Information Quality, Systematic Risk and the Cost of Capital*

Chris Armstrong† Snehal Banerjee‡ Carlos Corona§

August 2009

Abstract

Most models that examine the relationship between information quality and cost of capital do so in a single firm setting, and predict that the relationship is negative. Given the lack of consistent empirical support for this prediction, we reconsider this relationship in a setup where covariance risk, and not variance risk, determines the cost of capital. We allow investors to independently learn about systematic risk-factors and firm-specific betas. We show that an increase in information quality of either type (systematic or firm-specific) has two effects on a firm’s expected returns: (i) a “beta” effect which decreases expected returns when beta is positive, but increases expected returns when beta is negative, and (ii) a “convexity” effect which generally increases the expected returns. We test these predictions using proxies for systematic information quality based on the VIX and firm-specific information quality based on analyst forecasts and accruals quality, and find evidence consistent with our predictions.

---

*We have benefited from comments from Andres Almazan, John Core, Lorenzo Garlappi, Mike Faulkender, Mike Fishman, Yaniv Konchitchki, Bob Korajczyk, Arvind Krishnamurthy, Dimitris Papanikolaou, Jonathan Parker, Mitchell Petersen, Ernst Schaumburg, Viktor Todorov, and seminar participants at the University of Toronto. All errors are our own.

†The Wharton School, University of Pennsylvania, email: carms@wharton.upenn.edu

‡Kellogg School of Management, Northwestern University, email: snehal-banerjee@kellogg.northwestern.edu

§McCombs School of Business, University of Texas, email: Carlos.Corona@mccombs.utexas.edu
1 Introduction

Does information quality affect the cost of capital, and if so, how? Despite its theoretical and empirical importance, this has largely remained an open question. Most models in the finance and accounting literature predict that increasing information quality (IQ) reduces informational incompleteness or asymmetry and thus leads to lower costs of capital. However, the empirical evidence for this hypothesis has been inconclusive at best. One possible reason for the lack of empirical support is that these predictions are derived in single firm models, where variance risk determines the cost of capital. As a result, it is not clear whether these predictions extend to settings with multiple firms and diversifiable risk.

While much of the existing theoretical literature has focused on extending the analysis within the single firm framework, we consider a setup in which systematic risk, or covariance risk, determines the expected return on an asset. The premise of our paper is that if information quality affects the cost of capital for a firm, it must be through systematic risk. Furthermore, the effect of information quality must depend on the risk-factor loading, or beta, of the firm.

In order to parsimoniously capture the notion of systematic and idiosyncratic risk in a simple setup, we begin with a representative investor, endowment economy. As we show later, the results can be extended to more general settings, but the simple Lucas-tree economy provides clear intuition for our predictions. The representative investor can independently learn about the systematic risk factors that drive the aggregate uncertainty in the economy, and about the firm-specific factor loadings, or betas. In particular, systematic information quality refers to the quality of information available about the systematic risk factors, and firm-specific information quality refers to the quality of information available about the cash-flow beta of a specific firm.

Somewhat surprisingly, we find that an increase in either type of information quality has a similar effect on a firm’s cost of capital. An increase in information quality leads to a decrease in the expected return when the firm’s beta is positive, but an increase in expected return when the firm’s beta is negative - we call this the “beta” effect. In addition, an increase in systematic information quality leads to a decrease in aggregate uncertainty, which leads to an increase in the risk-free rate - we call this the “convexity” effect. Moreover, we find that there is a similar convexity effect of an increase in firm-specific information quality when systematic information quality is relatively low. As a result, the overall effect of increasing information quality on the cost of capital is negative only if the firm’s beta is large enough. We then show that the above results extend to a partial equilibrium setup with a multi-factor pricing kernel (or stochastic discount factor), and this provides the basis for our empirical analysis.

Since the existing empirical literature has largely focused on a negative relationship between information quality and cost of capital, it is unable to capture the non-monotonic

\footnote{There are important exceptions in the existing literature, including Pastor and Veronesi (2003) and Lambert, Leuz and Verrecchia (2007). We discuss these and other papers in the related literature section below.}
relationship generated by our model. We test our predictions directly, using various proxies for IQ, short-horizon four-factor betas as proxies for conditional betas and realized stock returns as a proxy for a firm’s cost of capital (or expected returns). Our firm-specific IQ proxies are based on mean analyst forecast errors, variance in earnings per share, and accruals quality, and our systematic IQ proxy is based on the CBOE Volatility Index (VIX). The empirical evidence is generally consistent with the above predictions across the various IQ proxies.

The intuition for our results is as follows. Firms generate stochastic cash-flows which can be priced using the representative agent’s marginal rate of substitution, and each firm’s cost of capital (or expected return) depends on the systematic risk, or beta, of its cash-flows. The representative investor can learn about the future realization of consumption growth (i.e. about the systematic risk-factor) and about how a firm’s cash-flows are correlated with consumption growth (i.e. about firm-specific betas). While we do impose some restrictions on the distribution of representative investor’s prior beliefs and learning process, we do not need to assume a specific information structure. To keep the analysis tractable, we assume that aggregate (log) consumption growth and firm-specific cash-flow betas are independent and conditionally normal. Hence the model is able to capture a large subset of standard information environments, and in particular, can allow for multiple signals about consumption growth and betas.

The reduction in the representative investor’s uncertainty as a result of learning from available information is a measure of the quality of the information. To capture this notion, we define systematic information quality as the decrease in the investor’s prior variance about consumption growth as a result of conditioning on (i.e. learning from) information available to him. Similarly, we define firm-specific information quality as the decrease in the investor’s prior variance about a firm’s beta due to the information available. Since the representative investor conditions on all the available information when pricing the cash-flows of each firm, both systematic and firm-specific information quality affects a firm’s expected return.

As we show, the effect of information quality on the expected return of the firm is two-fold and non-monotonic. First, increasing the information quality of either type attenuates the conditional covariance between cash-flows and the systematic risk-factor towards zero - this is the “beta” effect. Intuitively, this is because one can think of the conditional covariance as the correlation between cash-flows and the risk-factor multiplied by the product of the conditional standard deviations. Increasing information quality of either type leads to a decrease in the conditional standard deviations and therefore leads to an attenuation of the beta. This implies that higher information quality decreases expected returns (or cost of capital) for positive beta stocks, but increases expected returns for negative beta stocks.

The second effect is more subtle. An increase in the systematic information quality reduces aggregate uncertainty in the economy. Since the price of the risk-free asset is convex in aggregate uncertainty, an increase in systematic IQ leads to an increase in the risk-free rate, and therefore an increase in the expected return of all stocks. However, an increase in firm-specific information quality and the resulting decrease in firm-specific cash-flow uncertainty has two potentially off-setting effects. A decrease in uncertainty about beta reduces the systematic risk of the firm, which leads to a decrease in expected returns. However, a
decrease in uncertainty about beta also leads to a decrease in cash-flow volatility, which leads to an increase in expected returns since prices are convex in volatility. Moreover, we show the second effect dominates when the systematic information quality is low, but the first effect dominates when systematic information quality is high.

We show that these results extend to a partial equilibrium setup with a multi-factor pricing kernel. The factor pricing model is more appropriate for empirical tests of the model’s predictions. Moreover, our empirical tests specifically account for the non-monotonic nature of these relationships. Following Lewellen and Nagel (2006), we use short horizon, rolling window estimates of four-factor betas to directly proxy for conditional betas. For systematic information quality, we use a proxy based on the VIX volatility index, and for firm-specific information quality, we use measures based on analyst forecast errors and accruals quality. To test for the “beta” and “convexity” effects of information quality on expected returns, we use a generalized method of moments (GMM) specification that allows us to estimate both effects simultaneously. We find that the results are generally consistent with our predictions, regardless of which proxy we use. This is especially reassuring given that our proxies, though noisy, do not seem to be spuriously correlated for any obvious reasons.

Our empirical evidence also suggests that the convexity effect of firm-specific information quality is positive at the monthly frequency, but negative at the annual frequency. This is consistent with a prediction of the model that implies that the convexity effect of firm-specific information quality depends on the systematic information quality, and that investors learn more about systematic risk-factors at longer horizons (i.e. systematic information quality is larger at an annual horizon than at a monthly horizon).

The rest of the paper is organized as follows. The next subsection discusses some of the related theoretical and empirical literature. Section 2 is the theoretical section of the paper - section 2.1 presents the general equilibrium model and the basic analysis, and section 2.3 extends the results to a partial equilibrium, multi-factor pricing model. Section 2.4 studies the relationship between information quality and estimated betas, and presents the empirical predictions of the model. Section 3 presents the empirical analysis, beginning with data descriptions in section 3.1, empirical specifications in 3.2 and a discussion of the results in section 3.3. Section 4 concludes.

