Information Quality and the Expected Rate of Return: A Structural Equation Modelling Approach

Abstract

We use structural equation modelling (SEM) for a robust test of the role information quality plays in explaining the cost of equity capital (CoE). SEM allows us to reliably identify the direct and indirect effects that three information quality attributes, *quantity*, *asymmetry* and *precision*, have on CoE. The method also reduces the error-in-variables problem, which stems from selectivity in proxies for information quality and CoE. Using nine proxies to capture the variation in information quality attributes and nine CoE measures, we document that the direct effects of precision and asymmetry are equally important in explaining variation in CoE, while quantity has a negative direct effect. Quantity has a positive indirect effect on CoE mediated through asymmetry and precision. The strength of the relations we identify varies according to firm size, maturity, profitability and with proxies for CoE, which suggests that sample compositions and measurement choices affect the power of tests. Our results consolidate mixed evidence on the relation between information quality and CoE that is often based on a single measure of information quality and ignores indirect channels.

Keywords: cost of equity; implied cost of capital; information risk; information quality; structural equation modelling

1. Introduction

A large body of accounting literature investigates the extent to which information quality affects the cost of equity capital (CoE). The common proposition in this body of work is that firms with high-quality information environments, as captured by *quantity*, *precision* and *asymmetry*, should enjoy relatively low costs of equity (Easley et al. 2002, Francis et al. 2004, 2005).[1] However, the empirical validity of this assertion remains unclear for two reasons. First, because quantity, precision and asymmetry are not observable, a researcher has little guidance as to which empirical proxies are associated with the latent constructs and, thus, should be used in tests. To illustrate, the bulk of evidence on the association between CoE and information quality is based on tests that use the Dechow and Dichev (2002) accruals model to capture earnings quality (e.g. Francis et al. 2004, 2005, Ecker et al. 2006). However, Wysocki (2009, 1–2) finds that 'accounting quality measures derived from the DD model (and its extensions) show weak and often contradictory associations with other measures of accounting quality for U.S. and international firms' and concludes that 'overall, the DD model does not appear to reliably capture "high quality accruals" and, in some settings, will even reverse rank firms "earnings quality"'. To address the errors-in-variables problem, some studies use multiple proxies for information quality; however, results on their association with CoE vary substantially depending on the measure (Francis et al. 2004, 2005, Core et al. 2008, Mohanram and Rajgopal 2009, Shevlin 2013).

Second, previous studies largely ignore the indirect effects one information quality dimension can have on CoE that is mediated through another factor, e.g. information quantity

[1] Quantity reflects the amount of information available to investors, precision captures information precision, and asymmetry reflects information asymmetry between informed and uninformed investors.

matters only if it has a minimum level of precision.[2] If ignored, indirect effects are captured by the error term; this leads to the omitted correlated variable problem, which casts doubt on prior evidence. The challenges to reliably testing the relationship between CoE and information quality has resulted in an expansive literature where results vary depending on the researcher's choice of information quality proxies and regression specifications. Therefore, there is a need to aggregate existing findings using robust tests to help reconcile prior evidence. The consolidation of evidence is necessary to provide clarity and guidance for future work (Nezlobin et al. 2019).

This study uses structural equation modelling (SEM) to examine the *direct* and *indirect* effects the three information quality dimensions have on CoE. SEM mitigates the omitted correlated variables problem that originates from ignoring the indirect channel and accommodates the limitation arising from selectivity in proxies for information quality and CoE. Specifically, SEM uses factor analysis to identify shared variability between proxies for a specific information quality dimension, e.g. precision, and for CoE. This also mitigates bias in the estimated coefficients emanating from the errors-in-variable problem, increasing the empirical robustness of the results (Rao 1973).[3]

We employ three proxies for each information attribute and nine CoE measures, which include three valuation model–based estimates (i.e. the implied cost of capital measures), three proxies based upon risk factor models, and three proxies based on market sentiment augmented risk factor models. As with the independent variables, only the principal component from the CoE measures is retained. This point is important because there is little consensus in the literature on how to measure quantity, precision, asymmetry or CoE (Botosan and Plumlee 2005,

[2] To illustrate, Lambert et al. (2012) argue that Easley and O'Hara's (2004) negative association between information asymmetry and CoE is driven by changes in information precision.
[3] In other words, SEM reduces the likelihood of Type II error. Section 2 discusses SEM in more detail, particularly how it differs from principal component analysis and structural modelling.

Easton and Monahan 2016). SEM then uses the principal components to model indirect paths that link the latent information quality dimensions quantity, precision and asymmetry and direct paths that link the information quality dimensions to CoE. Because the direct and indirect effects are estimated simultaneously, SEM reduces the endogeneity caused by simultaneity (Hinson and Utke 2018) and – compared to equation-by-equation estimation – improves efficiency and produces unbiased estimates (Bollen 2014, Kline 2015).

The first part of the study models the associations between the nine empirical proxies and the respective information quality dimensions. The three proxies for asymmetry, probability of informed trading, bid-ask spreads and institutional investor concentration show a comparable contribution to common variation. Here, stock listing duration is the strongest predictor of quantity compared to media exposure intensity and firm age, explaining 76.7% of the common variance among the three proxies. Precision is largely explained by accrual quality and the value relevance of earnings, and it shows little association with analyst forecast accuracy.[4] Our results help explain the mixed evidence on the association between information quality proxies and CoE (e.g. Core et al. 2008, Mohanram and Rajgopal 2009, Shevlin 2013), as the power of the tests depends on the choice made by researchers regarding proxies for quantity, precision and asymmetry.

