Disclosure and Cost of Equity Capital Revisited

ABTRACT

We reexamine the relation between disclosure indices and cost of equity capital employing an empirical specification similar to that of Botosan (1997) for a substantially larger sample over an extended time frame made possible by textual analysis. Our results provide no support for a hypothesis of a negative relation between disclosure indices and implied cost of equity capital. Rather, consistent with a bias of implied cost of equity capital as a proxy for expected return depicted by Hughes, Liu, and Liu (2009), we find strong evidence of a positive relation.

Keywords: Disclosure; Implied cost of equity capital; Expected return; Textual analysis.

JEL Classification: G12; G14; M41.

I. INTRODUCTION

A prominent issue in accounting research is the relation between disclosures of firm-specific information such as that contained in financial reports and the cost of equity capital (i.e., expected return). Neoclassical theory on asset pricing argues that expected return is entirely composed of a risk-free return and a systematic risk premium. Idiosyncratic risks are presumed to be eliminated through diversification, implying that information on such risks has no effect on expected return. However, in apparent contradiction to theory, empirical studies such as Botosan (1997) offer evidence of a negative association between information contained in firm disclosures and implied cost of equity capital after controlling for effects of systematic risk. Limitations of her study are the small sample due to reliance on hand-collected data in creating indices as measures of firm disclosures and the questionable efficacy of implied cost of equity capital as a proxy for expected return.

In this study, we exploit recently developed textual analysis in a replication of Botosan's (1997) study applied to a greatly expanded sample in both cross-section and time. Our sample consists of 43,806 firm-year observations spanning years from 1995 to 2019, for which we have sufficient data to construct the disclosure indices that we employ in our analysis. Applying filters for implied cost of equity capital estimation and controls for systematic risk leaves 28,284 firm-year observations for data acquired from Value Line and 37,341 firm-year observations acquired from I/B/E/S. With Value Line, we compute implied cost of equity capital employing Brav, Lehavy, and Michaely's (2005) model parameterized by price targets, dividend forecasts, and dividend growth rates, while with I/B/E/S data, we take the average of implied cost of equity capital employing models by Claus and Thomas (2001), Gebhardt, Lee, and Swaminathan (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004) as suggested by Hail and Leuz (2006).

For most of our analysis, we use a modestly refined version of Botosan's original construction of disclosure indices introduced by Francis, Nanda, and Olsson (2008). The advantage of their version is enabling a validity check on the construction of indices by comparing descriptive statistics of index values. Francis et al.'s sample consists of 677 firms in fiscal 2001. Our textual analysis yielded 627 observations for 2001, for which the means, standard deviations, and quartile distributions closely correspond to those statistics in their study.[1] We extended the comparison of descriptive statistics to the full sample for which disclosure indices are feasibly constructed and find remarkably similar results. We emphasize analyses using Value Line data consistent with Botosan's original study. However, similar results are obtained using more extensive I/B/E/S data.

Similar to Botosan (1997), we regress estimates of implied cost of equity capital on proxies for systematic risk factors and disclosure indices for each of the 25 years composing our sample. Briefly summarizing our principal results under Value Line, we find that the coefficient of the total disclosure index is negative in only four years and never significantly negative in explaining the implied cost of equity capital as hypothesized by Botosan (1997). The annual coefficients are positive in the remaining 21 years and significantly positive at at least the 10% level in 10 years. We simulate the data generating process a thousand times using estimated parameters producing a similar pattern of signs and significance for yearly coefficients of the total disclosure index. Following Fama and MacBeth's (1973) procedure under the further assumption that coefficients are time-invariant, we find the average cross-sectional coefficient of the total disclosure index is significantly positive. A further panel regression with year and industry fixed effects for both industry and time also yields a similar result for the coefficient of the total disclosure index. Collectively, our results lead us to reject the hypothesis that firm-specific disclosures serve to reduce the cost of equity capital (expected return). Rather, we are left with strong evidence of a *positive* relation between such disclosures and the *implied* cost of equity capital.

---

[1] The coefficients of the total disclosure index with controls for systematic risks in both our study for 2001 and that of Francis et al. are similarly insignificant, suggesting that the effects of disclosure in their study on implied cost of equity capital, absent controls for systematic risk, may not be attributable to firm disclosures that contain idiosyncratic risks.

A plausible explanation for a positive relation between implied cost of equity capital and measures of disclosure is offered by Hughes, Liu, and Liu (2009), who model the difference between the implied cost of equity capital and expected return under the assumption that expected returns are stochastic. Merton's (1973) seminal study relates stochastic expected returns to random states of nature that induce changes in investors' opportunity set. In particular, Hughes et al. (2009) adopt Merton's illustration with stochastic expected returns in the form of stochastic betas in their analysis. The assumption of stochastic expected return is well supported by several empirical studies, including Campbell (1991), Jagannathan and Wang (1996), and Fama and French (1997). Hughes et al. (2009) identify factors including the volatility of cash flows that could account for a bias in implied cost of equity capital as a proxy for expected return. We extend the independent identically distributed expected return setting of Hughes et al. (2009) to a general mean-reverting expected return setting. We adopt a continuous-time model (in contrast to discrete-time) to parsimoniously depict a relation between cash flow volatility and bias in implied cost of equity capital.

Reflecting on the role of disclosure indices vis-à-vis the cost of equity capital, the information provided by financial data from EDGAR is, generally speaking, about past events accompanied by estimates of future events. Whether disclosures upon which disclosure indices are based are voluntary or mandatory is inextricably mixed, notwithstanding Botosan's and others' view of voluntary disclosure as a driver of cross-sectional variations by firms seeking to reduce their cost of equity capital.[2] We, alternatively, interpret disclosure indices drawn in this fashion as descriptive of the level of cash flow news, a significant portion of which pertains to idiosyncratic risks. In theory, under this interpretation, there should be no association of disclosure indices with the expected return after controlling for systematic risk. However, as evident from our model, this prediction does not apply to the implied cost of equity capital when the expected return is stochastic. In this case, the implied cost of equity capital is a biased proxy for expected return, a bias that could explain a positive association with disclosure. Empirically, consistent with this explanation, we find that idiosyncratic volatility estimated from residual returns, as a measure of idiosyncratic risks that manifest in cash flow news, is positively correlated with both disclosure indices and implied cost of equity capital.[3]

Literature

---

[2] A history of voluntary disclosure may be viewed as an implicit commitment, further blurring the distinction with mandatory disclosure.

[3] We acknowledge that empirical evidence discussed by Lambert (2009) as well as Hughes et al. (2009) is mixed with regard to conditions under which implied cost of capital as a measure of expected return is increasing in the volatility of cash flows, thereby recommending caution in interpreting these results.

Botosan (1997) provided a foundation for subsequent studies seeking to relate firm-specific information to the cost of equity capital. Botosan's sample consists of 122 hand-collected observations for manufacturing firms drawn from annual reports in 1990. The objective of her study is to empirically link indices thought to capture voluntary disclosure present in annual reports to the implied cost of equity capital as a proxy for expected returns after controlling for the effects of systematic risks. Results are consistent with a negative association. The study is seminal in the sense of lending impetus to various later studies examining pricing effects of information and its properties vis-à-vis the cost of equity capital. Botosan and Plumlee (2002) reaffirm the negative association between disclosure of information contained in annual reports and the implied cost of equity capital. A further study by Botosan, Plumlee, and Xie (2004) considers the relations between precisions of public and private information and the implied cost of equity capital. They find a negative association for the former that is more than offset by a positive association for the latter. Francis et al. (2008) refine the empirical specification in an attempt to better distinguish disclosures that are voluntary and consider the potential effects of earnings quality. They report a significant negative association between disclosure and implied cost of equity capital, but only before removing effects of systematic risks. This effect is diminished by conditioning on earnings quality, a feature that we do not explore.

Other empirical studies less directly linked to our platform include Francis, LaFond, Olsson, and Schipper (2004, 2005), who explore the effects of accruals quality and other properties of earnings on the cost of equity capital. The latter uses excess returns from a three-factor model to assess whether accruals quality is a priced risk factor, an implication challenged by Core, Guay, and Verdi (2008). Kothari, Li, and Short (2009) employ content analysis of reports by management, analysts, and the business press to assess the differential effects of favorable versus unfavorable disclosures on rolling forward estimates of a three-factor model from which they infer the cost of equity capital. Other dimensions of firm disclosures effects on the cost of equity capital include Daliwal, Li, Tsang, and Yang's (2011) study of the impact of initiating corporate disclosures of social responsibility, and Li's (2010) study of consequences mandated adoption of International Financial Reporting Standards further refined by Daske, Hail, Leuz, and Verdi (2013) to capture firm discretion in implementation.

Given a well-functioning market in which diversification eliminates pricing effects of idiosyncratic risks, we would not expect to find an impact of information pertaining to such risks to affect the cost of equity capital. In theory, the absence of pricing effects carries over to private information, as shown by Hughes, Liu, and Liu (2007) and Lambert, Leuz, and Verrecchia (2007). Although we do not address the distinction between voluntary and mandatory disclosure in assessing the effects of firm disclosures on implied cost of equity capital[4], Cheynel (2013) models the effects of voluntary disclosures on a firm's sensitivity to systematic risks and shows that betas, as a measure of that sensitivity, are lower conditional on disclosure than on non-disclosure. However, the inclusion of an interaction variable of betas and disclosure indices in order to capture the effects of disclosures on factor loadings does not alter our results in a qualitative sense.

Last, as noted, Hughes et al. (2009) model the effects of stochastic expected returns on implied cost of capital. In such a setting, they show that the implied cost of capital provides a biased estimate of expected returns. Their analysis depicts factors driving that bias, including the volatility of cash flows. In turn, we provide a continuous-time analog to their model that identifies conditions under which the bias in implied cost of equity capital is increasing in cash flow volatility and, hence, in expected disclosures as a reflection of that volatility.

