t+1} \mid \cdot), E(\text{ROE}_{i,t+1} \mid \cdot))] + 6 \times \text{COSKEW}(\text{VAR}(\text{ROE}_{i,t+1} \mid \cdot), E(\text{ROE}_{i,t+1} \mid \cdot), E(\text{ROE}_{i,t+1} \mid \cdot)) + 6 \times E[\text{VAR}(\text{ROE}_{i,t+1} \mid \cdot)] \times \text{VAR}(E(\text{ROE}_{i,t+1} \mid \cdot)) + M^4(E(\text{ROE}_{i,t+1} \mid \cdot))$$

In the above equations, $\text{COV}(\cdot; \cdot)$ is the covariance and $\text{COSKEW}(\cdot; \cdot; \cdot)$ is the coskewness.

With the above equations in mind, we do the following. First, for each firm $i$, year $t$ 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 $\text{ROE}_{i,t+1}$. Second, for a particular industry and year $t$ we identify all firms with non-missing values of the aforementioned estimates. Third, we use the estimates and the equations to predict the within-industry third and fourth moments of $\text{ROE}_{i,t+1}$. For example, when computing the third central moment, we sum: (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 (kurtosis) by dividing the aforementioned sum by the third (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 market-based variables.

B.1 Construction of Estimation and Prediction Samples

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 moments 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 first delete observations that have either missing values of the variables shown in equation (2) or negative equity book value in year t-1 (i.e., CEQt-1 ≤ 0). Next, we delete extreme observations, which we define as observations for which: |ROEt| > 2, |ROEt-1| > 2, |ACCt-1| > 2, LEVt-1 ∉ [1,20], and PAYOUTt-1 ∉ [0,1]. The estimation sample contains 174,215 firm-years with independent (dependent) variables drawn from the time-period spanning 1963 to 2010 (1964 to 2011).

In Figure B1 we graph the percentage of observations that are deleted after applying each of the separate criteria listed above—e.g., CEQt-1 ≤ 0, |ROEt| > 2, etc. We also graph the
percentage of observations that are deleted because one or more of the criteria apply. We refer to this group as “All.” It is important to note that the criteria are not mutually exclusive. For example, if $\text{CEQ}_{i,t} \leq 0$, $\text{LEV}_{i,t-1} \not\in [1,20]$. The graphs shown in Figure B1 lead to two conclusions. First, the percentage of observations deleted steadily increases during the sample period. This is consistent with results described in Fama and French [2004], Dichev and Tang [2008], and Srivastava [2014]. Second, this trend is primarily attributable the removal of observations with leverage that is either less than one or greater than 20.

To form our prediction sample we identify all firm-years with positive equity book value in year $t$ and non-missing values of $\text{ROE}_{i,t} = \text{IB}_{i,t} + \text{CEQ}_{i,t}, \text{LOSS}_{i,t}, \text{ROE}_{i,t} \times \text{LOSS}_{i,t}, \text{ACC}_{i,t}, \text{LEV}_{i,t}, \text{PAYER}_{i,t}$, and $\text{PAYOUT}_{i,t}$. We do not remove observations that have either extreme or 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 3 through 7 is lower. Finally, 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-based variables 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 four alternatives: (1) historical firm-level estimates; (2) historical matched-sample estimates; (3) analyst-based estimates; and, (4) historical industry-level estimates.

---

26 The results shown in Figures 1 through 9 and Tables 1 through 7 are not attributable to any specific time-period. In particular, we replicate our tests for separate ten-year time-periods—i.e., 1973-1982, etc.; and, we find that the untabulated results of these replications are similar to the results shown in the paper.
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. There are 134,552 firm-years in the prediction sample that have sufficient data to calculate historical firm-level moments. 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 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. There are 85,142 firm-years in the prediction sample that have sufficient data to calculate historical matched-sample moments. 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 estimates that can be used to construct industry-level estimates via the law of total moments.
We convert analysts’ forecasts into predicted moments as follows. First, for each firm-year we identify all of the analysts who made a forecast of year \( t+1 \) earnings within three months after the announcement of year \( t \) earnings. We obtain analysts’ forecasts from the I/B/E/S database. Second, we delete firm-years for which we identified less than five analysts, which leaves us with 36,462 firm-years drawn from the time period spanning 1983 and 2010. Third, we keep only the initial forecast made by each analyst identified and we convert it into a forecast of ROE by dividing it by equity book value in year \( t \). Finally, we group analyst-firm-year forecasts by industry and then we calculate the industry-level moments of these forecasts. There are 810 industry-years for which: (1) the realized industry-level moment in year \( t+1 \) is non-missing; (2) there are at least ten observations in the prediction sample; and, (3) there are at least ten non-missing values of forecasted ROE that can be used to calculate the industry-year estimate.