Next, we model the interrelations between the three information quality dimensions to identify the indirect paths through which they can affect CoE. Quantity has a significant impact on precision and asymmetry. The former result is consistent with the ability of investors to

[4] The lack of an association between precision and analyst forecast accuracy may reflect that analyst forecast accuracy is a 'catch-all' measure for the overall quality of a firm's information environment, rather than a proxy for precision alone. This result is intuitive as analyst forecast accuracy not only depends on the precision of information, but also on analysts' access to private compared to public information, which reflects the level of asymmetry, and on the amount of information available to them, which reflects quantity.

improve their assessment of future cash flow and discount rates news as the number of (perfectly uncorrelated) signals increases, even when the average precision of individual signals is low (Lipe 1986, Wild 1992, Ramakrishnan and Thomas 1998). The positive association between quantity and asymmetry reflects that, as the level of publicly available information increases, the cost of private information acquisition increases, which reduces information asymmetry between informed and uninformed investors (Welker 1995, Healy et al. 1999, Heflin et al. 2005). Further, we document a significant association between precision and asymmetry, which captures how higher public information precision reduces the information advantage of better-informed investors (Brown et al. 2004, Brown and Hillegeist 2007). Jointly, these results suggest that quantity and precision may have potential important indirect effects on CoE that are mediated through other dimensions.

The second part of the study examines the direct and indirect channels through which quantity, precision and asymmetry affect CoE. We document that quantity has a positive direct effect on CoE; however, the indirect effects mediated through precision and asymmetry are almost two and a half times as important, which produces an overall negative association between quantity and CoE. This result suggests that quantity's effect on CoE – as identified in previous studies – largely stems from its indirect associations with precision and asymmetry. Accounting for both indirect and direct channels, one standard deviation increase in quantity leads to a 10% smaller CoE. The opposite signs of indirect compared to direct associations between quantity and CoE are consistent with empirical evidence, e.g. Li (2008) finds that loss-making firms with lower earnings persistence and transient income increase the length of their annual reports and use longer sentences and complex words. This serves to obfuscate 'bad' news by making it more costly for investors to analyse the annual report, which reduces the price impact of disclosing bad news. Lawrence (2013) finds that individual investors' holdings decrease with longer and more complex financial disclosure, and Ertugrul et al. (2017) demonstrate that larger 10-K file size, a proxy for report length, is associated with stricter loan contract terms.

Precision has a strong direct association with CoE, and a small indirect effect mediated through asymmetry that accounts for 10% of the total effect precision has on CoE. A one standard deviation reduction in precision leads to a 52% lower CoE, an economically significant reduction. Finally, we find that firms that enact effective measures to reduce information asymmetry between privately informed and publicly uninformed investors have lower CoE. A one standard deviation reduction in asymmetry is associated with a 44.1% smaller CoE. Jointly, the results provide clarity as to the magnitude and direction of the effects quantity, precision and asymmetry have on CoE.

A sensitivity analysis shows that the strength of the associations vary according to firm characteristics such as firm size, maturity and profitability, which suggests that sample compositions affect the power of tests. To illustrate, precision's effect on CoE is almost two times greater for smaller than larger firms, and the effect of asymmetry is stronger in the latter sample period (2000–2010) than in the earlier sample period (1993–1999), which coincides with the increasing ownership by (on average better informed) institutional investors. The sensitivity results are important, given the high levels of heterogeneity in previous studies' samples. For example, Ecker (2014) focuses on the association between quantity and CoE only for IPO stocks, and Francis et al. (2005, 306) employ a sample that 'is restricted to firms with at least 7 years of data', which excludes smaller and younger firms. Further, we find that using risk factor measures alone to capture CoE tends to produce either statistically insignificant associations between information quality and CoE or of the opposite sign. These results are consistent with the conclusion in Fama and French (1997, 154) that the unexpected news component in realised returns tends to corrupt factor loadings and risk premia, which results in 'woefully imprecise estimates of the cost of equity', and that valuation model–based estimates outperform risk factor measures as proxies for CoE (Botosan and Plumlee 2005, Lee et al. 2010, 2015, Botosan et al. 2011). Using valuation model–based estimates or combining them with risk factor models produce results consistent with theoretical associations.

This study makes three contributions to the literature. First, this is the first study to provide simultaneous evidence on how quantity, precision and asymmetry affect CoE within a uniform structural equation modelling framework. Our results unify and consolidate previous evidence on the association between information quality and CoE and calibrate the effects quantity, asymmetry and precision have on CoE. This evidence is important given the numerous, often conflicting results that are frequently based on a single measure of information quality and CoE and ignore the indirect effects mediated through other dimensions. Further, we show how firm characteristics, such as firm size and age, affect the strength of the relationship between information quality dimensions and CoE. This evidence helps explain how sampling differences can produce inconsistent evidence and contributes novel insights to the debate on the pricing of information risk. Our study responds to the call in Beyer et al. (2010) for more research to disentangle the underlying complexity between firm-specific information and expected rates of return.

Second, our use of SEM addresses the concern that the subjective selection of proxies and errors-in-variables affects empirical results. To illustrate, Bhattacharya et al. (2012) measure information asymmetry as both the adverse selection component of the bid-ask spread and the probability of informed trading. However, Van Ness et al. (2001, 77) find that the adverse selection component of the bid-ask spread correlates weakly with multiple measures of information asymmetry, such as analyst forecast errors and market-to-book ratio, concluding that 'adverse selection models measure adverse selection weakly at best'. Lai et al. (2014) find no evidence that the probability of informed trading [PIN] is associated with expected returns in their sample of 47 countries. These results question the validity of employing PIN and bid-ask spread as information asymmetry measures, thus challenging the reliability of the conclusions in Bhattacharya et al. (2012) . However, this does not preclude that PIN and bid-ask spread share a common variation that captures information asymmetry. SEM acknowledges that information quality is latent and indirectly measurable and uses factor analysis to estimate *common variation* based on multiple proxies, which increases the validity of empirical tests.

Third, our analysis helps calibrate the effect each information quality dimension has on CoE. We find that information precision and asymmetry have the highest total effects (sum of indirect and direct channels) on CoE – these are almost five times stronger than the effect of quantity. This result is important for managers deciding where to allocate resources to reduce their firm's CoE. Further, our results suggest that increasing information quantity can have a negative direct effect on CoE, but a positive indirect effect mediated through asymmetry and precision. In contrast, improving information precision and reducing information asymmetry will have a direct positive effect on reducing CoE. Our results help guide future work on the association between information quality and CoE by identifying the sign and importance of each information quality dimension in explaining the variation in CoE.