1.1 Literature Review

The relationship between information quality and cost of capital remains inconclusive from both a theoretical and an empirical perspective. A standard approach has been to argue that increasing the quality of public information reduces informational asymmetry which, in turn increases liquidity and so reduces the cost of capital (e.g. Diamond and Verrecchia (1991), Baiman and Verrecchia (1996), Easley and O’Hara (2004)). There has been some empirical support for this hypothesis. For example, Easley, Hvidkjaer and O’Hara (2002) proxy for information risk using the Easley, Kiefer, O’Hara and Paperman (1996) PIN measure and show that this is priced. Francis, LaFond, Olsson and Schipper (2004, 2005) find that accruals quality has significant explanatory power for the cross section of expected returns.
Leuz and Verrecchia (2000) show that firms that voluntarily commit to switch from German financial reporting to higher quality international financial reporting exhibit a decline in the information asymmetry component of the cost of capital (which they capture using the bid-ask spread and trading volume). Barth, Landsman and Konchitchki (2007) show that firms with higher financial statement transparency, as measured by the covariance between earnings and returns, have lower expected returns and systematic risk. Finally, Francis, Nanda and Olsson (2008) show that the effect of voluntary disclosure on cost of capital can be partially explained by earnings quality.

On the other hand, it is not clear whether the predictions from these single-firm models should survive in a setup with multiple firms and diversifiable risk. In a multi-asset, noisy rational expectations model, Hughes, Liu and Liu (2007) show that information does not have any cross-sectional effects on the cost of capital. This is because firm-specific information is diversified away and systematic information affects factor risk-premia but not factor loadings in large economies. Moreover, other empirical work finds little or no evidence of a negative relation between cost of capital and information quality. For example, Botosan (1997) finds no evidence of a relationship between disclosure and the cost of capital for her full sample of firms and only weak evidence of a negative relationship for the sub-sample of firms with low analyst coverage. Botosan and Plumlee (2002) show that only the sub-component of AIMR scores related to annual report disclosures has a modest negative association with the cost of capital. Core, Guay and Verdi (2008) test Lambert, Leuz and Verrecchia’s (2007) prediction that asymmetric information effects should be diversified away, and empirically document that accruals quality is not a priced risk factor. Similarly, Duarte and Young (2007) decompose the PIN measure into asymmetric information and liquidity components and show that only the liquidity component is priced. It is important to note that since information quality is not a priced risk factor in our model, but affects expected returns through factor loadings, the results in these papers are not inconsistent with our theoretical predictions or empirical results.

In light of this conflicting evidence, Lambert, Leuz and Verrecchia (2007) (henceforth LLV) provide an important step in clarifying the relationship between cost of capital and information quality. They present a multi-firm model with systematic and diversifiable risk and show that information quality affects the cost of capital through two channels. The first effect, analogous to the “beta” effect in our model, is through the covariance of a firm’s cash-flows with those of the market. The second effect is an indirect one through the real decisions of the firm. As a result, the overall effect of increasing information quality on the cost of capital is ambiguous. LLV show that while the “variance” effects that drive the cost of capital results in single firm models vanish as the economy gets large, this is not the case for the “covariance” effects if the number of firms (i.e. the aggregate risk in the economy) and the number of investors (i.e. the aggregate risk bearing capacity) grow together. Finally, they show that although information quality is not a risk factor, it affects the cross-section of the cost of capital through its effect on the factor loadings.

We build on LLV’s intuition that information quality affects the cost of capital through systematic risk. However, our model is more general as it allows investors to update not only about the aggregate risk factors in the economy but also about the firm-specific cash-
flow betas. In particular, LLV’s setup can be studied as the special case of our model where investors have no uncertainty about the beta of a firm’s cash-flows. We are able to distinguish between systematic and firm-specific information quality and study how each affects the expected return, or cost of capital, of a firm.

The intuition for the convexity effect in our model is closely related to that in the learning model of Pastor and Veronesi (2003). In their model, investors are uncertain about a firm’s profitability. In a standard asset pricing model, the authors show that since prices are convex in the uncertainty about cash-flows (i.e. there is an option-value to having more volatile cash-flows), more firm-specific uncertainty leads to higher market to book ratios. A similar convexity effect in our model leads to a decrease in expected returns when firm-specific information quality increases. In particular, since investors do not learn about the risk-factor in Pastor and Veronesi (2003), their setup is analogous to a special case of our model where investors learn only about the beta of the firm. In general, our model also allows investors to learn about aggregate risk-factors and we find that the overall effect of an increase in firm-specific information quality depends on the extent to which investors are learning about the systematic risk-factors.

Our choice of using a standard asset pricing setting offers additional advantages. In particular, the model’s predictions are more applicable for empirical analysis - for instance, unlike in LLV’s normal-exponential setup, prices in our model are always non-negative. In addition, the intuition of our model carries through in “large economies” without any additional assumptions since expected returns are determined by the covariance of cash-flows with the pricing kernel.

Our theoretical analysis, coupled with our supporting empirical evidence, contributes to reconciling the mixed empirical evidence in the existing literature. Since the predicted relationship between cost of capital and information quality is non-monotonic, standard linear specifications used in earlier work are misspecified. Our model suggests that information quality affects both risk-factor loadings (through the “beta” effect) and risk premia (through the “convexity” effect), but often these affect a firm’s cost of capital in opposite directions. Consequently, finding either effect independently of the other is difficult. We are able to empirically identify these two confounding effects by jointly modeling them and using a GMM estimation approach. Finally, we find evidence consistent with our predictions using very different proxies for information quality, which suggests that these effects are quite robust.

Our paper is also related to the literature on estimation risk. The early literature in this area (e.g. Brown (1978), Bawa and Brown (1979)) suggested that estimation risk should be diversifiable and therefore not priced. However, subsequent work showed that the availability of different amounts of information across securities may yield a non-diversifiable effect on equilibrium prices, even in a CAPM setting (e.g. Barry and Brown (1985), Clarkson and Thomson (1990, 1996)). More recently, Kumar et al. (2008) extend this literature by proposing a conditional CAPM in which investors are uncertain about the higher moments of the distribution of returns and information signals. In another recent paper, Adrian and Franzoni (2008) extend the conditional CAPM by introducing unobservable long-run changes in conditional betas. In their setup, investors rationally learn about both the level of beta and its long-run mean from the history of realized returns. Instead of developing
predictions about specific models of estimation risk and learning about betas, we pursue a complementary approach by studying the effects of learning about betas and risk-factors jointly, and show that the effect of estimation risk may depend on the extent to which investors learn about the risk-factors.

2 Theoretical Results

2.1 General Equilibrium Consumption Economy

We consider a representative agent, single period, endowment economy. The representative agent has power utility with risk aversion $\gamma$ and a discount factor $\delta$. The aggregate consumption process for the agent is given by:

$$C_{t+1} = C_t \exp \{g + c_{t+1}\} \quad \text{where} \quad c_{t+1} \sim N(0, v_c)$$  \hspace{1cm} (1)

where $g$ represents the mean growth in aggregate consumption, and $c_{t+1}$ is the normally distributed shock to consumption growth. There are $N$ firms in the economy which produce cash-flows or dividends. Firm $i$’s cash-flows are given by

$$D_{i,t+1} = D_{i,t} \exp \{\mu_{i,t+1} + d_{i,t+1} - \frac{1}{2}v_d\} \quad \text{where} \quad d_{i,t+1} \sim N(0, v_d)$$  \hspace{1cm} (2)

and $d_{i,t+1}$ is normally distributed, independent of $c_{t+1}$ and across firms. The mean growth rate $\mu_{i,t+1}$ has a predictable component $\bar{\mu}_{i,t}$ and can always be projected on the consumption growth process as:

$$\mu_{i,t+1} = \bar{\mu}_{i,t} + \beta_{i,t+1} c_{t+1} + e_{i,t+1} \quad \text{where} \quad e_{i,t+1} \sim N(0, v_{e,i})$$  \hspace{1cm} (3)

and $e_{i,t+1}$ is assumed to be normally distributed and independent across firms. Note that the systematic risk of the firm’s cash-flows is captured by $\beta_{i,t+1}$ which determines the conditional covariance of the firm’s cash-flows with the aggregate consumption process. The representative investor may face uncertainty about the beta of firm $i$, and we denote the prior beliefs about $\beta_{i,t+1}$ as being given by a normal distribution:

$$\beta_{i,t+1} \sim N(\beta_i, v_{\beta,i})$$  \hspace{1cm} (4)

Conditional on all the information available at time $t$, the representative agent’s Euler equation can be used to price asset $i$’s dividend in the next period:

$$P_{i,t} = E_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} D_{i,t+1} \right]$$  \hspace{1cm} (5)

Here $E_t$ denotes the expectation of a random variable conditional on all the information at date $t$, and we assume that this information preserves the conditional normality of $c_{t+1}$ and $\beta_{i,t+1}$. We can then define the gross expected return of firm $i$’s cash-flows as:

$$E[R_{i,t+1}] = E \left[ \frac{D_{i,t+1}}{P_{i,t+1}} \right]$$  \hspace{1cm} (6)
2.2 Conditional Distributions and Expected Returns

In order to evaluate the unconditional expectation in (6), we begin by describing the conditional distribution of \(c_{t+1}\) and \(\beta_{i,t+1}\) given the information at time \(t\). First, we stack the two random variables into a single vector \(X_{i,t+1}\), and denote the investor’s unconditional beliefs about it by:

\[
X_{i,t+1} = \begin{pmatrix} c_{t+1} \\ \beta_{i,t+1} \end{pmatrix} \sim N \left( x_{i,0}, V_{i,0} \right) \quad \text{where} \quad x_{i,0} = \begin{pmatrix} 0 \\ \beta_i \end{pmatrix} \quad \text{and} \quad V_{i,0} = \begin{pmatrix} v_c & 0 \\ 0 & v_{b,i,t} \end{pmatrix}
\]

(7)

In particular, we assume that \(\beta_{i,t+1}\) and \(c_{t+1}\) are independent. This assumption is primarily made for analytical tractability and to clarify the intuition behind the results. In particular, this implies that the firm-specific information quality about \(\beta_{i,t+1}\) does not affect the representative investors beliefs about consumption growth \(c_{t+1}\), and vice versa.