---

[4] We note that the distinction between what is voluntary versus mandatory is unclear in the items that compose disclosure indices in both Botosan (1997) and Francis et al. (2008). We view the disclosure indices as a measure of the level of cash flow news that the firm is expected to provide but not limited to what is dictated by accounting standards.

Neoclassical theory argues that information on idiosyncratic risks should have no effect on expected returns. Whether theory holds in practice is an empirical question. Our study contributes to the literature in several ways. First, through textual analysis of firm disclosures, we substantially expand samples employed in earlier studies by Botosan (1997) and others employing similar sampling rules and tests. Second, we provide extensive evidence calling into question findings supporting the view that firms lower their cost of equity capital through greater disclosure of firm-specific information. Third, our findings of a positive relation between disclosure indices and implied cost of equity capital are traceable to a bias in implied cost of equity capital as a proxy for expected return characterized by Hughes et al. (2009).[5]

The remainder of this paper is organized as follows: Section II describes our disclosure indices, estimates of implied cost of equity capital, and sampling rules. Section III contains our empirical analysis of the relation between disclosure indices and the implied cost of equity capital. Section IV sets forth our model on bias in implied cost of equity capital and correlations of predicted relations. Section V concludes.

II. DISCLOSURE INDICES, ESTIMATES OF IMPLIED COST OF EQUITY CAPITAL, AND SAMPLING RULES

Disclosure Indices

---

[5] Empirically, allowing for interaction between disclosure indices and betas as a measure of systematic risk has no qualitative effects.

Our construction of disclosure indices is based on Francis et al.'s (2008) modified version of Botosan's (1997) approach.[6] The total disclosure measure consists of four categories, of which the contents are listed in Table 1. We began our coding by downloading all 10-K reports from the EDGAR system. For each 10-K report, according to the coding scheme, we assign the binary element (e.g., whether the firm discloses a forecasted cash flow or not) a value of one, if existing, or zero otherwise. For non-binary elements (e.g., number of quarters that firm discloses sales and net income), we convert it to binary variable depending on whether it is above or below the median value reported by all firms in the same year. Firms above the median receive a value of one and otherwise zero. The above procedures generate a value of zero or one for each of the elements in the coding scheme. We then scale this raw score by the maximum score in that year to obtain a percentage-based score for each firm year. Our results are robust to other ways to aggregate the disclosure score. For example, we have the same results if we assign equal weight to each category instead of each element. Neither are the results sensitive to directly using raw scores. Appendix B provides a detailed description of our textual analysis procedure for constructing disclosure indices.

In Table 2, we report descriptive statistics on our raw total/subcategory disclosure indices for 2001 (Panel B), the year for which their data were drawn, in comparison with those of Francis et al. (2008) (Panel A). Notwithstanding a small difference in sample sizes, the two distributions are nearly identical, lending confidence that our replication of disclosure indices captures essentially the same information. We further compare the above descriptive statistics with those for our entire sample (Panel C) and find the similarity distinctive. These comparisons establish a high level of validity for our textual-based determination of disclosure scores.

---

[6] Given that EDGAR did not exist in 1990, the year of Botosan's (1997) data, we cannot establish the validity of our replication of her coding scheme via textual analysis. Accordingly, we resort to employing the coding scheme of Francis et al. (2008). The latter excludes background information and MD&A arguing that these measures are less reflective of voluntary disclosure as a result of SEC restrictions. These categories were insignificant in Botosan's study, suggesting little impact of their exclusion. As noted earlier, this shift to Francis et al.'s coding enables a comparison of summary statistics as a means of verifying the validity of our recourse to textual analysis.

In Table 3, we report inter-correlations of raw disclosure indices in total and by category. The correlations are uniformly positive as one would expect given the likelihood that greater disclosure in one category would be accompanied by greater disclosure in another. These results are consistent with Botosan's. Furthermore, for each firm over time, our disclosure measure has an AR(1) coefficient of 0.9 and a coefficient of variation (standard deviation scaled by mean) of 0.2. Therefore, our disclosure measure is persistent over time for each firm.

Implied Cost of Equity Capital

We use an estimation model based on future price targets to construct our main proxy for the implied cost of equity capital. Specifically, we follow Brav, Lehavy, and Michaely (2005) and Francis, Nanda, and Olsson (2008) to form the implied cost of equity capital measure ICC_{VL}, which is derived from Value Line data on analysts' four-year out price targets (TP), dividend forecasts (DIV), and dividend growth rates (g). Assuming dividends are reinvested at the firm cost of equity capital ICC_{VL}, Brav, Lehavy, and Michaely (2005) suggest the following equation for the expected return:

$$(1 + ICC_{VL})^4 = \frac{TP}{P} + \frac{DIV\left[\frac{(1 + ICC_{VL})^4 - (1 + g)^4}{ICC_{VL} - g}\right]}{P}$$

where P = stock price nine days prior to the date of the Value Line report. For each firm year in our sample, we determine the value of ICC_{VL} that satisfies the above equation and use this as our estimate of the implied cost of equity capital. Following Francis, Nanda, and Olsson (2008), we use the average of the firm's four quarterly estimates of ICC_{VL} to form the annual estimation of the implied cost of equity capital for that year. We draw similar inferences if we use the quarterly ICC_{VL} estimation right after the annual report publication date.

We advance the implied cost of equity capital estimated from Value Line data as our principal measure for two reasons. First, it is essentially identical to the implied cost of equity capital measure used in Botosan (1997) and Francis, Nanda, and Olsson (2008) lending comparability to those previous studies on the relation between disclosure and cost of equity capital. Second, aside from the question of potential bias as a proxy for expected return, the Value Line based implied cost of equity capital has a higher construct validity than other measures regarding associations with firm risk attributes (Botosan and Plumlee 2005) and a significant correlation with future realized returns (Francis et al. 2004). Moreover, due to the use of the four-year out price targets, Value Line makes fewer assumptions of the long-term growth rate than other estimation models.

To reinforce our findings using Value Line, we also consider alternative measures implementable from I/B/E/S data. An advantage of I/B/E/S data is a significant enlargement of our sample. Following Hail and Leuz (2006), we use four different models in estimating implied cost of equity capital, including those by Claus and Thomas (2001), Gebhardt, Lee, and Swaminathan (2001), Ohlson and Juettner-Nauroth (2005) (as implemented by Gode and Mohanram 2003), and Easton (2004), as well as the average of these four estimates as our measure of implied cost of equity capital. Each of the four models calculates the implied cost of equity capital as the internal rate of return that equates current stock price with the discounted future dividends or earnings. We describe each model in more detail below.

Claus and Thomas (2001):

$$P_t = bv_t + \sum_{\tau=1}^{T} \frac{\hat{x}_{t+\tau} - ICC_{CT} \cdot bv_{t+\tau-1}}{(1 + ICC_{CT})^{\tau}} + \frac{(\hat{x}_{t+T} - ICC_{CT} \cdot bv_{t+T-1})(1+g)}{(ICC_{CT} - g)(1 + ICC_{CT})^{T}}$$

Gebhardt, Lee, and Swaminathan (2001):

$$P_t = bv_t + \sum_{\tau=1}^{T} \frac{\hat{x}_{t+\tau} - ICC_{GLS} \cdot bv_{t+\tau-1}}{(1 + ICC_{GLS})^{\tau}} + \frac{(\hat{x}_{t+T+1} - ICC_{GLS} \cdot bv_{t+T})}{ICC_{GLS}(1 + ICC_{GLS})^{T}}$$

Ohlson and Juettner-Nauroth (2005):

$$P_t = \left(\hat{x}_{t+1}/ICC_{OJ}\right) \cdot \left(g_{st} + ICC_{OJ} \cdot \hat{d}_{t+1}/\hat{x}_{t+1} - g_{lt}\right)/\left(ICC_{OJ} - g_{lt}\right)$$

Modified price-earnings growth (PEG) ratio model by Easton (2004):

$$P_t = \left(\hat{x}_{t+2} + ICC_{PEG} \cdot \hat{d}_{t+1} - \hat{x}_{t+1}\right)/ICC_{PEG}^{2}$$

The first two models are both special cases of the residual income valuation model. Specifically, $\hat{x}_{t+\tau}$ is the forecasted earnings per share of year t+τ, bv_t is the book value per share at year t, and g is the annualized inflation rate. The major difference between these two is the assumption of the residual income growth rate. Claus and Thomas (2001) assume residual income grows at the expected inflation rate after five years (T=5), while Gebhardt, Lee, and Swaminathan (2001) derive the residual incomes by linearly fading the forecasted return on equity to the industry median from T=3 to T=12, and assuming residual income to remain constant from T=12 on.

The latter two models are based on the abnormal earnings growth valuation model developed by Ohlson and Juettner-Nauroth (2005). Following Gode and Mohanram (2003), the short-term growth rate g_{st} is estimated as the average between the forecasted earnings growth rate from year *t+1* to *t+2* and the five-year growth forecast provided by the I/B/E/S analysts. We use the annualized inflation rate to proxy for the long-term growth rate g_{lt}. As suggested by Gebhardt, Lee, and Swaminathan (2001), we estimate the dividend payout ratio by dividing the actual dividends from the most recent fiscal year by earnings over the same time period and calculate the expected future net dividends per share $\hat{d}_{t+1}$ accordingly. Notably, both these two models require a positive change in forecasted earnings to yield a numerical solution. To be consistent with our Value Line based measure, we use stock price P_t nine days prior to the date of the I/B/E/S report in year *t*.