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. 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 at least ten observations in the prediction sample; and, (3) there are at least ten non-missing values of ROE that can be used to calculate the historical industry-year estimate.

B.2 Analyses of Market-based Variables

In our regressions that use equity-market variables as the dependent variables 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 151,454 (111,921) observations available for estimating the regression that excludes (includes) the control variables.

In our regressions that use credit-market variables as the dependent variables we use all observations from the prediction sample with non-missing values of the variables in the regression. For the regressions that exclude the control variables, there are 5,664, 10,602 and 29,245 observations with sufficient data to estimate the regression involving the CDS spreads, bond yields and bond ratings, respectively. When the control variables are included in the regression these numbers decline to 5,213, 10,061 and 26,868. We have CDS spread data, bond yield data and credit ratings for the years spanning 2000 through 2009, 1984 through 2010 and 1985 through 2011, respectively.
References


Figure 1 – Graph of $b_{0,AVG}^q$ on $q$

$b_{0,AVG}^q$ 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 $\pm 1.96$ multiplied by the standard error of the average. We use Newey-West adjusted standard errors assuming a lag-length of ten years.
Figure 2 – Graph of $b_{1, AVG}^q$ on q

$b_{1, AVG}^q$ 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 3 – Graph of $b_{2,AVG}^q$ on q

$b_{2,AVG}^q$ 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 4 – Graph of \(-\left( b_{1,AVG}^q + b_{3,AVG}^q \right)\) on q

\(b_{1,AVG}^q\) is the average coefficient on \(\text{ROE}_{t-1}\) and \(b_{3,AVG}^q\) is the average coefficient on \(\text{LOSS}_{t-1} \times \text{ROE}_{t-1}\). 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 \(\pm 1.96\) multiplied by the standard error of the temporal averages of 

\(-1 \times \left( b_{1,EY}^q + b_{3,EY}^q \right) \) and 

\(-1 \times \left( b_{1,OLS}^{LO} + b_{3,OLS}^{LO} \right)\) —i.e., we use the standard error of the average of the sum \textit{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 5 - Graph of $b_{4,AVG}^q$ on q

$b_{4,AVG}^q$ 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 6 - Graph of $b^q_{5,AVG}$ on q

$b^q_{5,AVG}$ is the average coefficient on LEV$_{t-1}$. 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.
$b_{6, AVG}^q$ is the average coefficient on $\text{PAYER}_{i,t-1}$. 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 8 - Graph of $b_{AVG}^q$ on $q$

$b_{AVG}^q$ 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 $\pm$ 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 9 - 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.
The graph plots the estimated conditional non-rearranged quantile function (dashed line) and rearranged quantile function (solid line) for Agnico Eagle Mines Ltd.’s ROE in 2003.
Figure B1 – Effect of Sample-selection Criteria

Years are shown on the x-axis and the percentage of observation deleted after applying a particular criterion or after applying all of the criteria is shown on the y-axis. The graphs relate to the estimation sample.
Table 1 – Descriptive Statistics for the Estimation Sample

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

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

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} ≥ 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