Similarly, since we have assumed that any information at time \(t\) available to the investor preserves conditional normality, we can denote the investor’s beliefs about \(X_{i,t+1}\) conditional on the information at time \(t\) by:

\[
X_{i,t+1} | I_t \sim N \left( x_{i,t}, V_{i,t} \right) \quad \text{where} \quad V_{i,t} = \begin{pmatrix} v_{c,t} & 0 \\ 0 & v_{b,i,t} \end{pmatrix}
\]

(8)

In turn, this implies that the unconditional distribution of the conditional mean \(x_{i,t}\) is given by

\[
x_{i,t} \sim N \left( x_{i,0}, V_{i,0} - V_{i,t} \right)
\]

(9)

Note that this follows from the Law of Iterated Expectations and the Law of Total Variance, and since we have assumed that the random variables are conditionally normal, the expected conditional variance is constant and given by \(V_{i,t}\). Note that the above expressions hold under a wide range of possible information structures, including those in which investors receive multiple signals about betas and risk-factors. As a result, by relying only on the conditional normality of beta and the aggregate risk-factors, we are able to study the effects of information quality on expected returns for a large class of models.

Finally, denote \(G \equiv \frac{1}{2} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\) and \(a \equiv -\begin{pmatrix} \gamma \\ 0 \end{pmatrix}\), and denote \(H_{i,t} = (I - 2GV_{i,t})\). This implies that the price can be expressed as

\[
P_{i,t} = \delta D_{i,t} E_t \left[ \exp \left\{ \gamma g + \bar{\mu}_{i,t} + e_{i,t+1} + X_{i,t+1}' G X_{i,t+1} + a' X_{i,t+1} \right\} \right]
\]

and the conditional expectation of the dividend can be expressed as

\[
E_t [D_{i,t+1}] = D_{i,t} E_t \left[ \exp \left\{ \bar{\mu}_{i,t} + e_{i,t+1} + X_{i,t+1}' G X_{i,t+1} \right\} \right]
\]

As a result, the unconditional expected gross return on firm \(i\)’s cash-flows is given by the following result.
Proposition 1 Suppose that firm $i$’s cash-flows are characterized by (2) - (3), and let $H_{i,t} = (I - 2GV_{i,t})$. If the determinant of $H_{i,0}$ is positive i.e.

$$|H_{i,0}| = 1 - v_c v_{b,i} > 0$$  \hspace{1cm} (10)

then the expected return on firm $i$’s cash-flows is given by:

$$E[R_{i,t+1}] = \frac{1}{\delta} \exp \left\{ \gamma g - x_{i,0}'H_{i,t}^{-1}a + \frac{1}{2} \left( H_{i,t}^{-1}a \right)' (V_{i,0} - V_{i,t}) H_{i,t}^{-1}a - \frac{1}{2} a'V_i H_{i,t}^{-1}a \right\} \hspace{1cm} (11)$$

$$= \frac{1}{\delta} \exp \left\{ \gamma g + \frac{\beta_i \gamma v_c}{1 - v_{b,i} v_c} + \frac{\gamma^2 (v_{b,i} v_c - v_{b,i} - v_c)^2}{2(1 - v_{b,i} v_c)^2} - \frac{\gamma^2 v_{b,i} v_c}{2(1 - v_{b,i} v_c)^2} \right\} \hspace{1cm} (12)$$

$$= \frac{1}{\delta} \exp \left\{ \gamma g + \frac{\beta_i \gamma v_c}{1 - v_{b,i} v_c} + \frac{\gamma^2 v_{b,i} v_c^2 - 2v_{b,i} v_c}{2(1 - v_{b,i} v_c)^2} \right\} \hspace{1cm} (13)$$

This follows from the fact that the expected return can be expressed as:

$$E[R_{i,t+1}] = E[D_{i,t+1}/P_{i,t}] = E[E_t(D_{i,t+1})/P_{i,t}]$$

The necessary condition for these conditional expectations to exist is given by the restriction (10) on the determinant of $H_{i,0}$. Intuitively, since these expectations involve the exponent of the product of two normal random variables, one must ensure that the variance of these random variables is small enough for the integral to converge.\(^2\)

While the expression in (11) is quite complicated, note that when there is no information about $c_{t+1}$ at time $t$ and no uncertainty about $\beta$, the firm’s expected return reduces to the familiar expression in the Lucas tree economy:

$$E[R_{i,t+1}] = \frac{1}{\delta} \exp \left\{ \gamma g + \gamma \beta_i v_c - \frac{1}{2} \gamma^2 v_c \right\} \hspace{1cm} (14)$$

The effect of allowing for uncertainty in $\beta$ changes the expression of the firm’s expected return to the following:

$$E[R_{i,t+1}] = \frac{1}{\delta} \exp \left\{ \gamma g + (\gamma \beta_i - \frac{1}{2} \gamma^2) \frac{v_c}{1 - v_{b,i} v_c} \right\} \hspace{1cm} (15)$$

Hence, intuitively, allowing for uncertainty about $\beta$ has the same effect as scaling the unconditional variance of consumption growth $v_c$ by $\frac{1}{1 - v_{b,i} v_c}$.

While we consider a static model for ease of exposition and tractability, the expression can be readily extended in a multi-period setup.\(^3\) Moreover, the expression in (11) is quite general - it holds under a broad range of information structures. To develop some intuition for the general expression, we consider special cases in order of increasing complexity.

\(^2\)If we relax the assumption that $\beta$ and $c_{t+1}$ are independent, then the restriction in (10) is given by:

$$(1 - v_{cb})^2 - v_c v_{b,i} > 0$$

Specifically, this implies that $\beta$ and $c_{t+1}$ cannot be positively correlated.

\(^3\)For instance, see Ang and Liu (2007) for a dynamic version of the basic setup we start with in continuous time. They calculate the price of an infinite stream of dividends and show that a dynamic Gordon growth model arises in which expected returns are determined by the covariance of dividend growth with the pricing kernel. Since we assume conditional normality, their model can be extended to allow for learning about risk-factors and cash-flow betas as in our model.
Case 1: No uncertainty about $\beta$

We begin with the most familiar setting and suppose there is no uncertainty about beta. This corresponds to the majority of models that embed learning in standard asset pricing setups. The representative investor’s information at date $t$ is only about consumption growth. Under the assumption of normality, the posterior variance of consumption growth can always be expressed as:

$$v_{c,t} = v_c(1 - \lambda_{c,t}) \quad \text{where} \quad \lambda_{c,t} \in [0, 1]$$

(16)

The coefficient $\lambda_{c,t}$ is a measure of the quality of systematic information about $c_{t+1}$ available to the representative investor at time $t$. In this case, the expected return on firm $i$’s cash-flows is given by:

$$E[R_{i,t+1}] = \frac{1}{\delta} \exp \{ \gamma g + \beta_i \gamma v_{c,t} + \gamma^2 (\frac{1}{2} v_c - v_{c,t}) \}$$

(17)

$$= \frac{1}{\delta} \exp \{ \gamma g + \beta_i \gamma v_c(1 - \lambda_{c,t}) - \gamma^2 v_c(\frac{1}{2} - \lambda_{c,t}) \}$$

(18)

Intuitively, higher information quality leads to a lower posterior variance in consumption growth (i.e. lower $v_{c,t}$). This implies that information quality has two potentially off-setting effects on the expected return:

- **The beta effect**: An increase in information quality decreases the cost of capital when $\beta_i > 0$, but increases the cost of capital when $\beta_i < 0$. This can be interpreted as an attenuation of the conditional covariance between the firm’s cash-flows and the aggregate consumption process. This leads to a decrease in the firm’s expected return when $\beta_i$ is positive, but an increase in the cost of capital when $\beta_i$ is negative.