Sample and Variable Description

We collect accounting information from Compustat, stock price from CRSP, analyst forecasts from Value Line and I/B/E/S, and 10-K file and annual report information from SEC EDGAR. Our sample spans calendar years from 1995 (the year EDGAR was fully implemented) to 2019, which corresponds to 10-K files for fiscal years from 1994 to 2018. Applying a filter based on all firm-years followed by Value Line and removing firms for which we lack sufficient data for disclosure indices and control variables leaves a final sample of 28,284 firm-years, or 83.6% of all firm-years with sufficient data for the Value Line implied cost of equity capital estimation.[7]

Table 4 summarizes our sample selection procedures (Panel A) and reports the sample distribution across industries (Panel B). Overall, our sample is well diversified across different industries and years. None of the industries takes up more than 10% of our sample, of which business services, retail, utilities, and electronic equipment take up 5% or above. Our sample slightly grows over time because Value Line is following more and more firms and providing necessary forecast data to calculate the implied cost of equity capital. There are 2,791 individual firms in our sample, of which 286 firms appear only once, and 163 firms appear in the sample for all 25 years.

---

[7] Sample size under I/B/E/S is 37,341 firm-year observations.

Table 5 reports the summary statistics for all variables used in this study, including the disclosure index, the implied cost of equity capital, and other control variables. We report both the raw disclosure score, *Raw Disc (Total),* and the percentage-based disclosure score, *Disc (Total)*, and a more detailed subcategory disclosure summary can be found in Table 2. Consistent with the extant literature, Value Line provides a higher implied cost of equity capital ICC_{VL} than all types of I/B/E/S measures (ICC_{IBES}, ICC_{GLS}, ICC_{CT}, ICC_{OJ}, ICC_{PEG}). The control variables include beta (measured with the CAPM model using a minimum of 24 monthly returns over the 60 months prior to the annual report publication date), size (proxied by the log of firms' market value of equity), book-to-market ratio (proxied by the log of firms' book-to-market ratio), ROA (return on assets), idiosyncratic risk (measured as the residual volatility of the Fama-French 3-Factor model as in Ang et al. 2006).

III. EMPIRICAL ANALYSIS OF DISCLOSURE EFFECTS ON IMPLIED COST OF EQUITY CAPITAL

Yearly Cross-Sectional Regression (Value Line)

For each calendar year from 1995 to 2019, we run the following cross-sectional regression where size, Beta, and book-to-market ratio are presumed to capture the effects of systematic risk, Disc denotes the total percentage-based disclosure score, and ICC is the implied cost of equity capital using Value Line data and Brav et al.'s (2003) model:

$$\mathrm{ICC}_{it} = \alpha_t + \gamma_{1t}\,\mathrm{Disc}_{it} + \gamma_{2t}\beta_{it} + \gamma_{3t}\,\mathrm{Size}_{it} + \gamma_{4t}\,\ln \mathrm{B/M}_{it} + \epsilon_{it}\,.$$

Table 6 contains the yearly cross-sectional regression results.[8] Notably, the slope coefficient for Disc is negative (significantly) only in 4 years (0 years), but is positive (significantly) in 21 years (10 years). The high frequency of positive coefficients, nearly half of which are statistically significant, cannot be attributed to chance. As shown below, when pooling the data over all the years together, the slope coefficient for Disc is significantly positive.

These results are clearly inconsistent with the hypothesis that disclosures reduce the cost of equity capital.[9] Indeed, we are not aware of any rational expectations pricing theory that would explain these findings. As mentioned earlier, in section IV, we offer a plausible explanation based on the potential bias contained in the implied cost of equity capital as a proxy for expected return.

Pooling Cross-Sectional Regressions (Value Line)

Table 7 reports pooling regression results with Value Line based estimates of implied cost of equity capital. We regress the implied cost of equity capital on factors assumed to capture effects of systematic risk (size, beta, book-to-market ratio) and disclosure scores. As noted, our full sample includes calendar years from 1995 to 2019.[10] Column (1) reports the average coefficient estimates from yearly cross-sectional regressions, as Fama and MacBeth (1973). We report the standard Fama-MacBeth intertemporal t-statistics based on the Newey-West consistent standard error. The significant positive coefficients on total disclosure indices complement the results from our yearly regressions as reported in Table 6.

---

[8] As we discuss later, the negative signs on size are consistent with smaller firms tending to have more volatile cash flows, thereby contributing to a bias in implied cost of equity capital as a proxy for expected return.

[9] Botosan (1997) reports a negative coefficient on Disc from a similarly specified regression. Her data was hand collected from 1990 annual reports for 122 manufacturing firms and unavailable.

[10] We also ran our analysis excluding 2001 for the tech bubble and 2007-2008 for the financial crisis with qualitatively unchanged results.

Panel regression results with year and industry fixed effects are presented in the last two columns of Table 7. Year fixed effects account for the time-series variation in the implied cost of equity capital across business cycles and control for any potential time trends of the disclosure index; the inclusion of industry fixed effects alleviate the concern that our finding is caused by industry-related disclosure norms and/or cross-industry variations of the implied cost of equity capital. Our results also persist after controlling for unobservable factors affecting a given industry in a given year, which are absorbed by year*industry fixed effects. With or without these fixed effects and/or including other fixed effects, the regression results are similar. Column (2) and (3) reports the panel regression results with year fixed effects and industry fixed effects and year*industry fixed effects to examine the extent to which time and/or industry variations contribute to our principal findings. Again, the slope coefficient on Disc is significantly positive.

Pooling Cross-Sectional Regressions (I/B/E/S)

In Table 8, we report similar pooling regression results to those in Table 7 using I/B/E/S estimates of implied cost of equity capital from 1995 to 2019. We calculate five different implied cost of equity capital measures from previous literature, including Claus and Thomas (2001), Gebhardt, Lee, and Swaminathan (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2005), and the average of these four measures. Column (1) reports Fama-Macbeth regression using the average. Column (2) to (5) reports panel regression results with five different estimates of implied cost of equity capital controlling for year*industry fixed effects. The results are qualitatively in line with those reported using Value Line data. The difference in magnitudes of coefficients is traceable to the relative scale difference in the implied cost of equity capital measures.

Overall, the findings reported in the previous three tables are remarkably robust, leaving little doubt that results in Botosan (1997) are special to the year for which she gathered data and do not generalize to samples that extend over time.

IV. DISCLOSURE AND BIAS IN IMPLIED COST OF EQUITY CAPITAL

Model of Bias in Implied Cost of Equity Capital as Proxy for Expected Return

We now show that the difference between the implied cost of equity capital and expected return due to the stochastic property of the latter produces a positive relation between disclosure measures and the implied cost of equity capital. Our model extends the identical independently distributed (in time-series) expected return setting of Hughes et al. (2009) to a more general mean-reverting expected return. We equate dividends in our model with cash flows available for distribution to shareholders in Hughes et al. (2009). We begin with a lemma that derives the implied cost of equity capital. We then show that the bias in implied cost of equity capital increases with dividend volatility, an analog to cash flow volatility in Hughes et al. (2009).

Lemma. *Suppose the dividend* D_t *of a stock follows a geometric Brownian motion*

$$dD_t = D_t\left(g dt + \sigma_d dB_t^d\right)$$

and the expected return μ_t *follows an OU process*

$$d\mu_t = -K(\mu_t - \bar{\mu})dt + \sigma dB_t$$

where $g, \sigma_d > 0$*,* $K > 0$*,* $\bar{\mu}$*, and* $\sigma > 0$ *are all constant,* B_t^d *and* B_t *are two standard Brownian motions with a constant correlation* ρ *under the physical measure. Then the price-dividend ratio* Φ_t *at time* t *is*

$$\Phi_t \equiv \frac{P_t}{D_t} = \int_0^\infty e^{-(\hat{\mu}-g)x} e^{-\frac{\mu_t-\hat{\mu}}{K}\left(1-e^{-Kx}\right)+\frac{\sigma^2}{2K^2}\left(x-\frac{2}{K}\left(1-e^{-Kx}\right)+\frac{1}{2K}\left(1-e^{-2Kx}\right)\right)} dx$$

where

$$\hat{\mu} = \bar{\mu} + \frac{\rho\sigma\sigma_d}{K},$$

The implied cost of equity capital is then

$$\nu_t = g + \frac{1}{\Phi_t}.$$

Proof. See appendix.

We assume the time-varying expected return follows a mean-reverting process, which is used in most empirical studies where the process of the expected return is needed to be explicitly assumed. An example of stochastic expected returns is to assume $\mu_t = r_f + \lambda_t\beta$, where the riskless return r_f and β are constant and the market risk premium λ_t follows

$$d\lambda_t = -K\left(\lambda_t - \bar{\lambda}\right)dt + \sigma_\lambda dB_t.$$

In this case,

$$\bar{\mu} = r_f + \bar{\lambda}\beta,$$

and

$$\sigma = \beta\sigma_\lambda.$$

The Lemma shows that, implied cost of equity capital ν_t does not equal the expected return μ_t. It is a non-linear function of expected return μ_t.[11] Some special cases will help to understand the relation between the cost of equity capital and expected return.

---

[11] Implied cost of equity capital depends on parameters $\rho, K, \sigma,$ and g. The sign of ρ is qualitatively important as we discussed above. The dependence of implied cost of equity capital on $K, \sigma,$ and g are non-monotonic, which is decided by the value of the expected return μ_t.