2. Theoretical Background and Research Hypotheses

Direct links between precision, asymmetry, quantity and CoE are supported by analytical work conducted by Easley and O'Hara (2004) and Lewellen and Shanken (2002). Indirect links are suggested by empirical evidence in Bhattacharya et al. (2013), Brown and Hillegeist (2007), Francis et al. (2008), Lang and Lundholm (1996) and Welker (1995). However, no study to date has attempted to directly model these links. This section reviews the literature on the association between the three information quality dimensions, precision, asymmetry and quantity, and CoE, and the most commonly used proxies.

2.1 Information Precision and Cost of Equity

Theory predicts that a firm's CoE decreases with the accuracy of available information about the future value of the firm (e.g. Easley and O'Hara 2004, Li 2005, Bloomfield and Fischer 2011), and this prediction has received much attention in the literature. One literature strand uses earnings quality metrics as a proxy for precision. Francis et al. (2004, 2005) report that CoE decreases as earnings quality increases, with the effects of accrual quality and value relevance metrics dominating other measures. Additional studies that report a positive association between earnings quality and CoE include Aboody et al. (2005), Barth et al. (2013), Francis et al. (2005), Gray et al. (2009), Kim and Qi (2010) and Ogneva (2012). However, Core et al. (2008) question the validity of this evidence and find that accruals quality is not priced. Ogneva (2012) reconciles the evidence and reports that accruals quality is positively associated with future cash flow shocks, which obfuscates the association between accruals quality and realised returns.

The second research strand uses the forecasts of security analysts as a proxy for information precision, with higher forecast accuracy indicating lower information uncertainty (Barry and Brown 1985, Barron and Stuerke 1998, Barron et al. 1998). Barron et al. (2005) show that the precision of analyst forecasts is a reliable measure for the general information precision of investors. The standard proxy for precision in this literature is analyst total forecast precision, a sum of public and private information precision, which is negatively associated with CoE (Barry and Brown 1985, Barron and Stuerke 1998, Barron et al. 2012, Botosan and Plumlee 2013).

2.2 Information Asymmetry and Cost of Equity

Easley and O'Hara (2004) propose that, as the fraction of uninformed investors and the number of private signals about the value of the firm increases, a firm's CoE decreases. This result is consistent with uninformed investors requiring compensation when trading against privately informed investors. Caskey et al. (2015), Hughes et al. (2007) and Lambert et al. (2007) also propose that greater asymmetry increases CoE by means of higher factor risk premia. The empirical literature examining this prediction has two streams.

The first stream uses market microstructure proxies to measure information asymmetry, with bid-ask spreads and the probability of informed trading used most commonly. Broad evidence suggests a positive relationship between these proxies and firms' CoE (Amihud and Mendelson 1986, Easley et al. 2002, Duarte et al. 2008, Easley et al. 2010, Bhattacharya et al. 2012, Levi and Zhang 2015, Brennan et al. 2016). However, Hughes et al. (2007) argue that investor under-diversification in a finite economy will produce evidence, consistent with Easley and O'Hara (2004). Lambert et al. (2012) maintain that Easley and O'Hara's (2004) negative association between information asymmetry and CoE is driven by changes in average information precision. Mohanram and Rajgopal (2009, 241) conclude that 'there is not much evidence to support the interpretation that information risk, proxied by PIN, is a source of priced information risk'. They acknowledge that while their paper suggests that 'PIN is not priced risk, it is difficult to make more general statements about the pricing of information risk since information risk can […] be proxied by different empirical variables' (Mohanram and Rajgopal

2009, 241).

The second literature stream proposes ownership-based measures, arguing that asymmetry decreases as market competition increases (Armstrong et al. 2011, Akins et al. 2012, Lambert et al. 2012). The underlying idea is that greater competition among informed investors leads to a quicker reflection of private information in prices, which reduces informational disadvantages between investors. Thus, greater investor concentration is associated with less competition and thus greater information asymmetry.

2.3 Information Quantity and Cost of Equity

If the amount of information about a firm is low (e.g. limited corporate disclosure, lack of analyst reports), investors encounter difficulties in accurately estimating return and cash flow parameters, which leads to higher CoE (Clarkson et al. 1996, Lewellen and Shanken 2002, Zhang 2006, Kumar et al. 2008). Ecker (2014) examines this prediction in an IPO setting, where information is naturally limited, and confirms that investors have initial difficulties in specifying the return-generating process for newly listed firms, as indicated by significant post-IPO abnormal returns. As information becomes more available over time, parameter uncertainty resolves, and abnormal returns disappear.

One stream of research testing the association between information quantity and CoE uses firm age and length of listing to measure information quantity and demonstrates that these proxies are negatively correlated with CoE (Barry and Brown 1984, 1985, Clarkson and Thompson 1990, Clarkson and Satterly 1997, Lee et al. 2003, Ecker 2014). A second stream uses firms' disclosure levels as proxies for quantity, which tends to have a negative association with CoE (Botosan 1997, Healy et al. 1999, Kothari et al. 2009, Baginski and Rakow 2012, Fu et al. 2012, Campbell et al. 2014, Cao et al. 2017).

2.4 Interrelationship between Information Attributes

Previous studies suggest a positive association between information quantity and information precision and indicate that causality runs from quantity to precision. For instance, disclosure levels tend to be positively associated with analyst forecast precision (Lang and Lundholm 1996, Byard and Shaw 2003), which supports the notion that greater disclosure leads to higher levels of accuracy in analyst forecasts, i.e. higher information precision. In addition, disclosure levels also show a significant association with earnings quality, although the sign of this association is unclear (Imhoff 1978, Waymire 1985, Cox 1985 and Francis et al. 2008 find a positive relationship, while Lang and Lundholm 1993 and Tasker 1998 find a negative one).