- **The convexity effect**: An increase in systematic information quality decreases the posterior variance in consumption growth ($v_{c,t}$). Since the price of the risk-free asset is convex in aggregate uncertainty, an increase in systematic IQ leads to a decrease in the price of the risk-free bond and so an increase in the risk-free rate. This becomes apparent from the expression of the gross risk-free rate in this case, which is given by:

$$R_{f,t} = \frac{1}{\delta} \exp \{ \gamma g + \gamma^2 (\frac{1}{2} v_c - v_{c,t}) \}$$

(19)

Case 2: No learning about $c_{t+1}$

In this case, suppose that there is uncertainty about $\beta_{i,t}$ and $c_{t+1}$, but these variables are independent and there is no information available at time $t$ about $c_{t+1}$ (i.e. $v_c = v_{c,t}$). This captures the notion that the public information about a firm is largely firm-specific and so affects beliefs about $\beta_{i,t}$, but not about systematic risk factors (i.e. $c_{t+1}$). Again, since the prior beliefs about $\beta_{i,t}$ are normally distributed, we can express the posterior variance about $\beta_{i,t}$ as:

$$v_{b,i,t} = v_{b,i}(1 - \lambda_{b,i,t}) \quad \text{where} \quad \lambda_{b,i,t} \in [0, 1]$$

(20)
where \( \lambda_{b,i,t} \) is a measure of the quality of firm-specific information about firm \( i \)'s beta available at time \( t \). The expected return expression in (11) reduces to the following:

\[
E[R_{i,t+1}] = \frac{1}{\delta} \exp \left\{ \gamma g + \frac{\beta_i \gamma v_c}{1-v_{b,i,t} v_c} + \frac{\gamma^2 v_c^2 (v_{b,i,t} - v_{b,i,t})}{2(1-v_{b,i,t} v_c)^2} - \frac{\gamma^2 v_c}{2(1-v_{b,i,t} v_c)^2} \right\}
\]

Again, the effect of increasing information quality is two-fold. The “beta” effect of an increase in information quality about a firm’s cash-flow beta has two potentially offsetting effects. The first effect is to increase the risk of these cash-flows, which leads to a lower option value in the price (since the price is convex in cash-flow growth, which leads to a lower option value in the price (since the price is convex in cash-flow uncertainty), and therefore a higher expected return - this is reflected by the \(- \frac{\gamma^2 v_c}{2(1-v_{b,i,t} v_c)^2}\) term. In the case where there is no learning about aggregate consumption, the second effect dominates, and the overall convexity effect of firm-specific information quality is to increase the expected return.

The “convexity” effect of an increase in firm-specific information quality is somewhat more subtle than in the case of systematic information quality. Increasing the information quality about a firm’s cash-flow beta has two potentially offsetting effects. The first effect is to increase the risk of these cash-flows, which leads to a decrease in the expected returns - this is captured by the \(- \frac{\gamma^2 v_c}{2(1-v_{b,i,t} v_c)^2}\) term in (21). In particular, the derivative of this term with respect to \( \lambda_{b,i,t} \) is negative. The second effect is to decrease the volatility in cash-flow growth, which leads to a lower option value in the price (since the price is convex in the cash-flow uncertainty), and therefore a higher expected return - this is reflected by the \( \frac{\gamma^2 v_c}{2(1-v_{b,i,t} v_c)^2}\) term. In the case where there is no learning about aggregate consumption, the second effect dominates, and the overall convexity effect of firm-specific information quality is to increase the expected return.

The intuition for the “convexity” effect is similar to that in the learning model of Pastor and Veronesi (2003). In their model, an increase in the uncertainty about idiosyncratic profitability leads to higher market to book ratios, because of the optionality of these firms. Over time, investors learn more about the profitability and the option value decreases, which leads to lower market to book ratios. In our model, the convexity is with respect to the systematic component of cash-flows - a decrease in the volatility of cash-flows leads to lower optionality in prices, which leads to an increase in expected returns.

The General Case: Systematic and Firm-specific Information Quality

Finally, suppose that there is uncertainty and learning about both consumption growth \( c_{t+1} \) and \( \beta_{i,t} \), but that these variables are independent. In this case, we can distinguish between firm-specific information quality and systematic information quality. Again, since \( \beta \) and \( c \) are normally distributed, we can express the conditional variances as

\[
v_{b,i,t} = v_{b,i}(1-\lambda_{b,i,t}) \quad \text{and} \quad v_{c,t} = v_c(1-\lambda_{c,t}) \quad \text{where} \quad \lambda_{b,i,t} \in [0,1] \quad \text{and} \quad \lambda_{c,t} \in [0,1]
\]

As before, \( \lambda_{c,t} \) is a measure of systematic information quality about consumption growth, while \( \lambda_{b,i,t} \) is firm-specific information quality about the firm’s beta. The expected return expression in (11) reduces to the following in this case:

\[
E[R_{i,t+1}] = \frac{1}{\delta} \exp \left\{ \gamma g + \frac{\beta_i \gamma v_c}{1-v_{b,i,t} v_c} + \frac{\gamma^2 v_c^2 (v_{b,i,t} - v_{b,i,t}) + (v_{b,i,t} - v_{b,i,t})}{2(1-v_{b,i,t} v_c)^2} - \frac{\gamma^2 v_c}{2(1-v_{b,i,t} v_c)^2} \right\}
\]

Again, the effect of increasing information quality about a firm’s cash-flow beta has two potentially offsetting effects. The first effect is to increase the risk of these cash-flows, which leads to a lower option value in the price (since the price is convex in cash-flow growth, which leads to a lower option value in the price (since the price is convex in cash-flow uncertainty), and therefore a higher expected return - this is reflected by the \(- \frac{\gamma^2 v_c}{2(1-v_{b,i,t} v_c)^2}\) term. In the case where there is no learning about aggregate consumption, the second effect dominates, and the overall convexity effect of firm-specific information quality is to increase the expected return.
In this case, the “beta” effect is similar to that in the earlier cases. Specifically, an increase in either systematic or firm-specific information quality leads to a decrease in expected returns when $\beta_i$ is positive, but an increase in expected returns when $\beta_i$ is negative.

In contrast to the earlier cases, however, now there are two sources of the “convexity” effect since increasing $\lambda_{c,t}$ and $\lambda_{b,i}$ both reduce the uncertainty about the cash-flows. It is easy to verify that the convexity effect of an increase in systematic information quality (i.e. $\lambda_{c,t}$) is to increase the expected return as before. However, an increase in the firm-specific information quality $\lambda_{b,i,t}$ only increases the expected return when the following condition is satisfied:

$$
(1 - 2\lambda_{c,t}) - (1 - \lambda_{c,t})^2 v_{b,i} v_{c} > 0 \iff 0 \leq \lambda_{c,t} < \frac{\sqrt{1 - v_{b,i} v_{c} - (1 - v_{b,i} v_{c})}}{v_{b,i} v_{c}}
$$

(26)

and decreases expected returns otherwise. As in Case 2 above, increasing firm-specific information quality has two potentially offsetting effects. The first effect is to reduce the risk of the cash-flows and hence to decrease expected returns - again, this is captured by the term in (24). Furthermore note that in this case, the derivative of this term with respect to $\lambda_{b,i,t}$ depends on the systematic information quality (i.e. $\lambda_{c,t}$) and can be shown to be decreasing in $\lambda_{c,t}$. The second effect, as before, is to decrease cash-flow volatility which decreases the option value in the price and so increases the expected return. As a result, the overall convexity effect for an increase in $\lambda_{b,i,t}$ is only positive when $\lambda_{c,t}$ is small enough.

In a sense, the general case of our model extends the intuition in the model of Pastor and Veronesi (2003). As in Case 2 discussed above, they find that learning about firm-specific cash-flow growth (or profitability) leads to a decrease in prices due to the fact that prices are convex in profitability. In our model, we allow investors to learn not only about the cash-flows of a specific firm, but also about aggregate risk-factors in the economy. Moreover, we find that the effect of learning about the cash-flows on expected returns depends crucially on what is being learned about the risk-factor.

### 2.3 Partial Equilibrium Setup with an Exogenous Pricing Kernel

The consumption CAPM model of the last section is very useful in developing intuition for our results. However, the consumption CAPM has had limited success in empirically explaining the cross-section of stock returns. In this section, we extend the results of the model to a partial equilibrium, linear factor pricing setup which allows us to test the empirical predictions of the model using the Fama French Carhart four factor model.