<table>
<thead>
<tr>
<th></th>
<th>ROE_{it}</th>
<th>ROE_{it-1}</th>
<th>LOSS_{it-1}</th>
<th>ROE_{it-1} \times LOSS_{it-1}</th>
<th>ACC_{it-1}</th>
<th>LEV_{it-1}</th>
<th>PAYER_{it-1}</th>
<th>PAYOUT_{it-1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROE_{it}</td>
<td>0.60</td>
<td>-0.42</td>
<td>0.42</td>
<td>0.11</td>
<td>0.01</td>
<td>0.20</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>ROE_{it-1}</td>
<td>0.70</td>
<td>-0.65</td>
<td>0.85</td>
<td>0.28</td>
<td>-0.06</td>
<td>0.25</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>LOSS_{it-1}</td>
<td>-0.43</td>
<td>-0.63</td>
<td>-0.61</td>
<td>-0.20</td>
<td>0.08</td>
<td>-0.30</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>ROE_{it-1} \times LOSS_{it-1}</td>
<td>0.44</td>
<td>0.65</td>
<td>-0.99</td>
<td>0.26</td>
<td>-0.15</td>
<td>0.22</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>ACC_{it-1}</td>
<td>0.12</td>
<td>0.22</td>
<td>-0.20</td>
<td>0.21</td>
<td>-0.16</td>
<td>-0.03</td>
<td>-0.06</td>
<td></td>
</tr>
<tr>
<td>LEV_{it-1}</td>
<td>0.08</td>
<td>0.05</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.21</td>
<td>-0.01</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>PAYER_{it-1}</td>
<td>0.23</td>
<td>0.28</td>
<td>-0.30</td>
<td>0.31</td>
<td>-0.06</td>
<td>0.07</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>PAYOUT_{it-1}</td>
<td>0.31</td>
<td>0.36</td>
<td>-0.29</td>
<td>0.29</td>
<td>-0.09</td>
<td>0.11</td>
<td>0.88</td>
<td></td>
</tr>
</tbody>
</table>

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 2 – Descriptive Statistics for Quantile-based Out-of-sample Forecasts

#### Panel A – Descriptive Statistics for One-year Forecasts Horizon

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>p1</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
<th>p99</th>
<th>Maximum</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_MEAN_i,t,t+1</td>
<td>0.042</td>
<td>0.237</td>
<td>-1.334</td>
<td>-0.845</td>
<td>-0.231</td>
<td>-0.043</td>
<td>0.097</td>
<td>0.166</td>
<td>0.233</td>
<td>0.520</td>
<td>2.047</td>
<td>170,522</td>
</tr>
<tr>
<td>Q_STD_i,t,t+1</td>
<td>0.123</td>
<td>0.131</td>
<td>0.002</td>
<td>0.006</td>
<td>0.023</td>
<td>0.042</td>
<td>0.076</td>
<td>0.159</td>
<td>0.279</td>
<td>0.655</td>
<td>1.231</td>
<td>170,522</td>
</tr>
<tr>
<td>Q_CV_i,t,t+1</td>
<td>2.366</td>
<td>221.964</td>
<td>0.012</td>
<td>0.039</td>
<td>0.138</td>
<td>0.292</td>
<td>0.716</td>
<td>1.267</td>
<td>2.240</td>
<td>13.000</td>
<td>90,342.540</td>
<td>170,522</td>
</tr>
<tr>
<td>Q_SKEW_i,t,t+1</td>
<td>0.192</td>
<td>2.032</td>
<td>-8.423</td>
<td>-4.147</td>
<td>-1.376</td>
<td>-0.881</td>
<td>-0.590</td>
<td>0.586</td>
<td>3.780</td>
<td>5.226</td>
<td>8.103</td>
<td>170,522</td>
</tr>
<tr>
<td>Q_KURT_i,t,t+1</td>
<td>6.243</td>
<td>9.215</td>
<td>-1.572</td>
<td>-1.036</td>
<td>-0.599</td>
<td>0.662</td>
<td>1.748</td>
<td>8.993</td>
<td>20.703</td>
<td>36.374</td>
<td>85.842</td>
<td>170,522</td>
</tr>
</tbody>
</table>

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 year t estimate of the standard deviation of ROE\_i,t+1. Q\_CV\_i,t,t+1 = Q\_STD\_i,t,t+1 ÷ |Q\_MEAN\_i,t,t+1|. 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