Empirical evidence proposes that as the quantity of information about a firm increases, information asymmetry decreases – with the direction flowing from quantity to asymmetry (Healy et al. 1999, Welker 1995, Brown et al. 2004, Heflin et al. 2005, Brown and Hillegeist 2007). Further, recent findings corroborate the proposition that asymmetry is negatively associated with precision, where causality is assumed from the former to the latter. Affleck-Graves et al. (2002), Bhattacharya et al. (2013) and Jayaraman (2008) find an inverse relationship between information asymmetry and analyst forecast precision and earnings quality. Furthermore, Bhattacharya et al. (2012, 472) conclude that ‘there appears to be a limited (in magnitude) feedback path from information asymmetry to earnings quality’. These findings conform to the notion that more precise public information lowers investors’ incentive to acquire additional private information, thereby reducing information asymmetry (Diamond and Verrecchia 1991).

This literature review illustrates the complexity of the field, which is marred by multiple proxies, inconsistent evidence and a failure to distinguish between direct and indirect effects.

Our study is the first to propose a uniform framework to examine the effect the three information quality dimensions have on CoE and to directly model both direct and indirect association channels.

2.5 Structural Equation Modelling vs Principal Component Analysis and Structural Modelling

SEM offers clear advantages over an OLS regression as it (1) identifies both direct and indirect effects between information qualities and CoE, (2) accommodates multiple proxies for information qualities and CoE and (3) is less affected by measurement error in proxies for quantity, precision and asymmetry, and CoE, which reduces the error-in-variables problem.

SEM should not be confused with principal component analysis, which is a data reduction method that creates a weighted factor from a predetermined set of variables. The resultant composite maximises variance from individual components, including the measurement errors inherent in the individual proxies. Thus, the composite may exhibit less correlation with the underlying latent variable than individual proxies, which reduces the validity of empirical tests (Hinson and Utke 2018). Bandalos and Boehm-Kaufman (2009) emphasise that the principal component does not represent the underlying latent construct and caution that researchers should interpret their results with care.[5] Importantly, principal component analysis does not examine the paths between variables. In contrast, SEM creates a common factor from proxy measures based on *common variance* and, in the process, removes the idiosyncratic error

[5] The partial least squares (PLS) technique is similar to principal component analysis, as the method first applies principal component analysis to the data and then performs least square regressions on the components (Huang et al. 2015, Giglio and Xiu 2017, Giovannelli, et al. 2018). Thus, PLS suffers from the same drawbacks as principal component analysis.

of individual variables from the common factor. The resultant factor can be interpreted as capturing the latent construct of interest.

SEM is also distinct from path analysis, which does not model the common factor underlying a latent variable. Further, path analysis fails to allow for simultaneous equations estimation, and complex path modelling generally increases Type I errors (LeBreton et al. 2009). Path analysis is also unsuitable for identifying multiple indirect effects (Hinson and Utke 2018).

There is a further distinction between SEM and structural modelling, which uses theory to model relationships among variables. These relations are later estimated empirically using statistical techniques, such as logistic regressions (Gow et al. 2016). Finally, Busemeyer and Jones (1983) argue that SEM is more likely to detect significant associations compared to models that rely on interaction terms as they suffer from the error-in-variable problem that is magnified by interacting noisy measures.

3. Methodology

This section presents the definitions of the proxy variables we use to capture precision, asymmetry and quantity. We also explain the methods we use to calculate CoE.

3.1 Information Precision

We measure information precision by accrual quality, earnings value relevance and the precision of analyst forecasts (Francis et al. 2008, Bhattacharya et al. 2012). Earnings quality metrics are a natural choice for precision, given that investors regard earnings as an important indicator of a firm’s future financial performance (Biddle et al. 1995, Liu et al. 2002). The accrual quality metric is based on the McNichols (2002) modification of the Dechow–Dichev (2002) model, which has the form:

$$TCA_{i,t} = \beta_{0,i} + \beta_{1,i}CFO_{i,t-1} + \beta_{2,i}CFO_{i,t} + \beta_{3,i}CFO_{i,t+1} + \beta_{4,i}\Delta Sales_{i,t} + \beta_{5,i}PPE_{i,t} + \varepsilon_{i,t}, \qquad \textbf{(1)}$$

where *TCA* are total accruals, *CFO* are operating cash flows, Δ*Sales* is change in sales and *PPE* is the gross value of property, plant and equipment. We then measure accrual quality, *AQ*, as the standard deviation of each firm's residual ($\varepsilon_{i,t}$) over the past four years, annually estimating equation (1) for each of the 48 Fama–French industries with at least 20 firms in each year. We multiply *AQ* by -1 to establish a more intuitive relationship with precision: the lower the information precision of a firm, the lower its accrual quality.

The value relevance of earnings is expressed as the degree to which both a firm's earnings and the change in its earnings explain stock returns, where greater explanatory power indicates more transparent and value relevant earnings, respectively (Francis et al. 2004, Barth et al. 2013). Specifically, we estimate earnings value relevance, *VR*, as the residual from regressing returns on earnings and changes in earnings:

$$RET_{i,t} = \beta_{0,i} + \beta_{1,i}NIBE_{i,t}/MV_{i,t-1} + \beta_{1,i}\Delta NIBE_{i,t}/MV_{i,t-1} + \varepsilon_{i,t}, \qquad \textbf{(2)}$$

where *RET* are firm *i*'s continuously compounded 15-month returns ending three months after the end of fiscal year *t*, *NIBE* is net income, Δ*NIBE* is change in net income and *MV* is the market value of equity. The lower the value relevance metric of earnings, the lower the precision.