- When $K \to +\infty$ and σ is finite. In this case, $\frac{\sigma}{K} \to 0$, the expected return μ_t is effectively constant, $\mu_t = \bar{\mu}$. In this limit, one can show that $\Phi_t = 1/(g + \bar{\mu})$, and the cost of equity capital $\nu_t = \bar{\mu}$.
- When $K \to +\infty$ and $\sigma \to +\infty$, such that $\frac{\sigma}{K} \neq 0$ (K and σ goes to infinity proportionally). In this limit, one can show that $\Phi_t = 1/(g + \bar{\mu} + \frac{\rho\sigma\sigma_d}{K})$, and the cost of equity capital $\nu_t = \bar{\mu} + \frac{\rho\sigma\sigma_d}{K}$. Intuitively, when $\rho > 0$, the expected return μ_t is high when D_t is high, and vice versa. So high dividend states receive high discount, which leads to lower P/D value than when μ_t is constant, which in turn leads to a higher cost of equity capital. This intuition holds in general (also as the 2-state model described in Lambert 2009). This case is the continuous-time analog of Hughes et al. (2009).

The risk of the dividend in general has two components, systematic and idiosyncratic. Thus the dividend volatility depends on beta as well as idiosyncratic volatility. For example, in a market model, $\sigma_d^2 = \beta_d^2\sigma_m^2 + {\sigma_d^i}^2$,where β_d is the beta of the dividend, σ_m is the market volatility, and σ_d^i is the idiosyncratic volatility. In neoclassical asset pricing models, the expected return depends on beta, but not on idiosyncratic volatility.

Proposition. *As in the above lemma, suppose the dividend of a stock follows a geometric Brownian motion*

$$dD_t = D_t\left(gdt + \sigma_d dB_t^d\right)$$

and the expected return follows a OU process

$$d\mu_t = -K(\mu_t - \bar{\mu})dt + \sigma dB_t$$

where g, $\sigma_d > 0$*,* $K > 0$*,* $\bar{\mu}$*, and* $\sigma > 0$ *are all constant,* B_t^d *and* B_t *are two standard Brownian motions with a constant correlation* ρ *under the physical measure. Then the implied cost of equity capital* ν_t *increases with the idiosyncratic volatility* σ_d^i *of the dividend if* $\rho > 0$.

Proof. Note that

$$\frac{\partial \nu_t}{\partial \sigma_d^i} = -\frac{1}{\Phi_t^2}\frac{\partial \Phi_t}{\partial \sigma_d}\frac{\partial \sigma_d}{\partial \sigma_d^i}$$

From $\hat{\mu} = \bar{\mu} + \frac{1}{K}\rho\sigma\sigma_d$, the derivative of Φ_t with respect to σ_d is

$$\frac{\partial \Phi_t}{\partial \sigma_d} = -\frac{\rho\sigma}{K}\int_0^{\infty}\left(x - \frac{1}{K}(1 - e^{-Kx})\right)e^{-(\hat{\mu}-g)x}e^{-\frac{\mu_t-\hat{\mu}}{K}(1-e^{-Kx})+\frac{\sigma^2}{2K^2}\left(x-\frac{2}{K}(1-e^{-Kx})+\frac{1}{2K}(1-e^{-2Kx})\right)}dx.$$

We used the assumption that the expected return does not depend on idiosyncratic volatility σ_d^i . Thus, the derivative of the expected return μ_t , as well as the parameter $\bar{\mu}$, with respect to the idiosyncratic volatility is zero. The result follows noting that $\frac{\partial \sigma_d}{\partial \sigma_d^i} > 0$ and $-\frac{\rho\sigma_d}{K} < 0$ because $\rho > 0$ and $\left(x - \frac{1}{K}(1 - e^{-Kx})\right) > 0$ for $x > 0$. ■

It is widely documented that the correlation (ρ) between dividends (cash flows) and expected returns is positive. Under this condition, the proposition establishes that the implied cost of equity capital increases with dividend volatility, given that the expected return is independent of the dividend volatility under neoclassical asset pricing theory. Plausibly assuming that firms with a higher disclosure index also have higher dividend volatility provides an explanation for our empirical finding of a positive association between implied cost of equity capital and disclosure indices as a measure of cash flow news.[12] Below, we take a closer look empirically at the links between disclosure indices, cash flow volatility, and potential bias in implied cost of equity capital.

Correlations between Disclosure, Idiosyncratic Risk, and Implied Cost of Equity Capital

In Table 9, we estimate the correlations between the Value Line implied cost of equity capital, cash flow volatility stemming from idiosyncratic risk, and disclosure indices. Idiosyncratic risk is measured as the idiosyncratic volatility of the Fama-French 3-factor model (Ang et al. 2006). Since the disclosure index slightly increases over time, we detrend our disclosure score at the year level. We get similar results if we control for both year and industry fixed effects.

The correlations reported in Table 9 support our explanation for a positive relation between implied cost of equity capital and disclosure indices. Consistent with the Proposition that implied cost of equity capital is increasing with idiosyncratic volatility, we find significant positive correlations between the implied cost of equity capital, idiosyncratic volatility as a measure of cash flow volatility attributable to idiosyncratic risk, and disclosure indices as a measure of news that gives rise to idiosyncratic volatility.

In short, we believe we have made a strong case for the prospect that the positive relation between disclosure indices and implied cost of equity capital documented in section III is likely driven by a bias in the latter as a proxy for expected return.

V. CONCLUSION

Understanding the relation between information and the cost of equity capital is central to accounting research. Botosan (1997) was seminal in its examination of the impact that disclosures contained in financial reports might have on a firm's cost of equity capital, i.e., its expected return. Her findings were suggestive of an effect of greater disclosure serving to reduce expected return after controlling for systematic risks. This interpretation was controversial. Neoclassical theory maintains that information on idiosyncratic risks should have no such effect on expected return. The dissonance between the empirical findings and theory prompted us to reexamine the linkage between disclosure and expected return, given the availability of an empirical technology that enabled a greatly enlarged sample extending over a quarter century. Our results strongly reject the hypothesis of a negative relation between disclosure and implied cost of equity capital as a proxy for expected return over that time frame. On the contrary, we found compelling evidence of a positive relation.

---

[12] Prices move with the news. We view disclosure indices as a reflection of cash flow news, a large component of which pertains to idiosyncratic risk.

A potential explanation for our results is the prospect of a systematic bias in implied cost of equity capital modeled by Hughes et al. (2009) and depicted in our parsimonious, continuous-time extension of their analysis. Specifically, in keeping with Merton's (1973) insight of stochastic expected returns, we identify conditions under which a bias in implied cost of equity capital as a proxy for expected return is increasing in the volatility of cash flows. Viewing disclosure indices as depicting the level of cash flow news suggests a positive relation between disclosure and cash flow volatility. Moreover, given that firms are small relative to the market suggests a positive relation between disclosure and the volatility of residual returns from a factor model as a measure of idiosyncratic volatility. We confirm these relationships empirically.

As with most empirical studies, we are limited in identifying defensible proxies for variables that we cannot observe. Although we believe our study poses a strong case for a bias in implied cost of equity capital as a measure of expected return driving a positive association between disclosure indices and implied cost of equity capital, we acknowledge that the conditions for our conclusions in that regard to hold as discussed in Hughes et al. (2009) and Lambert (2009) are not immutable.

Finally, we note that the advance in analyzing the contents of financial reports through textual analysis opens a new chapter for empirical studies addressed to the role that information firms provide to investors affects what they require as an expected return. The expansion in sampling enabled in this study calls into question the generalizability of earlier work limited by the available technology at those times. It seems to us that there are many opportunities for accounting researchers to employ textual analysis in revisiting previous studies or initiating new inquiries seeking to connect information with asset pricing.

References

Ang, A., R. Hodrick, Y. Xing, and X. Zhang. 2004. The cross-section of volatility and expected returns. *Journal of Finance* 61(1): 259-299.

Botosan, C. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72(3): 323-349.

Botosan, C., and M. Plumlee. 2002. A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research* 40(1): 21-40.

Botosan, C., M. Plumlee, and Y. Xie. 2004. The role of information precision in determining the cost of equity capital. *Review of Accounting Studies* 9: 233-259.

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

Brav, A., R. Lehavy, and R. Michaely. 2005. Using expectations to test asset pricing models. *Financial Management* 34(3): 31-64.

Campbell, J. 1991. A variance decomposition for stock return. *The Economic Journal* 101, 157-179.

Cheynel, E. 2013. A theory of voluntary disclosure and cost of capital. *Review of Accounting Studies* 18, 987-1020.

Claus, J., and J. Thomas. 2001. Equity premia as low as three percent? evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance* 56(5): 1629-1666.

Core, J., W. Guay, and R. Verdi. 2008. Is accruals quality a priced risk factor? *Journal of Accounting and Economics* 46(1): 2-22.

Daske, H., L. Hail, C. Leuz, and R. Verdi. 2013. Adopting a label: heterogeneity in the economic consequences around IAS/IFRS adoptions. *Journal of Accounting Research* 51(3): 495-547.

Dhaliwal, D., O. Li, A. Tsang, and Y. Yang. 2011. Voluntary nonfinancial disclosure and the cost of equity capital: the initiation of corporate social responsibility reporting. *The Accounting Review* 86(1): 59-100.

Easton, P. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79(1): 73-95.

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

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43(2): 153-193.

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

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

Francis, J., D. Nanda, and P. Olsson. 2008. Voluntary disclosure, earnings quality, and cost of capital. *Journal of Accounting Research* 46(1): 53-99.

Gebhardt, W., C. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39(1): 135-176.

Gode, D., and P. Mohanram. 2003. Inferring the cost of capital using the Ohlson-Juettner model. *Review of Accounting Studies* 8: 399-431.

Hail, L., and C. Leuz. 2006. International differences in the cost of equity capital: Do legal institutions and securities regulation matter? *Journal of Accounting Research* 44(3): 485-531.