In particular, we begin with an exogenous, $n$ factor pricing kernel, or stochastic discount factor, given by:

$$
M_{t+1} = M_t \exp \left\{ -r_f - \frac{1}{2} t' V_m t - t' m_{t+1} \right\} \text{ where } m_{t+1} \sim N(0, V_m)
$$

(27)

where $m_{t+1}$ is a $n \times 1$ vector of normal random variables, $V_m$ is the $n \times n$ covariance matrix of the risk-factors, $t$ is a vector of ones, and $r_f$ denotes the unconditional log risk-free rate.
Note that the existence of the pricing kernel relies only on the assumption that there is no arbitrage in the economy. This makes the current setup quite general - in particular, this representation can capture a variety of pricing models including consumption-based models and factor-based models (like the CAPM). As before, suppose that firm \( i \)'s cash-flows in the next period are given by:

\[
D_{i,t+1} = D_{i,t} \exp \left\{ \mu_{i,t+1} + d_{i,t+1} - \frac{1}{2} v_d \right\} \quad \text{where} \quad d_{i,t+1} \sim N(0, v_d)
\]  

(28)

and \( d_{i,t+1} \) is assumed to be purely idiosyncratic. The mean growth rate in cash-flows \( \mu_{i,t+1} \) has a predictable component \( \bar{\mu}_{i,t} \) and can always be projected on the pricing kernel as:

\[
\mu_{i,t+1} = \bar{\mu}_{i,t} + \beta_{i,t+1}' m_{t+1} + e_{i,t+1} \quad \text{where} \quad e_{i,t+1} \sim N(0, v_{e,i})
\]  

(29)

and \( e_{i,t+1} \) is assumed to be independent across firms. The risk-factor loadings of firm \( i \)'s cash-flows are now given by a vector of betas, \( \beta_{i,t+1} \), and there may be uncertainty about this vector. Specifically, suppose that the prior distribution on these \( \beta \)'s is given by

\[
\beta_{i,t+1} \sim N(\beta_i, V_{b,i})
\]  

(30)

where \( V_{b,i} \) is an \( n \times n \) covariance matrix of the \( \beta \)'s.

Conditional on the information available at date \( t \), the price of firm \( i \)'s cash-flows is given by:

\[
P_{i,t} = E_t \left[ \frac{M_{t+1}}{M_t} D_{i,t+1} \right]
\]  

(31)

To evaluate the above expression for price, and hence the expected return, we again begin by denoting the stack of \( m_{t+1} \) and \( \beta_{i,t+1} \) by \( X_{i,t+1} \). The unconditional distribution of \( X_{i,t+1} \) is given by

\[
X_{i,t+1} \sim N(x_{i,0}, V_{i,0})
\]  

(32)

and the conditional distribution of \( X_{i,t+1} \) at date \( t \) is given by

\[
X_{i,t+1} | Z_t \sim N(x_{i,t}, V_{i,t}).
\]  

(33)

As before, we assume that the correlation between \( \beta_{i,t+1} \) and the risk-factors \( m_{t+1} \) is zero.

Again, we denote \( G \equiv \frac{1}{2} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \) and \( a \equiv - \begin{pmatrix} t \\ 0 \end{pmatrix} \), and let \( H_{i,t} = (I - 2GV_{i,t}) \). Given this notation, one can express the expected return on firm \( i \)'s cash-flows as in the following result.

**Proposition 2** Suppose that firm \( i \)'s cash-flows are given by (28) - (29), and let \( H_{i,t} = (I - 2GV_{i,t}) \). If the determinant of \( H_{i,0} \) is positive i.e.

\[
|H_{i,0}| = |I - V_m V_{b,i}| > 0
\]  

(34)

then the expected return on firm \( i \)'s cash-flows is given by:

\[
E[R_{i,t+1}] = \exp \left\{ r_f + \frac{1}{2} V_m t - x_{i,0}' H_{i,t}^{-1} a + \frac{1}{2} (H_{i,t}^{-1} a)' (V_{i,0} - V_{i,t}) H_{i,t}^{-1} a - \frac{1}{2} a' V_{i,t} H_{i,t}^{-1} a \right\}
\]

\[
= \exp \left\{ r_f + \frac{1}{2} V_m t + \beta_i' V_{m,t} h_{i,t} + \frac{1}{2} h_{i,t}' (V_m - V_{m,t}) + V_{m,t}' (V_{b,i} - V_{b,i,t}) V_{m,t} \right\} h_{i,t} - \frac{1}{2} V_{m,t} h_{i,t}
\]  

(35)
where

\[ H_{i,t}^{-1}a = -\left( \frac{(I - V_{b,i,t}V_{m,t})^{-1} \lambda_t}{V_{m,t}(I - V_{b,i,t}V_{m,t})^{-1} \lambda_t} \right) \equiv \left( \begin{array}{c} h_{i,t} \\ V_{m,t}h_{i,t} \end{array} \right) \]  

(36)

Note that the expression for expected returns in Proposition 2 is similar to, but generalizes, the expression in Proposition 1. The link between the two expressions is more obvious when we consider the special case of a single factor pricing kernel. In this case, the expected return on firm \( i \)'s cash-flows is given by:

\[
E[R_{i,t+1}] = \exp \left\{ r_f + \frac{1}{2} V_m + \frac{\beta_i V_{m,t}^2}{1 - V_{b,i,t}V_{m,t}} + \frac{1}{2} \frac{V_{b,i,t}V_{n,t}^2 - 2V_{m,t} + V_m}{(1 - V_{b,i,t}V_{m,t})^2} \right\}
\]

Moreover, since \( \beta_{i,t+1} \) and \( m_{t+1} \) are normally distributed, their variances conditional on date \( t \) information can be represented as:

\[
V_{m,t} = V_m(I - \Lambda_{m,t}) \quad \text{and} \quad V_{b,i,t} = V_{b,i}(I - \Lambda_{b,i,t})
\]

(37)

where \( \Lambda_{m,t} \) and \( \Lambda_{b,i,t} \) are now \( n \times n \) matrices with eigenvalues less than 1. This implies that the expected returns can be expressed as

\[
E[R_{i,t+1}] = \exp \left\{ r_f + \frac{1}{2} V_m t + \beta_i V_m(I - \Lambda_{m,t})h_{i,t} + \frac{1}{2} h_{i,t} V_m \Lambda_{m,t} h_{i,t} + \frac{1}{2} \beta_i V_m(I - \Lambda_{m,t})h_{i,t} \right\}
\]

(38)

### 2.4 Estimates of \( \beta \) and Empirical Predictions

It is unlikely that the economist perfectly observes the pricing kernel or the conditional \( \beta \). Instead, suppose that we observe a vector of risk factors \( f_{t+1} \) which is a noisy version of the pricing kernel:

\[ f_{t+1} = m_{t+1} + w_{t+1} \quad \text{where} \quad w_{t+1} \sim N(0, V_w) \]

(39)

For simplicity, we assume the vector of errors, \( w_{t+1} \), is independent of \( m_{t+1} \) and each other (i.e. \( V_w \) is diagonal). Moreover, suppose that the conditional variances of \( \beta \) and \( m_{t+1} \) are denoted by (37) above, where \( \Lambda_{b,i,t} \) and \( \Lambda_{m,t} \) measure the quality of information about \( \beta_{i,t+1} \) and \( m_{t+1} \) respectively. This implies that the economist's estimate of the conditional beta with respect to factor \( f_{t+1} \) is given by:

\[
\hat{\beta}_i = E \left[ \frac{\text{cov}(f_{t+1}, \mu_{i,t+1}|I_t)}{\text{var}(f_{t+1}|I_t)} \right] = \beta_i(I - \Lambda_{m,t})V_m((I - \Lambda_{m,t})V_m + V_w)^{-1} \propto \hat{\beta}_i
\]

(40)

In terms of the observed \( \hat{\beta}_i \), the expected return is given by:

\[
E[R_{i,t+1}] = \exp \left\{ r_f + \frac{1}{2} V_m t + \hat{\beta}_i((I - \Lambda_{m,t})V_m + V_w)h_{i,t} + \frac{1}{2} h_{i,t} V_m \Lambda_{m,t} h_{i,t} + \frac{1}{2} \beta_i V_m(I - \Lambda_{m,t})h_{i,t} \right\}
\]

(41)

As a result, the "beta" and "convexity" effects translate immediately to empirically testable predictions, which are summarized in the following hypothesis.
Hypothesis 1 Suppose the pricing kernel is observed with error as in (39). Expected returns are:

(a) decreasing in the interaction between observed beta (i.e. \( \hat{\beta}_i \)) and information quality about the pricing kernel (i.e. \( \Lambda_{m,t} \))

(b) decreasing in the interaction between observed beta (i.e. \( \hat{\beta}_i \)) and information quality about the firm-specific beta (i.e. \( \Lambda_{b,i,t} \))

(c) increasing in systematic information quality about the pricing kernel (i.e. \( \Lambda_{m,t} \))

(d) increasing in firm-specific information quality (i.e. \( \Lambda_{b,i,t} \)) when systematic information quality is low enough