Barron et al. (1998) show that observable characteristics of analyst forecasts, namely forecast dispersion, squared error in the mean forecast and the number of forecasts, can be used to infer the degree of public and private information precision. The sum of public and private information precision, *AFP_Total*, reflects the total precision of analysts' information and captures the information precision of sophisticated investors (Barron et al. 2005). To calculate

AFP_Total, we first calculate the precision of analyst public and private signals as in Barron et al. (2005):

$$public = \frac{(SE - \frac{D}{N})}{\left[\left(SE - \frac{D}{N}\right) + D\right]^2} \quad \textbf{(3)}$$

$$private = \frac{D}{\left[\left(SE - \frac{D}{N}\right) + D\right]^2} \quad \textbf{(4)}$$

where SE is the squared error in the mean forecast $(\bar{F}_{it} - A_{it})^2$, D is forecast dispersion calculated as $\frac{1}{N-1}\sum_{i=1}^{N}(\bar{F}_{it} - F_{ijt})^2$, N is the number of forecasts, $\bar{F}_{it}$ is the mean forecast for firm *i* in quarter *t*, A_{it} are the actual earnings for firm *i* in quarter *t* and F_{ijt} is analyst *j*'s forecast of the earnings for firm *i*, in quarter *t*. *AFP_Total* is then calculated as the time-series average of the quarterly values for *public* and *private* over the three quarters q-1 to q+1, where q equals the quarter in which the fiscal year ends. Using time-series averages of lead, lag and current observations centred around the fiscal year-end reduces the noise in the proxies surrounding the release of fiscal year-end information and is consistent with past approaches (Bhattacharya et al. 2012, Botosan and Plumlee 2013). Firm-quarter observations for which *D* is zero, indicating stale forecasts, are excluded, and at least two unique analyst forecasts are required for each firm. Our estimation of *AFP_Total* is consistent with Botosan and Plumlee (2013), and *AFP_Total* should have a positive relationship with precision.

3.2 Information Quantity

Following previous studies, Barry and Brown (1984, 1985), Clarkson and Thompson (1990), Clarkson and Satterly (1997), Ecker (2014) and Lee et al. (2003), we use three information quantity measures: period of listing (*Listing*), firm age (*Age*) and relative media coverage (*RMC*).

The period of listing is the number of years since a firm's initial public offering, and firm age is the number of years since its incorporation. The media coverage index captures abnormally high firm media prominence (Beattie et al. 2004, Beretta and Bozzolan 2008). Greater media attention should correlate with higher information quantity as information about the firm becomesmore easily available (Kross and Schroeder 1989). We calculate *RMC* as firm's *i* residual in year *t* from regressing media coverage (*MC*) on a firm's industry indicator (*IND*) and firm size (*LNSIZE*) while controlling for time effects *T*:

$$\widehat{MC}_{i,t} = \beta_0 + \sum_{j=1}^{T-1} \tau_j T_j + \sum_{j=1}^{k-1} \beta_j IND_j + \beta_k LNSIZE_{i,t} + \varepsilon_{i,t}. \quad \textbf{(5)}$$

We calculate media coverage as the number of news entries per firm-year on Factiva, and we use four-digit SIC codes to assign each firm into one of the 48 Fama–French industries. Firm size is the natural logarithm of sales. The *RMC* index is positive for firms that enjoy more media coverage than the average firm in the same industry and of a similar size. *Listing*, *Age* and *RMC* should be positively associated with quantity.

3.3 Information Asymmetry

We use bid-ask spreads (*SPREAD*) and the probability of informed trading (*PIN*) as two microstructure proxies for information asymmetry (Brown et al. 2004, Bhattacharya et al. 2012, Li et al. 2016). Consistent with Stoll (1978), we calculate *SPREAD* as the average daily percentage spread over 252 trading days that are centred around the fiscal year-end date (i.e. *t-126* to *t+125*):

$$spread = \frac{|P^B - P^A|}{\frac{1}{2}(P^B + P^A)} \quad \textbf{(6)}$$

where $|P^B - P^A|$ is the absolute difference between closing bid (P^B) and closing ask (P^B) prices. Copeland and Galai (1983) and Glosten and Milgrom (1985) show that bid-ask spreads serve as a proxy for the exposure of market-makers to adverse selection and can capture the degree of information asymmetry between informed and uninformed investors.

The probability of informed trading, *PIN*, measures the likelihood that the next trade order is from a privately informed investor; a larger *PIN* score signifies a higher level of information asymmetry (Easley et al. 1996, 1997, Brown and Hillegeist 2007). The underlying notion of *PIN* is that, while it is impossible to directly observe which trades are based on private information, one can use imbalances between buy and sell orders to infer the probability of information-based trading for a given stock. *PIN* is calculated as the time-series average of quarterly *PIN* values over three quarters, q-1 to q+1, where q is the fiscal year-end quarter. Firm quarters for which the number of trading days falls below 30 are excluded.

We complement the two microstructure proxies with an investor concentration measure, *INV_Conc*, to capture investor competition. The estimation of *INV_Conc* is similar to Akins et al. (2012), and we use information on mutual fund holdings, which include the largest and most prominent institutional investor groups (Falkenstein 1996). Specifically, we first calculate quarterly investor concentration *INV_Conc_q* as

$$INV\_Conc\_q = \sum_{j=1}^{N} \left(\frac{Fund_{i,j}}{Fund_i} \right)^2 \qquad \textbf{(7)}$$

where $Fund_{i,j}$ is the number of shares held by mutual fund *j* in firm *i*, $Fund_i$ is the total number of shares held by all mutual funds in firm *i* and *N* is the total number of mutual funds invested in firm *i*. We then average quarterly *INV_Conc_q* over the three quarters q-1 to q+1 to calculate *INV_Conc*. Higher values of *INV_Conc* denote less competition in the trading of a firm's stock, and *INV_Conc* should be positively associated with information asymmetry.

3.4 Cost of Equity

There is no consensus among researchers regarding how to measure CoE. Early studies use risk factor models to estimate expected return (Elton, 1999); however, Fama and French (1997, 154) conclude that risk factor CoE proxies are ‘woefully imprecise estimates of the cost of equity’. More recent literature infers CoE from valuation models, where the key input are analyst consensus earnings forecasts (Easton and Monahan 2005, Guay, Kothari and Shu 2011). We calculate nine proxies for a firm’s CoE: three risk factor–based measures (RFB), three market sentiment augmented proxies (FVIX) and three valuation model–based (VMB) estimates, i.e. the implied cost of capital.