Hughes, J., J. Liu, and J. Liu. 2007. Information asymmetry, diversification, and cost of capital. *The Accounting Review* 82(3): 705-729.

Hughes, J., J. Liu, and J. Liu. 2009. On the relation between expected returns and implied cost of capital. *Review of Accounting Studies* 14: 246-259.

Jagannathan, R., and Z. Wang. 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51(1), 3-53.

Kothari, S., X. Li, and J. Short. 2009. The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis. *The Accounting Review* 84(5): 1639-1670.

Lambert, R., C. Leuz, and R. Verrecchia. 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45(2): 385-420.

Lambert, R. 2009. Discussion of “on the relation between expected returns and implied cost of capital”. *Review of Accounting Studies* 14: 260–268.

Li, S. 2010. Does mandatory adoption of international financial reporting standards in the European union reduce the cost of equity capital? *The Accounting Review* 85(2): 607-636.

Merton, R. 1973. An intertemporal capital asset pricing model. *Econometrica* 41(5): 867-887.

Ohlson, J., and B. Juettner-Nauroth. 2005. Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies* 10: 349-365.

Appendix: Proof of Lemma

Proof. Without loss of generality, we will prove for the case of t=0. The price-dividend ratio Φ_0 is given by

$$\Phi_0 = \mathrm{E}_0\left[\int_0^{\infty} e^{\int_0^t -\mu_v dv + \left(g - \frac{1}{2}\sigma_d^2\right)dv + \sigma_d dB_v^d}\right].$$

The expected return μ_t satisfies

$$\mu_t = \bar{\mu} + (\mu_0 - \bar{\mu})e^{-Kt} + \int_0^t e^{-K(t-v)}\sigma dB_v \, .$$

The cumulative expected return is

$$\int_0^t \mu_v dv = \bar{\mu}t + \frac{\mu_0 - \bar{\mu}}{K}(1 - e^{-Kt}) + \int_0^t \int_0^v e^{-K(v-u)}\sigma dB_u dv.$$

Changing the integration order,

$$\int_0^t \int_0^v e^{-K(v-u)}\sigma dB_u dv = \int_0^t \int_u^t e^{-K(v-u)}\sigma dv dB_u = \frac{\sigma}{K}\int_0^t \left(1 - e^{-K(t-u)}\right)dB_u.$$

So we have

$$\int_0^t -\mu_v dv + g dv - \frac{1}{2}\sigma_d^2 dv + \sigma_d dB_v^d = \int_0^t -\mu_v dv + g dv - \frac{1}{2}\sigma_d^2 dv + \sigma_d\left(\rho dB_v + \sqrt{1-\rho^2}dB_v^{\perp}\right)$$
$$= -\left(\bar{\mu} - g + \frac{1}{2}\sigma_d^2\right)t - \frac{\mu_0 - \bar{\mu}}{K}(1 - e^{-Kt}) + \int_0^t \left(-\frac{\sigma}{K}\left(1 - e^{-K(t-u)}\right) + \rho\sigma_d\right)dB_u + \sigma_d\sqrt{1-\rho^2}dB_u^{\perp}.$$

This is a normal random variable. The mean is given by the first two terms. The variance is given by Ito' Isometry

$$\int_0^t \left(-\frac{\sigma}{K}\left(1 - e^{-K(t-u)}\right) + \rho\sigma_d\right)^2 du + \sigma_d^2(1-\rho^2)du$$
$$= \frac{\sigma^2}{K^2}\left(t - \frac{2}{K}(1 - e^{-Kt}) + \frac{1}{2K}(1 - e^{-2Kt})\right) + \frac{\rho\sigma\sigma_d}{K}\frac{1}{K}(1 - e^{-Kt}) - \frac{2\rho\sigma\sigma_d}{K}t + \sigma_d^2 t.$$

Using the moment generating function of a normal random variable, we get

$$
\begin{aligned}
\Phi &= \int_0^\infty \mathrm{E}_0\left[e^{\int_0^t -\mu_v dv+\left(g-\frac{1}{2}\sigma_d^2\right)dv+\sigma_d dB_v^d}\right] \\
&= \int_0^\infty e^{-(\hat{\mu}-g)t} e^{-\frac{\mu_0-\hat{\mu}}{K}(1-e^{-Kt})+\frac{\sigma^2}{2K^2}\left(t-\frac{2}{K}(1-e^{-Kt})+\frac{1}{2K}(1-e^{-2Kt})\right)} dt\,.
\end{aligned}
$$

■

We need the following transversality condition

$$\hat{\mu} - g - \frac{\sigma^2}{2K^2} > 0.$$

The transversality condition implies that the limit either $K \to 0$ or $\sigma \to \infty$ is not appropriate.

Next, implied cost of equity capital ν is given by

$$\Phi_0 = \frac{1}{\nu_0 - g}$$

which gives

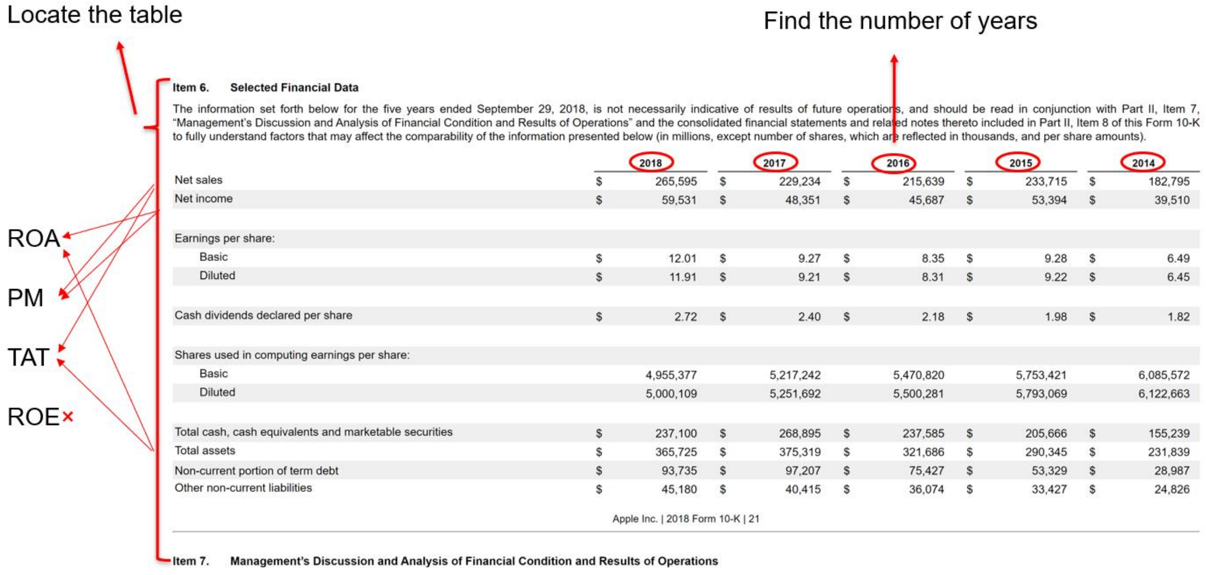
$$\nu_0 = g + \frac{1}{\Phi_0}.$$

Appendix: Disclosure Measure Construction

This section describes our textual analysis procedure. We mainly apply two coding schemes to replicate Botosan's disclosure index: table coding that features figure positioning and key information extraction; and text coding that features fuzzy search and number matching.

We use table coding for most items (a-e) in Category I of Table 1 since the summary of historical results is typically listed in an integrated table within "Item 6. Selected Financial Data" in Form 10-K. Essentially, we want to count the number of years/quarters that a firm provides sufficient information to calculate the financial ratios, including return on assets (ROA), net profit margin (PM), asset turnover (TAT), return on equity (ROE), etc. Figure A.1 shows an example of the table coding procedure for Apple's 10-K file in fiscal 2018. Apple provides ROA information for five years since we can extract total assets and net income from 2014 to 2018 (similar results for PM and TAT). As a comparison, ROE is missing since there is no information on total shareholders' equity for the past five years.

FIGURE A.1: An Example of the Table Coding



Item 6. Selected Financial Data

The information set forth below for the five years ended September 29, 2018, is not necessarily indicative of results of future operations, and should be read in conjunction with Part II, Item 7, "Management's Discussion and Analysis of Financial Condition and Results of Operations" and the consolidated financial statements and related notes thereto included in Part II, Item 8 of this Form 10-K to fully understand factors that may affect the comparability of the information presented below (in millions, except number of shares, which are reflected in thousands, and per share amounts).