3 Empirical Analysis

3.1 Data Description and Variable Measurement

The first challenge in testing our two hypotheses is that they are about betas conditional on information available to investors. The standard approach to testing conditional beta models has been to “condition down” to a set of relevant state variables (e.g. Jagannathan and Wang (1996), Lettau and Ludvigson (2001)). An alternative approach, proposed by Lewellen and Nagel (2006), avoids the problem of having to define the relevant set of conditioning information by using short-window regressions to directly estimate the conditional betas. We follow their approach and estimate conditional betas from quarterly regressions (using daily observations) of returns on the three Fama and French (1993) factors and Carhart’s (1997) momentum factor. These factors are the excess market return (mktrf), the small minus big (smb) portfolio, the high minus low (hml) portfolio and the up minus down (umd) portfolio. Since we estimate the factor loadings using daily data, we use Dimson (1979) betas with three day lags to adjust for asynchronous trading effects that may arise at this frequency. We estimate the following specification for each firm:

\[
r_{j,t} - r_{f,t} = \sum_{s=0}^{3} b_{mktrf,s}r_{mktrf,t-s} + b_{smb,s}r_{smb,t-s} + b_{hml,s}r_{hml,t-s} + b_{umd,s}r_{umd,t-s} + u_{j,t}
\] (42)

and calculate the estimates for betas as:

\[
\hat{\beta}_{mktrf,j,t} = \sum_{s=0}^{3} b_{mktrf,t-s}, \quad \hat{\beta}_{smb,j,t} = \sum_{s=0}^{3} b_{smb,s}, \quad \hat{\beta}_{hml,j,t} = \sum_{s=0}^{3} b_{hml,s}, \quad \hat{\beta}_{umd,j,t} = \sum_{s=0}^{3} b_{umd,s}
\] (43)

We trim each cross-section on each of the beta estimates at the 1st and 99th percentiles to avoid over-weighting outliers.
The second challenge in testing our predictions is identifying firm-specific measures of the level of information quality. We rely on three complementary proxies, the first two constructed at a monthly frequency and the third calculated annually. The reason for using proxies from very different sources is to ensure that our empirical results are robust to some degree of noise and misspecification. Moreover, information quality is a complex construct, so using multiple measures helps to ensure the robustness of our findings. Our first proxy for firm-specific information quality is based on the mean forecast error across analyst estimates of earnings per share. This captures the notion that when the quality of a firm’s public information is higher, analysts should be in greater agreement. Each month, we compute the mean absolute forecast error of annual earnings and scale it by the median estimate. We use the exponential of the negative mean forecast error, denoted by \( \lambda_{mfe} \), as our first proxy for firm-specific information quality.\(^4\) The transformation ensures that \( \lambda_{mfe} \) lies between 0 and 1 (which corresponds to the theoretical range of \( \lambda \) in our model), is increasing in information quality, and has more desirable distributional properties.

Our second proxy for firm-specific information quality is based on the quality of information available to financial analysts which we measure based on the approach developed in Barron et al. (1998) (hereafter BKLS). Their approach estimates both the quality of public information (i.e., the quality of the information that is common to all analysts following the firm) and the quality of private information (i.e., the quality of information that is available to a particular analyst as a result of his information search efforts). The BKLS setup assumes that there are \( N \) financial analysts that issue earnings forecasts and that each analyst has two pieces of information: (1) common information that is available to all analysts (denoted as \( h \)) and (2) information that is available only to the specific analyst (denoted as \( s \)). The quality of each piece of information is captured by its precision. BKLS show that forecast dispersion and the error in the mean forecast are functions of (1) the quality of public information, (2) the quality of private information and (3) the number of analysts issuing a forecast. They then show that the quality of public and private information can be recovered by the following equations, respectively:

\[
h = \frac{SE - D/N}{[(1 - \frac{1}{N})D + SE]^2}, \quad s = \frac{D}{[(1 - \frac{1}{N})D + SE]^2}
\]

where \( SE \) is the expected squared error of the mean forecast, \( D \) is the expected forecast dispersion, and \( N \) is the number of analysts issuing a forecast. We use the exponential of the negative quality of public information (\( h \)), denoted by \( \lambda_{bkls} \), as our second proxy for firm-specific information quality.

Our third proxy for firm-specific information quality is accruals quality.\(^5\) This proxy focuses on an important subset of the information available about a firm and has been shown

\(^4\)We restrict attention to annual earnings forecasts to maximize the sample size and scale by the absolute median estimate to control for the mechanical relationship between size and absolute forecast errors.

\(^5\)There are other measures used in the literature (e.g. share price volatility used by Lang and Lundholm (1993) and Leuz and Verrecchia (2000)) and probability of informed trade used by Easley et al. (1996)) that could be used as proxies for information quality. While we anticipate the results from using these to be qualitatively similar, we believe that our proxies are less model dependent and so have a tighter link to information quality. In particular, both price volatility and PIN are potentially affected by liquidity and other micro-structure effects, which is unlikely to be a problem with analyst forecasts or accruals quality.
to be an important proxy for information quality in the accounting literature.\footnote{Ng (2008) uses the accruals quality measure of Dechow and Dichev (2002) as a measure of the “reliability” of earnings, which he argues is an important component of information quality.} Our measure of accruals quality is based on the Dechow and Dichev (2002) model that relates current accruals to lagged, contemporaneous, and next period’s operating cash flows as follows:

$$\frac{TCA_{j,t}}{Assets_{j,t}} = \gamma_0 + \gamma_1 \frac{CFO_{j,t-1}}{Assets_{j,t}} + \gamma_2 \frac{CFO_{j,t}}{Assets_{j,t}} + \gamma_3 \frac{CFO_{j,t+1}}{Assets_{j,t}} + \varepsilon_{j,t}$$ \hspace{1cm} (45)$$

where $TCA_{j,t}$ is the current accruals of firm $j$ in year $t$; $Assets_{j,t}$ is the average total assets of firm $j$ for years $t$ and $t-1$; and $CFO_{j,t}$ is the cash-flow from operations for firm $j$ in year $t$. For each firm in our sample, we estimate (45) using a rolling 10-year window. The measure of accruals quality ($AQ$) is given by the standard deviation of the 10 residuals $\varepsilon_{j,t}$ for each estimation period. As with the other proxies, we use the exponent of the negative, denoted by $\lambda_{AQ}$, as our proxy for firm-specific information quality. The main drawback of using $\lambda_{AQ}$ is that, unlike the other proxies, it is only available at the annual frequency.

In addition to idiosyncratic IQ, we require a proxy for systematic information quality in the economy. We capture this construct using the VIX Volatility Index provided by the Chicago Board Options Exchange (CBOE). The VIX is constructed from eight S&P 100 index call and put contracts as the implied volatility of a synthetic at-the-money option on the S&P 100 with a one-month maturity.\footnote{The CBOE makes to versions of its Volatility Index available. The first is based on the S&P 100 and is available beginning in January of 1986 and is denoted by the ticker VXO. The second is based on the S&P 500 and uses a broader range of index options than the VXO and is denoted by the ticker VIX. The CBOE started calculating the VIX in September of 2003 and backfilled this measure to January of 1990. Ang et al. (2006) report that the two measures are very highly correlated (98%) and they use the VXO because of its longer time series. We also use the VXO to maximize our sample period.} This measure is available daily, so we use the closing value at the end of each month (which corresponds to the frequency of our tests) and transform it according to $\lambda_{VIX} = \exp \{-VIX/10\}$, since this yields a proxy with a comparable distribution to the firm-specific variables.

### 3.2 Empirical Specifications

The relationship between expected returns, beta and information quality given by (41) is highly non-linear, and difficult to estimate robustly. To test the basic relationships outlined in Hypothesis 1, while maintaining a relatively low number of estimated parameters, we consider two possible specifications. The first assumes that the two types of information quality enter additively, and in particular, we estimate the following regression:

$$E[R_{i,t} - R_{f,t}] = d_0 \lambda_{i,t} + m_0 \lambda_{VIX,t} + \left( a_1 \tilde{\beta}_{hml,i,t} + a_2 \tilde{\beta}_{mktrf,i,t} + a_3 \tilde{\beta}_{smb,i,t} + a_4 \tilde{\beta}_{umd,i,t} \right) \times \left( 1 + d_1 \lambda_{i,t} + m_1 \lambda_{VIX,t} \right)$$ \hspace{1cm} (46)$$

where $\lambda_{i,t}$ denotes one of the three firm-specific IQ proxies and $\lambda_{VIX,t}$ denotes our proxy for systematic IQ. Hypothesis 1 predicts that the coefficients $d_1$ and $m_1$ are negative, and the
coefficient $m_0$ is positive. The sign of the coefficient $d_0$ can either be positive or negative depending on whether the systematic information quality is large enough, and is therefore empirically indeterminate.

Our second specification assumes that the two types of information quality enter multiplicatively, and we estimate the following regression:

$$E[R_{i,t} - R_{f,t}] = id_0\lambda_{i,t} \times \lambda_{VIX,t} + \left( a_1\hat{\beta}_{hml,i,t} + a_2\hat{\beta}_{mktrf,i,t} + a_3\hat{\beta}_{smb,i,t} + a_4\hat{\beta}_{umd,i,t} \right) \times \left( 1 + id_1\lambda_{i,t} \times \lambda_{VIX,t} \right)$$

(47)

Our model predicts that the coefficient $id_0$ is positive and $id_1$ is negative. We use iterated generalized method of moments (GMM) to estimate the above non-linear specifications.\textsuperscript{8} The iterated GMM procedure weights the moment conditions using the sample covariance matrix estimated from the last iteration of parameter estimates, and re-iterates until the parameter estimates converge. A GMM estimation allows us to directly test the non-monotonic relationship and to impose the restriction that the effect of systematic and firm-specific information quality ($\hat{\lambda}_{VIX,t}$ and $\lambda_{i,t}$) is the same across all $\hat{\beta}$'s. This increases the power of the test and makes our empirical specification more closely aligned to the theoretical prediction in Hypothesis 1.