<table>
<thead>
<tr>
<th></th>
<th>Q_MEAN_{i,t,t+1}</th>
<th>Q_STD_{i,t,t+1}</th>
<th>Q_SKEW_{i,t,t+1}</th>
<th>Q_KURT_{i,t,t+1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_MEAN_{i,t,t+1}</td>
<td>-0.62</td>
<td>0.20</td>
<td>0.27</td>
<td>(-9.97)</td>
</tr>
<tr>
<td></td>
<td>(6.56)</td>
<td>(3.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q_STD_{i,t,t+1}</td>
<td>-0.47</td>
<td>-0.26</td>
<td>-0.38</td>
<td>(-6.90)</td>
</tr>
<tr>
<td></td>
<td>(-9.50)</td>
<td>(-14.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q_SKEW_{i,t,t+1}</td>
<td>0.21</td>
<td>-0.34</td>
<td>0.67</td>
<td>(3.70)</td>
</tr>
<tr>
<td></td>
<td>(-9.03)</td>
<td>(4.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q_KURT_{i,t,t+1}</td>
<td>0.38</td>
<td>-0.63</td>
<td>0.14</td>
<td>(2.83)</td>
</tr>
<tr>
<td></td>
<td>(-7.93)</td>
<td>(1.24)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 3 – Regressions of $R_{\text{STD}_{\text{IND},t+1}}$ on $Q_{\text{STD}_{\text{IND},t,t+1}}$ and Alternative Estimates

<table>
<thead>
<tr>
<th></th>
<th>A = Historical Firm-level</th>
<th>A = Historical Matched-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$Q_{\text{STD}_{\text{IND},t,t+1}}$</td>
<td>1.41  ***</td>
<td>1.39  ***</td>
</tr>
<tr>
<td></td>
<td>(16.39)</td>
<td>(14.56)</td>
</tr>
<tr>
<td>$A_{\text{STD}_{\text{IND},t,t+1}}$</td>
<td>0.13</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(1.87)</td>
<td>(-0.23)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.01</td>
<td>0.31  ***</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(7.63)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.29</td>
<td>0.04  ***</td>
</tr>
<tr>
<td>Vuong</td>
<td>5.68  ***</td>
<td></td>
</tr>
<tr>
<td>Industry-years</td>
<td>1,960</td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A = Analysts' Forecasts</th>
<th>A = Historical Industry-level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>$Q_{\text{STD}_{\text{IND},t,t+1}}$</td>
<td>5.43</td>
<td>1.47  ***</td>
</tr>
<tr>
<td></td>
<td>(1.85)</td>
<td>(10.68)</td>
</tr>
<tr>
<td>$A_{\text{STD}_{\text{IND},t,t+1}}$</td>
<td>0.13</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(1.37)</td>
<td>(1.03)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.93</td>
<td>0.27  ***</td>
</tr>
<tr>
<td></td>
<td>(-1.38)</td>
<td>(4.31)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.35</td>
<td>0.18</td>
</tr>
<tr>
<td>Vuong</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>Industry-years</td>
<td>810</td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

$R_{\text{STD}_{\text{IND},t+1}}$ is the dependent variable. It equals the realized within-industry standard deviation of ROE$_{i,t+1}$. $Q_{\text{STD}_{\text{IND},t,t+1}}$ 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) $A_{\text{STD}_{i,t+1}}$ 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) $A_{\text{STD}_{i,t,t+1}}$ 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) $A_{\text{STD}_{i,t+t+1}}$ equals the within-industry standard deviation of FROE$_{ij,t}$. FROE$_{ij,t}$ equals the ratio of analysts’ $j$’s forecast of firm $i$’s earnings in year $t+1$ to firm $i$’s year $t$ equity book value. In columns (11) and (12) $A_{\text{STD}_{i,t,t+1}}$ equals the within-industry standard deviation of ROE$_{i,t}$. 