We use *rCAPM*, *rFF3* and *rFF4* as our RFB proxies (Fama and French 1993, Carhart 1997). Specifically, *rCAPM* is calculated as:

$$r_{CAPM} = \bar{r}_{f,t} + \hat{\beta}_{RMRF,i,t}\left(\overline{R_M - R_f}\right)_t \quad \textbf{(8)}$$

where $\bar{r}_{f,t}$ is the expected annual risk-free rate and $\left(\overline{R_M - R_f}\right)_t$ is the annual market risk premium. We calculate the annual risk-free rate and market risk premium by compounding 12 months of returns and estimate equation (8) using 12-month rolling windows to avoid outdated estimates. $\hat{\beta}_{RMRF,i,t}$ is the market beta estimated over a 60-month window using rolling time-series regressions (Barth et al. 2013), *rFF3* expands the CAPM by including the size (SMB) and book-to-market factors (HML) and *rFF4* includes the momentum factor. The approach for calculating *rFF3* and *rFF4* is similar to *rCAPM*.

Recent studies suggest that investor sentiment has an important impact on the cost of capital (Tetlock 2007). As a result, we re-estimate *rCAPM*, *rFF3* and *rFF4* with an additional risk factor for expected market volatility (i.e. FVIX) to capture market sentiment: *rFVIX*, *rFVIX3* and *rFVIX4* (Ang et al. 2006). The underlying notion is that companies with higher negative sensitivity to VIX index changes have higher CoE. FVIX reflects the monthly excess return on a factor-mimicking portfolio that tracks daily changes in the VIX index (Barinov 2013).

Because previous studies document that RFB proxies tend to be imprecise estimates of the cost of equity (Fama and French 1997, Ferson and Locke 1998), we also estimate three implied cost of capital (ICC) estimates: *rPE*, *rPEG* and *rAEGM* (Easton 2004). The price-earnings-ratio-based implied CoE is calculated as:

$$r_{PE} = \left(\frac{P}{eps_1}\right)^{-1} \qquad \textbf{(9)}$$

where eps_1 is the one year ahead earnings consensus forecast and *P* is the current stock price. Observations for which eps_1 is negative are excluded to avoid negative CoE. *rPEG* is the price-earnings-growth CoE, calculated as:

$$r_{PEG} = \sqrt{\frac{eps_2 - eps_1}{P}} \qquad \textbf{(10)}$$

where eps2 are earnings forecasts for two years ahead. All observations in which eps1 is larger than *eps2* are excluded. *rAEGM* is the abnormal earnings growth of CoE, calculated as:

$$r_{AEGM} = A + \sqrt{A^2 + \frac{eps_1}{P}\left(\frac{eps_2 - eps_1}{eps_1} - G_{AEG}\right)} \qquad \textbf{(11)}$$

where $A = \frac{1}{2}\left(G_{AEG} + \frac{dps_1}{P}\right)$, and G_{AEG} is the perpetual growth rate in abnormal earnings set to the current expected annual risk-free rate, minus three percent (Easton 2004). The expected annual risk-free rate is calculated by first taking the average of the one-month U.S. Treasury bill rate over rolling past 12-month windows and then annualising it by compounding it over 12 months. *dps1* is analysts' mean dividends per share estimate for *t+1*. Missing dividend forecasts are estimated as *eps1* × *current dividend payout ratio*, where the dividend payout ratio is equal to $\frac{Dividends\ (Common)}{Income\ Before\ Extraordinary\ Items}$ for firms with positive earnings and $\frac{Dividends\ (Common)}{0.06\ x\ Total\ Assets}$ for firms with currently negative earnings (Easton 2004). Risk factor–based proxies are calculated six months after the end of a firm's fiscal year, and valuation model–based proxies are derived from the first available analyst consensus forecast after a firm's earnings announcement. Table 1 summarises the proxies for precision, asymmetry, quantity and CoE.

(Table 1 about here)

4. Sample and Data Selection

The sample includes all NYSE, AMEX and NASDAQ listed securities that are covered on the CRSP/Compustat Merged (CMM) database and have fiscal year ends from 1993–2010. The accounting information is drawn from Compustat, return data from CRSP and analyst forecasts from I/B/E/S. We use SDC Platinum, Osiris and CRSP to identify founding and listing dates and Factiva to count the number of news articles. CDA/Spectrum (s12) is the source of institutional holdings data. Eighteen years is the longest possible sampling period for the required information. The final sample includes 60,995 firm years for 7,091 unique firms.

Table 2 reports the descriptive statistics for our sample. The average firm is incorporated for 27.8 years and listed on one of the three major U.S. exchanges for 14.7 years. These figures are similar to the IPO age data provided on Jay Ritter's website. A negative median value for the *RMC* index suggests that firms in the sample tend to enjoy less media coverage than the average company in the same industry and of similar size. To provide a simple measure of how well *Listing*, *Age* and *RMC* capture quantity, we calculate Cronbach's alpha. Small Cronbach's alphas indicate low internal consistency among measurements (Little et al. 1999, Kline 2015). *Listing*, *Age* and *RMC* are significantly positively correlated, with a Cronbach's alpha of 0.66.[6]

The values for earnings quality measures are comparable to previous studies. An average *AQ* metric of -0.061 and *VR* metric of -0.410 are similar to Francis et al. (2004, 2005). Our mean of 1,059 for total analyst forecast precision (*AFP_Total*) is comparable with Botosan and Plumlee (2013). The Cronbach's alpha for AQ and VR is 0.54, though it is only 0.43 when including *AFP_Total*, which suggests that *AFP_Total* may be a poor proxy for precision.

The mean *PIN* score of 0.205 falls within the range of 0.150 to 0.300 in previous studies (Easley et al. 2002, Brown and Hillegeist 2007, Duarte et al. 2008, Bhattacharya et al. 2012), and the sample mean *SPREAD* is comparable with Corwin and Schultz (2012). Our average *INV_Conc* is similar to Akins et al. (2012). All three measures are significantly positively correlated, with a Cronbach's alpha of 0.65.