	2018	2017	2016	2015	2014
Net sales	$ 265,595	$ 229,234	$ 215,639	$ 233,715	$ 182,795
Net income	$ 59,531	$ 48,351	$ 45,687	$ 53,394	$ 39,510
Earnings per share:					
Basic	$ 12.01	$ 9.27	$ 8.35	$ 9.28	$ 6.49
Diluted	$ 11.91	$ 9.21	$ 8.31	$ 9.22	$ 6.45
Cash dividends declared per share	$ 2.72	$ 2.40	$ 2.18	$ 1.98	$ 1.82
Shares used in computing earnings per share:					
Basic	4,955,377	5,217,242	5,470,820	5,753,421	6,085,572
Diluted	5,000,109	5,251,692	5,500,281	5,793,069	6,122,663
Total cash, cash equivalents and marketable securities	$ 237,100	$ 268,895	$ 237,585	$ 205,666	$ 155,239
Total assets	$ 365,725	$ 375,319	$ 321,686	$ 290,345	$ 231,839
Non-current portion of term debt	$ 93,735	$ 97,207	$ 75,427	$ 53,329	$ 28,987
Other non-current liabilities	$ 45,180	$ 40,415	$ 36,074	$ 33,427	$ 24,826

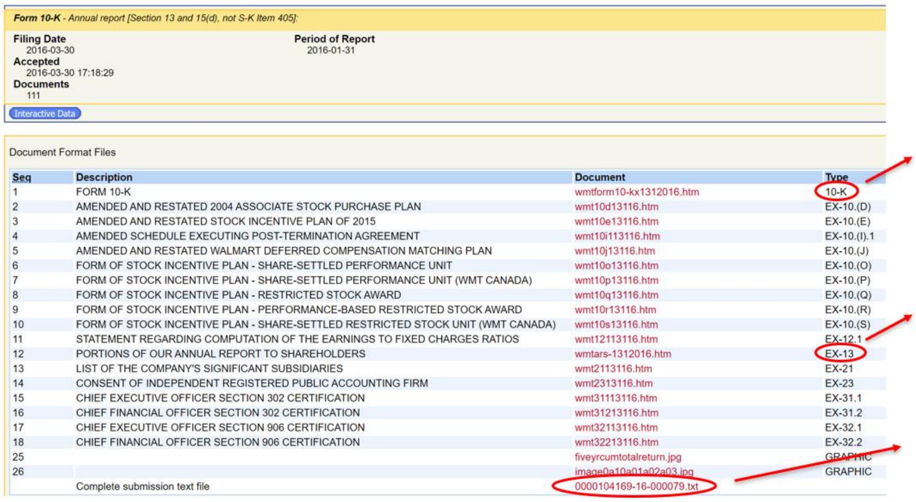
Apple Inc. | 2018 Form 10-K | 21

Item 7. Management's Discussion and Analysis of Financial Condition and Results of Operations

For the table coding procedure, there are three files that may contain the desired annual report information in the EDGAR system. Most of our annual report information comes from the traditional 10-K files. However, firms may also choose to put quantitative and MD&A information in the EX-13 form or only uploaded an all-in-one text file at the early stage of EDGAR implementation. Figure A.2 shows an example of Walmart's 2015 annual report page with all three different files uploaded simultaneously. For each firm in each year, we first go over the balance sheet, income statement, and cash flow statement in the original 10-K file. If the above information is missing, we would turn to the EX-13 form and the all-in-one text file and check whether that firm used an alternative reporting method at the time.

FIGURE A.2: An Example of the Annual Report Page in EDGAR

Walmart 2015 Form 10-K



Form 10-K *- Annual report [Section 13 and 15(d), not S-K Item 405]:*

Filing Date
2016-03-30
Accepted
2016-03-30 17:18:29
Documents
111

Period of Report
2016-01-31

Document Format Files

Seq	Description	Document	Type
1	FORM 10-K	wmtform10-kx1312016.htm	10-K
2	AMENDED AND RESTATED 2004 ASSOCIATE STOCK PURCHASE PLAN	wmt10d13116.htm	EX-10.(D)
3	AMENDED AND RESTATED STOCK INCENTIVE PLAN OF 2015	wmt10e13116.htm	EX-10.(E)
4	AMENDED SCHEDULE EXECUTING POST-TERMINATION AGREEMENT	wmt10i113116.htm	EX-10.(I).1
5	AMENDED AND RESTATED WALMART DEFERRED COMPENSATION MATCHING PLAN	wmt10j13116.htm	EX-10.(J)
6	FORM OF STOCK INCENTIVE PLAN - SHARE-SETTLED PERFORMANCE UNIT	wmt10o13116.htm	EX-10.(O)
7	FORM OF STOCK INCENTIVE PLAN - SHARE-SETTLED PERFORMANCE UNIT (WMT CANADA)	wmt10p13116.htm	EX-10.(P)
8	FORM OF STOCK INCENTIVE PLAN - RESTRICTED STOCK AWARD	wmt10q13116.htm	EX-10.(Q)
9	FORM OF STOCK INCENTIVE PLAN - PERFORMANCE-BASED RESTRICTED STOCK AWARD	wmt10r13116.htm	EX-10.(R)
10	FORM OF STOCK INCENTIVE PLAN - SHARE-SETTLED RESTRICTED STOCK UNIT (WMT CANADA)	wmt10s13116.htm	EX-10.(S)
11	STATEMENT REGARDING COMPUTATION OF THE EARNINGS TO FIXED CHARGES RATIOS	wmt12113116.htm	EX-12.1
12	PORTIONS OF OUR ANNUAL REPORT TO SHAREHOLDERS	wmtars-1312016.htm	EX-13
13	LIST OF THE COMPANY'S SIGNIFICANT SUBSIDIARIES	wmt2113116.htm	EX-21
14	CONSENT OF INDEPENDENT REGISTERED PUBLIC ACCOUNTING FIRM	wmt2313116.htm	EX-23
15	CHIEF EXECUTIVE OFFICER SECTION 302 CERTIFICATION	wmt31113116.htm	EX-31.1
16	CHIEF FINANCIAL OFFICER SECTION 302 CERTIFICATION	wmt31213116.htm	EX-31.2
17	CHIEF EXECUTIVE OFFICER SECTION 906 CERTIFICATION	wmt32113116.htm	EX-32.1
18	CHIEF FINANCIAL OFFICER SECTION 906 CERTIFICATION	wmt32213116.htm	EX-32.2
25		fiveyrcumtotalreturn.jpg	GRAPHIC
26		image0a10a01a02a03.jpg	GRAPHIC
	Complete submission text file	0000104169-16-000079.txt	

traditional 10-K that contains most annual report information

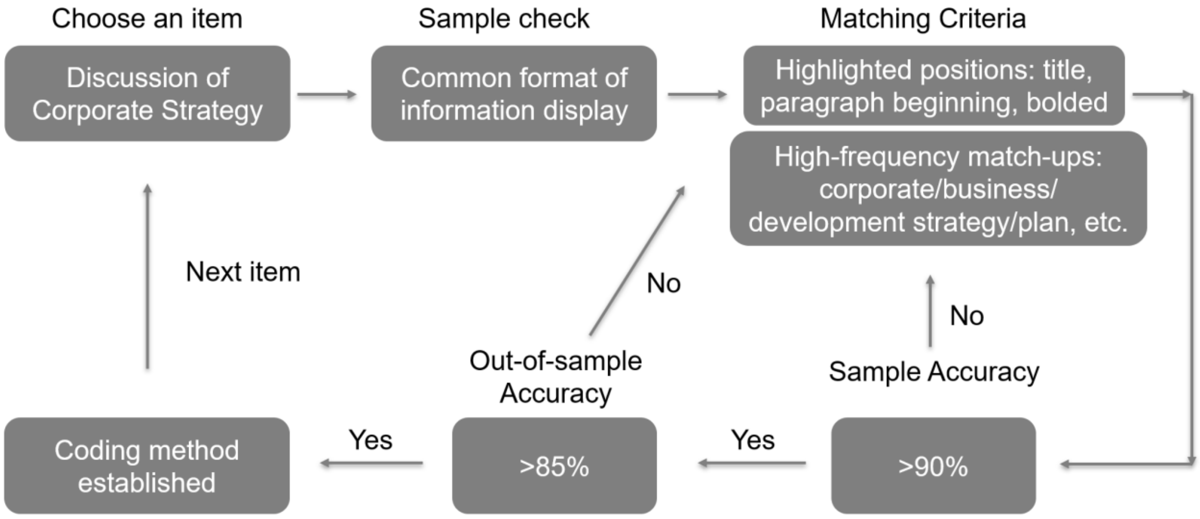
optional form that covers quantitative and MD&A information which should have been included in 10-K

all-in-one file, uploaded solely at the early stage of EDGAR

For the remaining items (I.f to IV.g) in Table 1, we mainly use the text coding method. Specifically, we construct a list of conclusive keywords and collocations for each qualitative item and also require a corresponding number when necessary for each quantitative item. Figure A.3 shows an example of the text coding procedure. For instance, for Item I.g "discussion of corporate strategy", we first check the standard information display format. Next, we summarize the highlighted positions where this information typically shows up (title, paragraph beginning, bolded font) and high-frequency match-up phrases (corporate/business/development strategy/plan, etc.). We require an in-sample accuracy of 90% and an out-of-sample accuracy of 85% for a fully established coding method (200 firms for each). Of the 20 individual items for the text coding procedure, we manually establish and verify the coding method and automatically calculate the final score for each firm in each year.

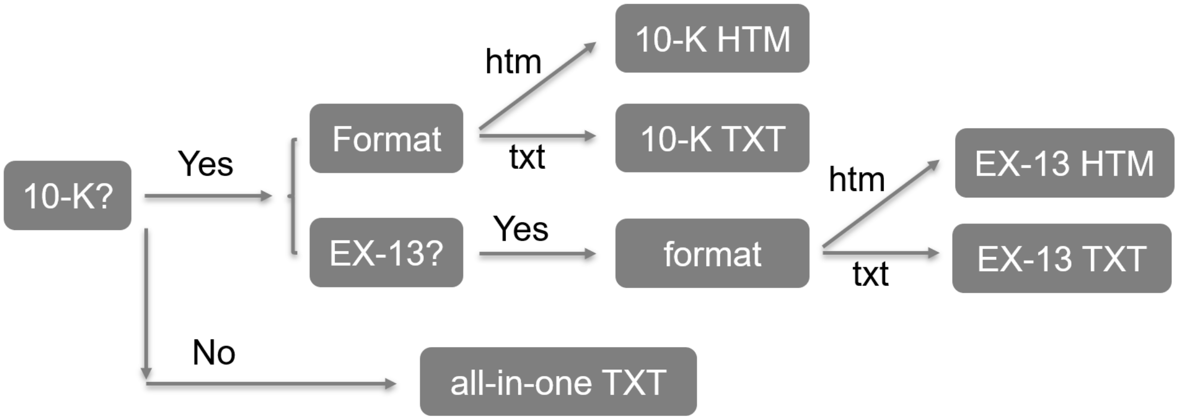
Our textual analysis has some superior features compared to the traditional word searching method. First, our coding scheme has loose matching tolerance. For example, if we confirm "forecasted sales" is a keyword, a sentence like "based on analyst forecasts, our sales value next year will…" is also taken into consideration in our coding scheme. Similarly, selected match-up phrases are captured as long as they are positioned in the same paragraph or within a one-sentence gap as appropriate. Second, our coding method is time-efficient, online analytical, and easily applicable to other disclosure measures such as ESG, CSR, and innovation. We benefit from the multiprocessing frame and improve our running efficiency from 300 hours (single process) to 20 hours for the entire EDGAR universe. To avoid getting banned by EDGAR for excessive requests, we use an online real-time coding frame that can directly generate the disclosure score without downloading the 10-K file for each firm.