63
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 4 – Regressions of $R_{\text{SKEW}_{\text{IND},t+1}}$ on $Q_{\text{SKEW}_{\text{IND},t,t+1}}$ and Alternative Estimates

<table>
<thead>
<tr>
<th></th>
<th>A = Historical Firm-level</th>
<th>A = Historical Matched-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$Q_{\text{SKEW}_{\text{IND},t,t+1}}$</td>
<td>0.56 ***</td>
<td>0.56 ***</td>
</tr>
<tr>
<td></td>
<td>(14.70)</td>
<td>(14.88)</td>
</tr>
<tr>
<td>$A_{\text{SKEW}_{\text{IND},t,t+1}}$</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(1.61)</td>
<td>(-0.10)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.51 ***</td>
<td>-1.43 ***</td>
</tr>
<tr>
<td></td>
<td>(-4.42)</td>
<td>(-7.19)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Vuong</td>
<td>10.77 ***</td>
<td></td>
</tr>
<tr>
<td>Industry-years</td>
<td>1,960</td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A = Analysts’ Forecasts</th>
<th>A = Historical Industry-level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>$Q_{\text{SKEW}_{\text{IND},t,t+1}}$</td>
<td>0.47 ***</td>
<td>0.42 ***</td>
</tr>
<tr>
<td></td>
<td>(8.35)</td>
<td>(6.88)</td>
</tr>
<tr>
<td>$A_{\text{SKEW}_{\text{IND},t,t+1}}$</td>
<td>0.00</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(-0.24)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.42 *</td>
<td>-0.26 *</td>
</tr>
<tr>
<td></td>
<td>(-2.11)</td>
<td>(-2.53)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>Vuong</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Industry-years</td>
<td>810</td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

$R_{\text{SKEW}_{\text{IND},t+1}}$ is the dependent variable. It equals the realized within-industry skewness of ROE$_{i,t+1}$. $Q_{\text{SKEW}_{\text{IND},t,t+1}}$ 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) $A_{\text{SKEW}_{i,t,t+1}}$ 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) $A_{\text{SKEW}_{i,t,t+1}}$ 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) $A_{\text{SKEW}_{i,t,t+1}}$ equals the within-industry skewness of FROE$_{i,j,t}$. FROE$_{i,j,t}$ equals the ratio of analysts’ j’s forecast of firm i’s earnings in year t+1 to firm i’s year t equity book value. In columns (11) and (12) $A_{\text{SKEW}_{i,t,t+1}}$ equals the within-industry skewness deviation of ROE$_{i,t}$.
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 5 – Regressions of $R_{KURT_{IND,t+1}}$ on $Q_{KURT_{IND,t,t+1}}$ and Alternative Estimates

<table>
<thead>
<tr>
<th></th>
<th>A = Historical Firm-level</th>
<th>A = Historical Matched-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$Q_{KURT_{IND,t,t+1}}$</td>
<td>0.67 ***</td>
<td>0.64 ***</td>
</tr>
<tr>
<td></td>
<td>(4.43)</td>
<td>(4.54)</td>
</tr>
<tr>
<td>$A_{KURT_{IND,t,t+1}}$</td>
<td>0.03 *</td>
<td>0.03 *</td>
</tr>
<tr>
<td></td>
<td>(2.06)</td>
<td>(2.03)</td>
</tr>
<tr>
<td>Intercept</td>
<td>3.77 ***</td>
<td>10.92 ***</td>
</tr>
<tr>
<td></td>
<td>(3.66)</td>
<td>(17.97)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.17</td>
<td>0.08</td>
</tr>
<tr>
<td>Vuong</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>Industry-years</td>
<td>1,960</td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A = Analysts' Forecasts</th>
<th>A = Historical Industry-level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>$Q_{KURT_{IND,t,t+1}}$</td>
<td>0.63 ***</td>
<td>0.28 ***</td>
</tr>
<tr>
<td></td>
<td>(4.01)</td>
<td>(7.67)</td>
</tr>
<tr>
<td>$A_{KURT_{IND,t,t+1}}$</td>
<td>0.19 ***</td>
<td>0.16 ***</td>
</tr>
<tr>
<td></td>
<td>(10.93)</td>
<td>(9.01)</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.15 ***</td>
<td>3.32 ***</td>
</tr>
<tr>
<td></td>
<td>(3.78)</td>
<td>(10.00)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.21</td>
<td>0.31</td>
</tr>
<tr>
<td>Vuong</td>
<td>-1.44</td>
<td></td>
</tr>
<tr>
<td>Industry-years</td>
<td>810</td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