Summary statistics for the VMB and RFB are also consistent with former evidence: an average *rPEG* of 0.124 is similar to Barron et al. (2012) and Easton and Monahan (2005), and a mean *rFF4* of 0.103 is in line with Barth et al. (2013) and Kothari et al. (2009). The VIX-augmented RFB proxies *rFVIX*, *rFVIX3* and *rFVIX4* are equal to their VIX-free counterparts in terms of average levels but with somewhat greater standard deviations. The Cronbach's alphas for RFB measures are 0.77: 0.75 for FVIX and 0.32 for VBM measures. High Cronbach's alphas for RFB and FVIX reflect they share the same component, market risk premium, which explains a significant portion of the return variation (Fama and French 1993). VBM measures are, by construct, more heterogenous, which explains the lower Cronbach's alpha and higher value range.

[6] Before estimating the SEM models, we standardise all variables to mean zero and variance one to make coefficients comparable between variables. Cronbach's alphas are based on standardised variables.

4.1 SEM Estimation Results

SEM estimation proceeds in three steps. First, the model performs factor analysis to capture the impact proxy variables have on each information quality dimension, e.g. the effect *Listing, Age* and *RMC* have on quantity. Second, it captures the common variation in CoE proxies. Third, it uses the resultant factor component to examine the association between information quality dimensions and CoE. Figure 1 illustrates (1) the relationship between the proxy variables and information quality dimensions and CoE, (2) the predicted interrelationship between quantity, precision and asymmetry, and (3) the predicted effects the three information quality dimensions have on CoE. To facilitate the presentation of our estimation results, we map Figure 1 to SEM estimation output and report the regression results in Figure 2 with coefficient estimates and t-statistics included in parentheses.

(Figure 1 about here)

4.2 The Impact of Proxy Variables on Quantity, Precision, Asymmetry and CoE

Figure 2 reports the SEM regression results mapped to Figure 1. We standardise all variables to mean zero and a variance of one to allow for the easier comparability of coefficients. Standard errors are adjusted for firm clusters. We document that all information quality proxies are positively correlated with their respective information attributes. Quantity loads significantly positively on *Listing, Age* and *RMC* with most variance explained by *Listing* (R^2 = (coefficient estimate)^2 = $(0.876)^2$ = 0.767) and *Age* (R^2 =0.385).[7] Factor loadings on *PIN*, *SPREAD* and *INV_Conc* are all positive and significant, with *PIN* explaining the largest proposition of common variation (R^2 =0.613). Precision is strongly correlated with earnings quality *AQ* and *VR*

[7] The reported R^2 calculations are independent and do not add up to 100%.

but weakly associated with *AFP_Total* ($R^2 = 0.011$).

One explanation for the weak association between *AFP_Total* and *AQ* and *VR* is that *AFP_Total* captures the overall quality of firms' information environments, rather than precision alone. Analysts take a holistic approach in appraising companies, and their forecast accuracy depends on information precision, their access to private information – which captures asymmetry – and the amount of information available to them, which captures quantity. Figure 2 validates the use of the nine information quality proxies to capture quantity, precision and asymmetry; however, it also highlights significant variations in how each proxy contributes to the common variation. This result suggests that measurement choices are likely to have a significant impact on the power of tests for the relationship between information quality and CoE, which helps explain the often-conflicting evidence in the previous literature.

Botosan et al. (2011) examine 11 VBM CoE estimates and one RFB proxy for their association with common risk factors and the explanatory power of future realised returns. They find that the price-earnings-growth CoE, *rPEG*, demonstrates the greatest construct validity among all tested proxies. Lee et al. (2010) examine the predictive power of seven VBM and two RFB proxies and conclude that the former outperforms the latter. Based on past evidence, we start by selecting the *rPEG* and the Carhart (1997) model as proxies for CoE. We use the Carhart model as it is one of the most popular empirical asset pricing models. Sensitivity tests examine if our conclusions change depending on the choice of CoE proxies. *rPEG* explains close to 60% of the common variation in CoE. Controlling for rPEG, the four-factor model estimates have a negative association with CoE. The former result is consistent with previous literature showing significant correlations between the implied cost of capital measures and CoE (Botosan et al. 2011). The latter result is consistent with Lee et al. (2010, 2015) and Botosan et al. (2011), who find that VMB and RFB proxies are negatively correlated and conclude that they 'do not capture the same underlying construct' (Botosan et al. 2011, 1102).

(Figure 2 about here)

4.3 The Interrelatedness among Information Quality Dimensions

Figure 2 documents the significant interrelationship between information quality dimensions. We document a significant positive association between quantity and precision (coefficient of 0.277; p-value of 0.000) and conclude a significant negative association between quantity and asymmetry (-0.242; p-value: 0.000). Intuitively, as the amount of information about a firm increases, investors can better evaluate the accuracy of cash flow, and the discount rates information and informational disadvantage of uninformed investors decreases. Further, we document a negative association between precision and asymmetry (-0.114; p-value: 0.000), which is consistent with the prediction that more precise information available to the public reduces the advantage of informed investors. Together, the Figure 2 results suggest that there may be important indirect channels through which the three information quality dimensions affect CoE.

4.4 The Effects Quantity, Precision and Qsymmetry have on CoE

Figure 2 reports only the direct effects of quantity, precision and asymmetry on CoE because indirect links are difficult to graph clearly. Table 3 presents the total effects information quality dimensions have on CoE and their disaggregation into direct and indirect effects. The direct effect of quantity on CoE diverges from its assumed negative association (0.145; p-value: 0.000). However, the total effect (direct + indirect) remains significantly negative (-0.105; p-value: 0.000), given that the indirect effects (i.e. quantity on CoE through precision and asymmetry) are almost two and a half times as important and negative (-0.250; p-value: 0.000). This result suggests that quantity's effect on CoE – identified in previous studies – largely stems from its indirect association with CoE mediated by precision and asymmetry.

With respect to the direct association between precision and CoE, as firms increase the accuracy of their information, they can expect to enjoy lower CoE (-0.520; p-value: 0.000). The direct effect of precision on CoE accounts for almost 90% of the total effect, with the indirect effect explaining 10% of the total association (precision via asymmetry on CoE: -0.050 = -0.114 x 0.441).