3.3 Empirical Results

Table 1 presents descriptive statistics for the variables in our sample. The first panel reports statistics related to the distribution of the firm-specific information quality proxies ($\lambda_{mfe}$, $\lambda_{bkls}$, and $\lambda_{AQ}$), the systematic information quality proxy ($\lambda_{VIX}$), factor loadings (betas), returns, and the control variables. The data spans the period 1984 - 2006, and all the variables are measured monthly, except the information quality proxy $\lambda_{AQ}$ which is measured at an annual frequency.\textsuperscript{9} The mean and median $\hat{\beta}_{mktrf}$ estimates are close to 1, and the corresponding estimates for the other betas suggests that the sample is tilted towards larger firms. This is partly due to the fact that we require firms to appear in IBES and have more than one analyst following them.

Tables 2 - 4 presents the results of our firm-level tests of Hypothesis 1 based on (46) and (47) using $\lambda_{mfe}$, $\lambda_{bkls}$, and $\lambda_{AQ}$ respectively. As a benchmark, we report the coefficients of returns on the estimates for $\hat{\beta}$. Consistent with prior empirical studies, we find that the in-sample risk premia on $\hat{\beta}_{hml}$ and $\hat{\beta}_{smb}$ are negative for the firms in our sample.\textsuperscript{10} This is likely to be due to the fact that our sample is tilted towards large firms and our data is from 1984 onwards. We then report the estimates of the specifications in (46) and (47) which

\textsuperscript{8}See Hansen (1982) and Hansen and Singleton (1982) for a discussion of the GMM procedure, and Ferson and Foerster (1994) for a discussion of iterated GMM.

\textsuperscript{9}Note that the betas and control variables are also measured annually when $\lambda_{AQ}$ is used as the proxy for information quality.

\textsuperscript{10}Jagannathan and Wang (1996) and Lettau and Ludvigson (2001) also document negative estimates of risk premia for various risk factors in their samples.
include all four factor $\hat{\beta}$'s. The results provide statistically significant evidence consistent with our predictions, especially for the monthly specifications. In particular, when using either $\lambda_{mfe}$ or $\lambda_{bkls}$ as the proxy for firm-specific information quality, we find that both the additive (i.e. (46)) and multiplicative specifications (i.e. (47)) yield statistically significant evidence consistent with the beta effect for firm-specific and systematic information quality (i.e. $d_1$, $m_1$, and $id_1$ are negative and significant), and with the convexity effect for systematic information quality (i.e. $m_0$ and $id_0$ are positive and significant). Moreover, we find that the convexity effect of firm-specific information for both monthly specifications is also positive (i.e. $d_0$ is positive and significant), suggesting that the systematic information quality over this horizon is relatively small.

The evidence from the annual specification, which uses $\lambda_{AQ}$ as the proxy for firm-specific information quality, is less statistically significant but still consistent with the predictions of the model. This is because we are now restricted to annual observations instead of monthly observations, which makes the problem of estimating expected returns particularly difficult. However, the annual specification yields an interesting effect different from that in the monthly specifications. In particular, the coefficient for the firm-specific convexity effect in the additive specification (i.e. (46)) is significantly negative for the annual proxy $\lambda_{AQ}$ (i.e. $d_0$ is negative and significant). This is consistent with the idea that over an annual horizon, investors learn more about the systematic risk-factor (i.e. systematic information quality is relatively large for this period) and consequently the convexity effect is negative as predicted by the model.

Although the economic significance of these effects is somewhat difficult to quantify as a result of the non-linearity of the specifications, it is an important exercise. Our coefficient estimates suggest that, in the additive monthly specifications, the direct effects of a one standard deviation increase in systematic IQ and firm-specific IQ correspond to increases in expected returns of 80 basis points and 30 basis points respectively. Assuming an average expected return of 2% per month, a one standard deviation increase in systematic and firm-specific IQ correspond to decreases of 32 basis points and 30 basis points respectively. Thus, the effects of information quality on expected returns identified by our model appear to be small, but economically significant.

4 Conclusions

We study the relationship between information quality and cost of capital in standard asset pricing models where systematic risk determines the cost of capital. Our premise is that if information quality affects cost of capital, it must be through systematic risk. We allow investors to learn about aggregate risk-factors and about firm-specific factor loadings (or betas) and study the effects of both systematic and firm-specific information quality on expected returns.

Unlike single firm models which generally predict a negative relationship between the two, our model predicts a non-monotonic relationship between expected returns and information quality. However somewhat surprisingly, the effect of both types of information quality
are similar. We show that both systematic and firm-specific information quality decrease expected returns for positive beta stocks but increase expected returns for negative beta stocks - this is what we call the “beta” effect of information quality. There is also a direct, “convexity” effect of information quality which may increase or decrease expected returns. We show that the convexity effect of systematic information quality unambiguously increases expected returns, by lowering aggregate uncertainty and hence increasing the risk-free rate. However, the convexity effect for firm-specific information quality is positive only when systematic IQ is low, and negative otherwise.

As a result of these potentially offsetting effects, the overall effect of increasing information quality on expected returns is non-linear and is negative only for firms with sufficiently high betas. We test these non-linear predictions directly using realized stock returns, short horizon four factor betas as proxies for systematic risk, and three proxies for information quality based on analyst forecasts and accruals quality. We find evidence consistent with the predictions of our model. Moreover, we find evidence consistent with the notion that investors learn little about aggregate risk-factors over short horizons (i.e. one month), but substantially over longer horizons (i.e. one year).

Our objective is to provide a first step in analyzing the relationship between information quality and cost of capital in a standard asset pricing framework with multiple firms and multiple signals. While extremely simple, our model captures the notion that non-diversifiable, covariance risk is important when studying the effects of information on expected returns and we find that the empirical evidence is consistent with this prediction. Moreover, the setup lends itself to extension, and provides a benchmark for more sophisticated models of endogenous disclosure and information quality.
5 References


Clarkson, P. J. Guedes, R. Thompson (1996), ”On the Diversification, Observability, and Measure-


6 Tables

Table 1: Summary statistics for relevant variables. The proxies for information quality are given by \( \hat{\lambda}_{mfe} = \exp\{-10 \times \frac{mfe}{\text{abs}(mest)}\} \), \( \hat{\lambda}_{BKLS} = \exp\{-h\} \), and \( \hat{\lambda}_{AQ} = \exp\{-AQ\} \). Equal weighted averages of firm characteristics for information quality quintiles are also reported.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Std Dev</th>
<th>5th Pctl</th>
<th>95th Pctl</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{i,t} ) (in percent)</td>
<td>633797</td>
<td>2.114</td>
<td>3.267</td>
<td>23.272</td>
<td>-38.853</td>
<td>38.638</td>
</tr>
<tr>
<td>( \beta_{mktf} )</td>
<td>633797</td>
<td>1.076</td>
<td>0.998</td>
<td>9.132</td>
<td>-1.503</td>
<td>3.993</td>
</tr>
<tr>
<td>( \beta_{hml} )</td>
<td>633797</td>
<td>0.120</td>
<td>0.188</td>
<td>7.557</td>
<td>-4.809</td>
<td>4.747</td>
</tr>
<tr>
<td>( \beta_{smb} )</td>
<td>633797</td>
<td>0.686</td>
<td>0.525</td>
<td>11.811</td>
<td>-2.727</td>
<td>4.714</td>
</tr>
<tr>
<td>( \beta_{umd} )</td>
<td>633797</td>
<td>-0.059</td>
<td>-0.023</td>
<td>3.400</td>
<td>-3.713</td>
<td>3.569</td>
</tr>
<tr>
<td>mfe/\text{abs}(mest)</td>
<td>443484</td>
<td>0.473</td>
<td>0.090</td>
<td>3.314</td>
<td>0.007</td>
<td>1.458</td>
</tr>
<tr>
<td>BKLS h</td>
<td>589009</td>
<td>5.236</td>
<td>0.011</td>
<td>333.02</td>
<td>0.000</td>
<td>2.141</td>
</tr>
<tr>
<td>AQ</td>
<td>193432</td>
<td>0.461</td>
<td>0.436</td>
<td>0.219</td>
<td>0.143</td>
<td>0.861</td>
</tr>
<tr>
<td>spread/prc</td>
<td>517912</td>
<td>0.016</td>
<td>0.011</td>
<td>0.018</td>
<td>0.000</td>
<td>0.050</td>
</tr>
<tr>
<td>( \hat{\lambda}_{mfe} )</td>
<td>443484</td>
<td>0.412</td>
<td>0.407</td>
<td>0.329</td>
<td>0.000</td>
<td>0.932</td>
</tr>
<tr>
<td>( \hat{\lambda}_{bkls} )</td>
<td>589009</td>
<td>0.875</td>
<td>0.989</td>
<td>0.254</td>
<td>0.124</td>
<td>1.000</td>
</tr>
<tr>
<td>( \hat{\lambda}_{AQ} )</td>
<td>193432</td>
<td>0.646</td>
<td>0.646</td>
<td>0.136</td>
<td>0.423</td>
<td>0.867</td>
</tr>
<tr>
<td>( \hat{\lambda}_{spread} )</td>
<td>517912</td>
<td>0.418</td>
<td>0.329</td>
<td>0.336</td>
<td>0.006</td>
<td>0.965</td>
</tr>
<tr>
<td>VIX</td>
<td>383027</td>
<td>20.118</td>
<td>19.464</td>
<td>7.019</td>
<td>11.991</td>
<td>41.706</td>
</tr>
<tr>
<td>( \hat{\lambda}_{VIX} )</td>
<td>383027</td>
<td>0.161</td>
<td>0.143</td>
<td>0.085</td>
<td>0.016</td>
<td>0.301</td>
</tr>
</tbody>
</table>
Table 2: This table presents the results from the GMM estimation for the following specifications using the firm specific information quality based on mean forecast error ($\hat{\lambda}_{mfe}$). We use monthly observations of all the variables for a total of 258129 observations.