$R_{KURT_{IND,t+1}}$ is the dependent variable. It equals the realized within-industry kurtosis of $ROE_{it+1}$. $Q_{KURT_{IND,t,t+1}}$ 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 $ROE_{it+1}$. In columns (2) and (3) $A_{KURT_{t,t+1}}$ 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 $ROE_{it+1}$. In columns (5) and (6) $A_{KURT_{t,t+1}}$ 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 $ROE_{it+1}$. In columns (8) and (9) $A_{KURT_{t,t+1}}$ equals the within-industry kurtosis of $FROE_{ij,t}$. $FROE_{ij,t}$ equals the ratio of analysts’ $j$’s forecast of firm $i$’s earnings in year $t+1$ to firm $i$’s year $t$ equity book value. In columns (11) and (12) $A_{KURT_{t,t+1}}$ equals the within-industry kurtosis of $ROE_{it}$. 

[67]
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 6 – Regressions of Equity-market Variables on Quantile-based Predictions

<table>
<thead>
<tr>
<th></th>
<th>Implied Cost of Capital</th>
<th>Forward-earnings-to-price</th>
<th>Book-to-price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Q_MEAN&lt;sub&gt;i,t,t+1&lt;/sub&gt;</td>
<td>0.066 ***</td>
<td>0.087 ***</td>
<td>0.456 ***</td>
</tr>
<tr>
<td></td>
<td>(4.74)</td>
<td>(10.14)</td>
<td>(12.88)</td>
</tr>
<tr>
<td>Q_STD&lt;sub&gt;i,t,t+1&lt;/sub&gt;</td>
<td>-0.114 ***</td>
<td>-0.139 ***</td>
<td>-0.318 ***</td>
</tr>
<tr>
<td></td>
<td>(-11.33)</td>
<td>(-15.88)</td>
<td>(-8.13)</td>
</tr>
<tr>
<td>Q_SKEW&lt;sub&gt;i,t,t+1&lt;/sub&gt;</td>
<td>-0.004 ***</td>
<td>-0.002 *</td>
<td>-0.006 ***</td>
</tr>
<tr>
<td></td>
<td>(-4.09)</td>
<td>(-2.44)</td>
<td>(-5.54)</td>
</tr>
<tr>
<td>Q_KURT&lt;sub&gt;i,t,t+1&lt;/sub&gt;</td>
<td>0.001 ***</td>
<td>0.001 *</td>
<td>0.003 ***</td>
</tr>
<tr>
<td></td>
<td>(3.08)</td>
<td>(2.28)</td>
<td>(4.07)</td>
</tr>
<tr>
<td>SIZE&lt;sub&gt;i,t&lt;/sub&gt;</td>
<td>-0.006 ***</td>
<td>-0.006</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(-7.37)</td>
<td>(-1.66)</td>
<td>(-6.61)</td>
</tr>
<tr>
<td>BETA&lt;sub&gt;i,t&lt;/sub&gt;</td>
<td>0.005 ***</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(7.32)</td>
<td>(-0.20)</td>
<td>(7.91)</td>
</tr>
<tr>
<td>RET_MEAN&lt;sub&gt;i,t&lt;/sub&gt;</td>
<td>-0.005 *</td>
<td>0.013 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.26)</td>
<td>(2.40)</td>
<td></td>
</tr>
<tr>
<td>RET_STD&lt;sub&gt;i,t&lt;/sub&gt;</td>
<td>-0.077 ***</td>
<td>-0.108 ***</td>
<td>-1.325 ***</td>
</tr>
<tr>
<td></td>
<td>(-2.99)</td>
<td>(-3.30)</td>
<td>(3.91)</td>
</tr>
<tr>
<td>RET_SKEW&lt;sub&gt;i,t&lt;/sub&gt;</td>
<td>0.001 ***</td>
<td>0.003 ***</td>
<td>0.018 ***</td>
</tr>
<tr>
<td></td>
<td>(4.03)</td>
<td>(3.85)</td>
<td>(4.03)</td>
</tr>
<tr>
<td>RET_KURT&lt;sub&gt;i,t&lt;/sub&gt;</td>
<td>0.000</td>
<td>0.000</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(1.70)</td>
<td>(0.29)</td>
<td>(1.81)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.093 ***</td>
<td>0.123 ***</td>
<td>0.021 *</td>
</tr>
<tr>
<td></td>
<td>(9.13)</td>
<td>(11.43)</td>
<td>(2.05)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.35</td>
<td>0.46</td>
<td>0.56</td>
</tr>
<tr>
<td>Firm-years</td>
<td>151,454</td>
<td>111,921</td>
<td>151,454</td>
</tr>
<tr>
<td>Years</td>
<td>39</td>
<td>39</td>
<td>39</td>
</tr>
</tbody>
</table>