FIGURE A.3: An Example of the Text Coding


Choose an item
Discussion of Corporate Strategy
Sample check
Common format of information display
Matching Criteria
Highlighted positions: title, paragraph beginning, bolded
High-frequency match-ups: corporate/business/ development strategy/plan, etc.
Next item
No
No
Out-of-sample Accuracy
Sample Accuracy
Coding method established
Yes
>85%
Yes
>90%

Third, we allow for file-sensitive and format-sensitive table coding schemes. As mentioned above, annual report information can be positioned in different files (10-K, EX-13, and all-in-one file) with different formats (HTM or TXT). Since there is a vast encoding difference between TXT and HMT and a huge layout difference between 10-K, EX-13, and the all-in-one text file, we must reconstruct our table coding scheme for each specific file/format, which ends up with five different coding methods in total. Figure A.4 illustrates how we choose the coding method under each circumstance. For each firm in each year, we first identify whether the firm uploads a 10-K file; if no, we directly go to the all-in-one text file. If a 10-K file is uploaded, we would first check the format of the 10-K file and use the corresponding method to search information encoded in HTM or TXT. If no desired information is found in the 10-K, we would explore the EX-13 form and redo the abovementioned steps. As the EDGAR developed over the years, more and more information tends to be disclosed through the file/format of 10-K HTM.

FIGURE A.4: Different Coding Methods for Different File/Format



In summary, our textual analysis generates a firm-year disclosure measure for each firm that ever submitted a 10-K file to the EDGAR system in that year. We end up with 144,778 firm-year observations from 1994 to 2019, of which a great portion is not included in this study since they are either not covered by Value Line or I/B/E/S or missing formula inputs for the implied cost of equity capital estimation.

TABLE 1 : The Coding Scheme to Analyze 10-K Filings

I. Summary of historical results

a. Return on assets or sufficient information to compute ROA (net income, tax rate, interest expense, and total assets)
b. Net profit margin or sufficient information to compute PM (net income, tax rate, interest expense, and sales)
c. Asset turnover or sufficient information to compute TAT (sales and total assets)
d. Return on equity or sufficient information to compute ROE (net income and total equity)
e. Number of quarters that firm discloses sales and net income
f. Trends in the industry
g. Discussion of corporate strategy

II. Other financial measures

a. Free cash flow (or cash flow other than those reported in SCF)
b. Economic profit, residual income type measure
c. Cost of capital (wacc, hurdle rate, EVA target rate)

III. Nonfinancial measures

a. Number of employees
b. Average compensation per employee
c. Percentage of sales in products designed in the past few (3–5) years
d. Market share
e. Units sold (or other output measure, e.g., production)
f. Unit selling price
g. Growth in units sold (or growth in other output measure, e.g., production)
h. Growth in investment (expansion plans, number of outlets, etc.)

IV. Projected information

a. Forecasted market share
b. Cash flow forecast
c. Capital expenditures, R&D expenditures, or general investment forecast
d. Profit forecast
e. Sales forecast
f. Other output forecast
g. Industry forecast

We use the same coding scheme as in Francis, Nanda, and Olsson (2008), including four major categories and 25 individual items.

TABLE 2 : Summary Statistics of the Disclosure Index

Panel A: FNO’s Disclosure Index

Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Raw Disc (Total)	677	7.98	2.97	4	6	8	10	12
Raw Disc Cat. I	677	2.30	1.50	1	1	2	3	5
Raw Disc Cat. II	677	0.15	0.38	0	0	0	0	1
Raw Disc Cat. III	677	3.18	1.52	1	2	3	4	5
Raw Disc Cat. IV	677	2.35	1.58	0	1	2	3	5

Panel B: Our Disclosure Index in FNO’s Sample

Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Raw Disc (Total)	627	7.95	2.99	4	6	8	10	12
Raw Disc Cat. I	627	2.24	1.54	1	1	2	3	5
Raw Disc Cat. II	627	0.23	0.48	0	0	0	0	1
Raw Disc Cat. III	627	3.20	1.35	2	2	3	4	5
Raw Disc Cat. IV	627	2.28	1.61	0	1	2	3	5

Panel C: Our Disclosure Index in the Full Sample

Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Raw Disc (Total)	43,806	10.27	3.43	6	8	10	13	15
Raw Disc Cat. I	43,806	2.22	1.25	1	1	2	3	3
Raw Disc Cat. II	43,806	0.39	0.58	0	0	0	1	1
Raw Disc Cat. III	43,806	4.07	1.43	2	3	4	5	6
Raw Disc Cat. IV	43,806	3.59	1.78	1	2	4	5	6

This table summarizes the descriptive statistics of FNO’s (Francis, Nanda, and Olsson 2008) disclosure index as well as our disclosure index in FNO’s original sample (fiscal 2001) and the entire Value Line/IBES sample from 1995 to 2019. We report both the total raw disclosure index and individual disclosure indexes for each of the four subcategories. Detailed construction of these individual subcategories can be found in Tabel 1.

TABLE 3 : Inter-correlations of raw disclosure indices

	Total	*Cat. I*	*Cat. II*	*Cat. III*	*Cat. IV*
Raw Disc (Total)	1.0000	0.4687	0.4624	0.7711	0.8428
		0.0000	0.0000	0.0000	0.0000
Raw Disc Category I	0.5036	1.0000	0.0910	0.1031	0.1125
	0.0000		0.0000	0.0000	0.0000
Raw Disc Category II	0.4636	0.1242	1.0000	0.2146	0.3364
	0.0000	0.0000		0.0000	0.0000
Raw Disc Category III	0.7690	0.1606	0.2186	1.0000	0.5554
	0.0000	0.0000	0.0000		0.0000
Raw Disc Category IV	0.8510	0.2074	0.3472	0.5632	1.0000
	0.0000	0.0000	0.0000	0.0000	

This table presents the inter-correlations of raw disclosure indices in total and by category. The lower left shows the Spearman correlation, while the upper right shows the Pearson correlation.

TABLE 4 : Sample Selection Procedures

Panel A: Sample Selection

	Value Line		I/B/E/S	
	N	%	N	%
Firm-years with sufficient data for the implied cost of equity capital	33,837	100	51,464	100
No disclosure score	3,569	10.5	8,896	17.3
Insufficient data for controls	1,984	5.9	5,227	10.2
Total firm-year observations	28,284	83.6	37,341	72.6

Panel B: Sample divided by industry (top 20)

	Value Line		I/B/E/S	
	N	%	N	%
Business Services	2,543	8.99	4,130	11.06
Retail	2,064	7.30	2,258	6.05
Utilities	1,728	6.11	1,581	4.23
Electronic Equipment	1,649	5.83	2,180	5.84
Banking	1,302	4.60	4,182	11.20
Machinery	1,217	4.30	1,355	3.63
Insurance	1,141	4.03	1,758	4.71
Pharmaceutical Products	1,044	3.69	1,029	2.76
Wholesale	976	3.45	1,345	3.60
Petroleum and Natural Gas	971	3.43	980	2.62
Computers	917	3.24	1,316	3.52
Trading	847	2.99	1,456	3.90
Medical Equipment	832	2.94	1,079	2.89
Chemicals	831	2.94	814	2.18
Measuring and Control Equipment	714	2.52	802	2.15
Transportation	701	2.48	1,053	2.82
Construction Materials	692	2.45	683	1.83
Food Products	644	2.28	635	1.70
Automobiles and Trucks	621	2.20	589	1.58
Consumer Goods	532	1.88	592	1.59

This table shows our sample selection procedures and the distribution of firm-year observations by industry. Specifically, we use the Fama-French 48 industry classification and report the top 20 industries based on the Value Line sample.