\[
E[R_{i,t} - R_{f,t}] = d_0\lambda_{i,t} + m_0\lambda_{VIX,t} + \left( a_1\beta_{hml,i,t} + a_2\beta_{mktrf,i,t} + a_3\beta_{smb,i,t} + a_4\beta_{umd,i,t} \right) (1 + d_1\lambda_{i,t} + m_1\lambda_{VIX,t})
\]

\[
E[R_{i,t} - R_{f,t}] = id_0\lambda_{i,t} \times \lambda_{VIX,t} + \left( a_1\beta_{hml,i,t} + a_2\beta_{mktrf,i,t} + a_3\beta_{smb,i,t} + a_4\beta_{umd,i,t} \right) (1 + id_1\lambda_{i,t} \times \lambda_{VIX,t})
\]

<table>
<thead>
<tr>
<th>$\hat{\lambda}_{mfe}$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$d_0$</th>
<th>$d_1$</th>
<th>$m_0$</th>
<th>$m_1$</th>
<th>$id_0$</th>
<th>$id_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>parms</td>
<td>-0.014</td>
<td>0.630</td>
<td>-0.253</td>
<td>0.545</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-0.670</td>
<td>20.870</td>
<td>-10.362</td>
<td>24.353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>0.196</td>
<td>0.076</td>
<td>-0.341</td>
<td>0.657</td>
<td>3.548</td>
<td>-0.778</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>-0.020</td>
<td>0.113</td>
<td>-0.718</td>
<td>0.971</td>
<td></td>
<td></td>
<td>11.772</td>
<td>-2.986</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-0.374</td>
<td>1.194</td>
<td>-10.291</td>
<td>15.067</td>
<td></td>
<td></td>
<td>44.864</td>
<td>-22.267</td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>0.256</td>
<td>-0.524</td>
<td>-0.893</td>
<td>1.037</td>
<td>1.682</td>
<td>-0.475</td>
<td>8.788</td>
<td>-3.108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>-0.112</td>
<td>0.361</td>
<td>-0.350</td>
<td>0.693</td>
<td></td>
<td></td>
<td>23.116</td>
<td>-3.898</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: This table presents the results from the GMM estimation for the following specifications using the firm specific information quality based on the BKLS measure of uncertainty ($\hat{\lambda}_{bkls}$). We use monthly observations of all the variables for a total of 222346 observations.

\[
E[R_{i,t} - R_{f,t}] = d_0 \lambda_{i,t} + m_0 \lambda_{VIX,t} + \left( \frac{a_1 \hat{\beta}_{hml,i,t} + a_2 \hat{\beta}_{mktrf,i,t} + a_3 \hat{\beta}_{smb,i,t} + a_4 \hat{\beta}_{umd,i,t}}{1 + d_1 \lambda_{i,t} + m_1 \lambda_{VIX,t}} \right) (1 + d_1 \lambda_{i,t} + m_1 \lambda_{VIX,t})
\]

\[
E[R_{i,t} - R_{f,t}] = id_0 \lambda_{i,t} \times \lambda_{VIX,t} + \left( \frac{a_1 \hat{\beta}_{hml,i,t} + a_2 \hat{\beta}_{mktrf,i,t} + a_3 \hat{\beta}_{smb,i,t} + a_4 \hat{\beta}_{umd,i,t}}{1 + id_1 \lambda_{i,t} \times \lambda_{VIX,t}} \right) (1 + id_1 \lambda_{i,t} \times \lambda_{VIX,t})
\]

<table>
<thead>
<tr>
<th>$\hat{\lambda}_{bkls}$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$d_0$</th>
<th>$d_1$</th>
<th>$m_0$</th>
<th>$m_1$</th>
<th>$id_0$</th>
<th>$id_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>parms</td>
<td>0.032</td>
<td>0.564</td>
<td>-0.218</td>
<td>0.449</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>1.786</td>
<td>21.474</td>
<td>-10.336</td>
<td>23.213</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>1.076</td>
<td>-2.109</td>
<td>-1.028</td>
<td>0.785</td>
<td>2.242</td>
<td>-0.972</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>0.035</td>
<td>0.055</td>
<td>-0.627</td>
<td>0.784</td>
<td>10.927</td>
<td>-2.921</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>0.777</td>
<td>0.675</td>
<td>-10.370</td>
<td>14.161</td>
<td>47.761</td>
<td>-20.589</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>1.345</td>
<td>-3.153</td>
<td>-1.259</td>
<td>0.190</td>
<td>0.971</td>
<td>-0.774</td>
<td>5.141</td>
<td>-1.848</td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>0.386</td>
<td>-0.694</td>
<td>-0.690</td>
<td>0.525</td>
<td></td>
<td></td>
<td>14.849</td>
<td>-4.844</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: This table presents the results from the GMM estimation for the following specifications using the firm specific information quality based on accruals quality ($\hat{\lambda}_{AQ}$). We use annual observations of all the variables for a total of 11568 observations.

\[
E[R_{i,t} - R_{f,t}] = d_0 \lambda_{i,t} + m_0 \lambda_{VIX,t} + \left( a_1 \hat{\beta}_{hml,i,t} + a_2 \hat{\beta}_{mktrf,i,t} + a_3 \hat{\beta}_{smb,i,t} + a_4 \hat{\beta}_{umd,i,t} \right) (1 + d_1 \lambda_{i,t} + m_1 \lambda_{VIX,t})
\]

\[
E[R_{i,t} - R_{f,t}] = id_0 \lambda_{i,t} \times \lambda_{VIX,t} + \left( a_1 \hat{\beta}_{hml,i,t} + a_2 \hat{\beta}_{mktrf,i,t} + a_3 \hat{\beta}_{smb,i,t} + a_4 \hat{\beta}_{umd,i,t} \right) (1 + id_1 \lambda_{i,t} \times \lambda_{VIX,t})
\]

<table>
<thead>
<tr>
<th>$\hat{\lambda}_{AQ}$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$d_0$</th>
<th>$d_1$</th>
<th>$m_0$</th>
<th>$m_1$</th>
<th>$id_0$</th>
<th>$id_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>parms</td>
<td>-0.201</td>
<td>1.011</td>
<td>-0.205</td>
<td>1.532</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-2.175</td>
<td>11.312</td>
<td>-1.867</td>
<td>14.068</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>-0.149</td>
<td>0.539</td>
<td>-0.208</td>
<td>1.993</td>
<td>1.359</td>
<td>-0.376</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-1.183</td>
<td>2.259</td>
<td>-1.359</td>
<td>3.623</td>
<td>6.401</td>
<td>-1.219</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>-0.374</td>
<td>0.226</td>
<td>-0.611</td>
<td>1.902</td>
<td>8.684</td>
<td>-1.591</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-2.343</td>
<td>0.951</td>
<td>-2.753</td>
<td>5.528</td>
<td>11.906</td>
<td>-2.513</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>-0.392</td>
<td>0.830</td>
<td>-0.491</td>
<td>1.518</td>
<td>-2.932</td>
<td>0.241</td>
<td>15.541</td>
<td>-1.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-1.896</td>
<td>1.972</td>
<td>-1.838</td>
<td>2.125</td>
<td>-7.295</td>
<td>0.340</td>
<td>13.359</td>
<td>-1.059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>parms</td>
<td>-0.371</td>
<td>0.423</td>
<td>-0.494</td>
<td>1.651</td>
<td>11.091</td>
<td>-1.260</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-2.623</td>
<td>2.001</td>
<td>-2.644</td>
<td>5.118</td>
<td>10.272</td>
<td>-1.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>