Implied Cost of Capital, ICC<sub>i,t</sub>, earnings-to-price, EP<sub>i,t</sub>, and book-to-price, BP<sub>i,t</sub>, are the dependent variables. ICC<sub>i,t</sub> is based on the model described in Gebhardt et al. [2001]. However, following Hou et al. [2012], our forecasts of ROE for years t+1 through t+3 are obtained from the OLS version of equation (2). EP<sub>i,t</sub> (BP<sub>i,t</sub>) 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<sub>i,t,t+1</sub> is the year t quantile-based estimate of the mean of ROE<sub>i,t+1</sub>. Q_STD<sub>i,t,t+1</sub> is the year t quantile-based estimate of the standard deviation of ROE<sub>i,t+1</sub>. Q_SKEW<sub>i,t,t+1</sub> is the year t quantile-based estimate of the skewness of ROE<sub>i,t+1</sub>. Q_KURT<sub>i,t,t+1</sub> is the year t quantile-based estimate of the excess kurtosis of ROE<sub>i,t+1</sub>. SIZE<sub>i,t</sub> equals firm i’s equity market value at the end of year t. BETA<sub>i,t</sub> is estimated in three steps. First, for each industry based on Fama-French 48 classifications, we estimate the slope coefficient obtained from an OLS weighted regressions of industry-level monthly returns on the contemporaneous return on the market portfolio. The weights equal the contemporaneous
equity market values. We use monthly returns drawn from a 60-month period ending three months after the last month of fiscal-year t. Second, we unlever this estimated levered industry beta using the industry leverage ratio. Third, we estimate the levered firm beta using firm i’s leverage ratio to re-lever the unlevered industry beta. RET_{it} is the annual stock return of firm i in year t. RET_{STD_{it}} is the volatility of historical market-model residuals for firm i. RET_{SKEW_{it}} is the skewness of historical market-model residuals for firm i. RET_{KURT_{it}} 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.
Credit-default-swap, CDS spreads, bond ratings and credit ratings are the dependent variables. CDS spreads are obtained from the MarkIt Group and equal the quoted spread on 5-year CDS contracts of senior unsecured debts with modified restructuring clauses. Bond yields are obtained from either the Trace or Mergent database. Credit ratings range between 1 and 24 and are
obtained from Standard and Poors. All three variables are measured four months after the end of
fiscal year t. MEAN \_ROA_{i,t+1} is the year t quantile-based estimate of the mean of \textit{ROA}_{i,t+1}. 
STD \_ROA_{i,t+1} is the year t quantile-based estimate of the standard deviation of \textit{ROA}_{i,t+1}. 
SKEW \_ROA_{i,t+1} is the year t quantile-based estimate of the skewness of \textit{ROA}_{i,t+1}. 
KURT \_ROA_{i,t+1} is the year t quantile-based estimate of the excess kurtosis of \textit{ROA}_{i,t+1}. BP_{i,t} equals the ratio of firm i’s year t equity book value to its equity market value at the end year t. 
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 \_SKEW_{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. T2MAT_{i,t} is the remaining number of years to the 
date on which the bond matures. BOND \_SIZE_{i,t} is the natural log of the aggregate par value of the 
bond at the date of issuance. Industry fixed effects based on Fama-French 12 classifications 
are included.

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. 
Temporal standard errors in the analyses of bond yields and bond ratings reflect the Newey-West 
adjustment assuming a ten-year lag length. Temporal standard errors in the analyses of CDS 
spreads reflect the Newey-West adjustment assuming a nine-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.