TABLE 5 : Summary Statistics

Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Raw Disc (Total)	43,806	10.269	3.433	6	8	10	13	15
Disc (Total)	43,806	0.613	0.175	0.357	0.500	0.632	0.737	0.833
ICC_{VL}	28,284	0.142	0.083	0.048	0.086	0.132	0.188	0.255
ICC_{IBES}	37,341	0.098	0.027	0.069	0.080	0.094	0.110	0.132
ICC_{GLS}	41,820	0.063	0.030	0.027	0.043	0.062	0.081	0.098
ICC_{CT}	41,949	0.094	0.032	0.062	0.077	0.091	0.108	0.129
ICC_{OJ}	37,998	0.115	0.032	0.083	0.095	0.110	0.129	0.155
ICC_{PEG}	41,029	0.119	0.049	0.076	0.090	0.108	0.135	0.177
Realized Return	41,668	0.152	0.625	-0.403	-0.156	0.085	0.338	0.677
Beta	43,806	1.125	0.719	0.318	0.620	1.018	1.494	2.060
Idio. Risk	43,806	0.101	0.055	0.048	0.063	0.088	0.125	0.170
Size	43,806	7.148	1.710	4.986	5.927	7.047	8.243	9.467
B/M	43,806	0.535	0.375	0.166	0.282	0.460	0.692	0.977
ln B/M	43,806	-0.858	0.728	-1.796	-1.266	-0.776	-0.369	-0.023
ROA	43,804	0.040	0.081	-0.023	0.011	0.040	0.079	0.123
Analysts	41,585	9.238	7.276	2	4	7	13	20
Audit Fee	32,951	14.083	1.195	12.517	13.296	14.067	14.857	15.640

This table summarizes the descriptive statistics of all variables used in this study. *Raw Disc (Total)* = the raw disclosure score of the 10-K file; *Disc (Total)* = the percentage-based disclosure score scaled by the maximum disclosure index in that year; ICC_{VL} = the Value Line implied cost of equity capital as in Brav, Lehavy, and Michaely (2005); ICC_{GLS}, ICC_{CT}, ICC_{OJ}, ICC_{PEG} are respectively the implied cost of equity capital as described in Gebhardt, Lee, and Swaminathan (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004);ICC_{IBES} = the average of the above four measures of the implied cost of equity capital using the I/B/E/S data; *Beta*: beta coefficients estimated with the CAPM model using a minimum of 24 monthly returns over the 60 months prior to the annual report publication date ; *Idio. Risk*: idiosyncratic risk, measured as the residual volatility of the Fama-French 3-Factor model as in Ang et al. (2006); *Size*: the log of firms' market value of equity in millions of dollars; (*ln*) *B/M*: (the log of) the firm's book-to-market ratio; *ROA*: return on assets; *Analysts*: the number of analysts following the firm; *Audit Fee*: the log of the firms' auditing fees.

TABLE 6 : Yearly Cross-Sectional Regressions

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Size	-0.670***	-1.001***	-1.278***	-1.767***	-1.879***	-0.995***	-0.548***	-0.187	-0.214	-0.128	-0.261	-0.116	0.015
	(-2.74)	(-4.22)	(-6.03)	(-8.64)	(-9.66)	(-4.26)	(-2.99)	(-1.02)	(-1.16)	(-0.77)	(-1.54)	(-0.77)	(0.10)
Beta	1.929***	2.269***	1.845***	3.160***	1.905***	2.253***	3.694***	3.281***	0.349	1.150***	2.312***	1.204***	1.117***
	(2.95)	(3.82)	(3.81)	(6.33)	(3.63)	(3.72)	(9.02)	(10.38)	(1.19)	(4.46)	(8.80)	(4.94)	(4.26)
ln B/M	1.294**	0.462	-0.840*	0.181	1.812***	3.463***	1.421***	0.764*	1.112***	-0.100	0.004	-0.436	0.100
	(2.06)	(0.77)	(-1.71)	(0.38)	(4.54)	(7.78)	(4.11)	(1.96)	(2.73)	(-0.25)	(0.01)	(-1.22)	(0.26)
Disc	1.960	3.546**	4.771***	4.111***	1.275	-0.229	1.987	2.371*	2.008	0.163	0.337	1.814	-0.133
	(1.17)	(2.11)	(3.22)	(2.89)	(0.90)	(-0.14)	(1.59)	(1.80)	(1.45)	(0.13)	(0.26)	(1.42)	(-0.10)
R^2	0.035	0.043	0.055	0.115	0.176	0.147	0.118	0.122	0.020	0.023	0.082	0.034	0.017
Firms	782	858	968	976	1,086	975	1,059	1,116	1,149	1,139	1,112	1,105	1,147

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Size	-0.297*	0.026	0.045	-0.214	-0.271*	-0.364***	-0.489***	-0.219*	-0.419***	-0.438***	-0.311**	-0.495***
	(-1.92)	(0.17)	(0.33)	(-1.47)	(-1.89)	(-3.07)	(-4.22)	(-1.83)	(-3.19)	(-3.34)	(-2.16)	(-3.00)
Beta	2.267***	1.360***	0.747***	2.705***	2.997***	1.747***	0.787***	3.037***	3.593***	1.769***	1.624***	4.013***
	(7.49)	(3.54)	(2.62)	(8.51)	(9.06)	(6.38)	(3.03)	(9.45)	(10.04)	(5.14)	(4.47)	(9.00)
ln B/M	1.220***	0.409	0.968***	1.064***	1.772***	0.566**	0.152	1.004***	0.555**	1.042***	2.145***	3.738***
	(3.84)	(1.33)	(3.25)	(3.45)	(6.22)	(2.36)	(0.64)	(4.68)	(2.45)	(4.45)	(8.56)	(13.46)
Disc	2.527*	2.260	2.346*	1.542	3.113**	-0.147	0.987	3.113**	1.107	-0.046	5.048***	4.388**
	(1.80)	(1.60)	(1.70)	(1.12)	(2.00)	(-0.11)	(0.73)	(2.18)	(0.70)	(-0.03)	(2.88)	(2.22)
R^2	0.087	0.021	0.020	0.083	0.133	0.061	0.030	0.114	0.112	0.059	0.091	0.243
Firms	1,179	1,160	1,205	1,215	1,219	1,245	1,247	1,264	1,291	1,284	1,296	1,207

This table reports the results of cross-sectional regressions of the Value Line implied cost of equity capital (in percentage) on known risk proxies (*Beta*, *Size*, and *ln B/M*), and the percentage-based disclosure indices *Disc*, for each calendar year from 1995 to 2019. The t-statistics are reported in the parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

TABLE 7 : Disclosure and the Implied Cost of Equity Capital (Value Line)

	Implied Cost of Equity Capital		
	(1)	(2)	(3)
Size	-0.499***	-0.388***	-0.358***
	(-3.08)	(-5.76)	(-5.31)
Beta	2.125***	1.399***	1.492***
	(12.11)	(10.47)	(10.54)
ln B/M	0.955***	1.589***	1.450***
	(3.66)	(12.11)	(10.92)
Disc	2.009***	1.634***	1.770***
	(5.95)	(3.44)	(3.72)
R^2	0.082	0.204	0.270
Observations	28284	28284	28218
Year FE		Yes	No
Industry FE		Yes	No
Year*Ind FE		No	Yes
Method	Fama-MacBeth	Panel	Panel

This table reports results from regressing the implied cost of equity capital (in percentage) on known risk proxies (*Beta*, *Size*, and *ln B/M*) and the disclosure indices. *Disc* is the percentage-based total disclosure score. In columns (1), we follow Fama and MacBeth (1973) and report the mean of the annual coefficient estimates; t-statistics are calculated based on the Newey-West consistent standard error. In columns (2) and (3), we report panel regressions with year, industry, and year*industry fixed effects. Robust t-statistics adjusted for firm-level clustering are reported in the parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

TABLE 8 : Disclosure and the Implied Cost of Equity Capital (I/B/E/S)

	Implied Cost of Equity Capital					
	(1) ICC_{IBES}	(2) ICC_{IBES}	(3) ICC_{GLS}	(4) ICC_{CT}	(5) ICC_{OJ}	(6) ICC_{PEG}
Size	-0.259***	-0.324***	-0.144***	-0.195***	-0.423***	-0.645***
	(-10.15)	(-18.77)	(-11.18)	(-9.68)	(-20.27)	(-22.40)
Beta	0.578***	0.480***	0.180***	0.172***	0.583***	1.178***
	(6.15)	(15.25)	(7.05)	(4.20)	(14.84)	(19.23)
ln B/M	1.049***	0.942***	1.451***	0.378***	0.739***	1.266***
	(12.60)	(24.66)	(37.48)	(8.07)	(16.48)	(20.14)
Disc	0.762***	0.660***	0.220**	0.396***	0.858***	1.799***
	(5.32)	(5.65)	(2.47)	(2.79)	(6.00)	(8.28)
R^2	0.195	0.382	0.693	0.224	0.294	0.288
Observations	37341	37287	41774	41902	37944	40989
Year*Ind FE		Yes	Yes	Yes	Yes	Yes

This table reports results from regressing the implied cost of equity capital (in percentage) on known risk proxies (*Beta*, *Size*, and *ln B/M*) and the disclosure indices. *Disc* is the percentage-based total disclosure score. In column (1), we follow Fama and MacBeth (1973) and report the mean of the annual coefficient estimates; t-statistics are calculated based on the Newey-West consistent standard error. In columns (2)-(6), we report panel regressions with year*industry fixed effects. ICC_{GLS}, ICC_{CT}, ICC_{OJ}, ICC_{PEG} are respectively the implied cost of equity capital as described in Gebhardt, Lee, and Swaminathan (2001), Claus and Thomas (2001), Ohlsonand Juettner-Nauroth (2005), and Easton (2004); ICC_{IBES} = the average of the above four measures of the implied cost of equity capital using the I/B/E/S data. Robust t-statistics adjusted for firm-level clustering are reported in the parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

TABLE 9: Relation between disclosure, idiosyncratic volatility, and implied cost of equity capital

	Disc	*Idiosyncratic VOL*	*Implied CoC*
Disc	1.0000	0.1879	0.0527
		0.0000	0.0000
Idiosyncratic VOL	0.1838	1.0000	0.2224
	0.0000		0.0000
Implied CoC	0.0501	0.2515	1.0000
	0.0000	0.0000	

This table presents the correlations between the disclosure index (*Disc*), idiosyncratic volatility (*Idiosyncratic VOL*) measured as the residual volatility from a Fama-French 3-factor model, and Value Line implied cost of equity capital (*Implied CoC*). The disclosure index is detrended at the year level. The lower left shows the Spearman correlation,while the upper right shows the Pearson correlation.