Firms that reduce information asymmetry between investors who are privately informed and those who are publicly uninformed (e.g. by decreasing investor concentration) benefit from lower CoE (0.441; p-value: 0.000). As indicated by Figure 2, there is no indirect effect of information asymmetry on CoE, and the direct effect equals the total effect. Bhattacharya et al. (2012) recommend channelling corporate activity towards improving precision rather than decreasing asymmetry. However, given that both effects exhibit similar strength, the question is rather over which attribute a firm has greater discretion. In other words, a firm should dedicate scarce resources to those activities for which the potential of improvement is expected to be greatest.[8]

(Table 3 about here)

Our sample period ends in 2010 as we are constrained by the availability of PIN data. To ensure the robustness of our conclusions, Appendix A repeats the analysis depicted in Table 3 for the sample period 1989–2015 but without PIN as one of the information asymmetry measures.

[8] In the untabulated results, we find that fit statistics for our model are consistent with suggestions in Hu and Bentler (1999) with good levels of fit for the structural model (SRMR: 0.048) and acceptable levels for the measurement model (CFI: 0.912; TLI: 0.873; RMSEA: 0.013).

The conclusions from this test are virtually the same as for our main findings, with the magnitudes of coefficient estimates similar to Table 3 and, on average, slightly higher. To illustrate, the direct effect of asymmetry on CoE increases slightly from 0.441 in Table 3 to 0.565 in Appendix A table. These results suggest that extending the sample period strengthens the evidence for a significant association between information asymmetry and CoE.

5. Sensitivity Analysis

5.1 The Effect of Sample Composition

Previous studies vary significantly in sample compositions, with their focus ranging from IPO stocks (Ecker 2014) to firms listed for a minimum of seven years (Francis et al. 2005). To better understand the impact sample composition has on the association between information quality and CoE, we re-estimate the SEM regression for firms with characteristics that vary across the listing exchange, firm size, profitability, age and for the subperiods 1993/99 and 2000/10.

As a first test, we group firms by the listing exchange, which allows us to compare SEM estimates for firms stratified simultaneously by size, age and profitability. The average NYSE firm is much larger (market cap: USD6,569 m), older (31 years since founding), listed for longer (17.89 years) and more profitable (ROA: 7.2%) than firms trading on AMEX (Market value=USD310 m, Age=28, Listing=12.31, ROA=-1.47%) and NASDAQ (Market value=USD1,494 m, Age=24, Listing=11.75, ROA=-2.5%).[9] The columns 'Exchange' in Table 4 report SEM results by listing exchange. The general conclusion is that coefficient estimates for quantity, precision and asymmetry are smaller for NYSE than NASDAQ and AMEX. To

[9] Unreported ANOVA tests show that mean levels for *Age*, *Listing*, *Market Cap* and *ROA* are significantly different between exchanges.

illustrate, the total effect of asymmetry on CoE for NASDAQ (0.452, p-value: 0.000) and AMEX (0.490, p-value: 0.000) stocks is twice the size of NYSE stocks (0.205, p-value: 0.000). Further, SEM explains just 18.7% of CoE variance for NYSE firms but a substantial 60.4% for AMEX and 34.5% for NASDAQ. This result suggests that samples geared towards NYSE stocks will have less power to identify significant associations between information quality and CoE as coefficients converge towards zero.

(Table 4 about here)

Next, we split the sample into terciles based on firms' market capitalisation. The columns 'Size' in Table 4 indicate that the indirect and direct effects of quantity are higher for smaller stocks, and a similar picture emerges for the direct effects of precision – asymmetry associates with CoE only among the largest firms. Larger firms tend to have higher institutional ownership, which can explain why uninformed investors may require a premium to invest in them (Akins et al. 2012).

The column 'Profitability' in Table 4 highlights that the effects of the three information quality dimensions on CoE are higher for less profitable firms, which is consistent with investors requiring a premium to invest in loss-making or less profitable firms, where future profitability is at risk. We find no evidence that the three quality dimensions are associated with CoE for the most profitable firms; thus, tests on samples geared towards such firms will tend to produce insignificant results.

The column 'Firm Age' in Table 4 demonstrates that the effect of asymmetry on CoE is higher for older firms with more years on a given exchange. This result is consistent with higher institutional ownership levels for these firms (Dahlquist and Robertsson 2001), and institutional investors tend to be more informed. The total effect of precision is stronger for younger firms, but direct and indirect quantity effects on CoE dominate among more established firms. These results are consistent with the shorter availability of financial information, impairing investors' ability to forecast their future performance (Ecker 2014).

Finally, we split the sample into the subperiods 1993 to 1999 and 2000 to 2010 and estimate SEM for each subperiod. We run this test for two reasons. First, the bulk of studies on the association between information quality and CoE use samples comparable to the first half of our sample period. Thus, we want to ensure that our conclusions are not different if we use comparable periods. Second, regulatory reforms in the early 2000s, such as the Regulation Fair Disclosure of 2000 and the Sarbanes–Oxley Act of 2002, reduced information asymmetry between investors and increased information precision and quantity. Thus, the effect of information quality on CoE may be harder to identify in the latter sample period, where information quality should be higher. Columns 'Subperiods' in Table 4 document significant coefficients on the three information quality dimensions for both subperiods, with the effect of asymmetry and precision on CoE increasing in importance over time. The former result is consistent with higher ownership and the more dominant role played by institutional investors in equity markets (Hartzell and Starks 2003), which can correlate with higher asymmetry. The latter result is consistent with the reducing value relevance of accounting numbers (Francis and Schipper 1999, Lev and Zarowin 1999, Core et al. 2003), which leads to the lower precision of accounting information. Thus, our conclusions are valid even for the more recent sample period.

Overall, sensitivity tests substantiate our conclusion regarding the significant association between quantity, precision and asymmetry and CoE. They also highlight that the strength of the association varies with firm characteristics and tends to be stronger for smaller, younger and less profitable firms.

5.2 Variation in the Choice of CoE Proxies

Our main tests use a combination of rPEG and rFF4, and this section tests the sensitivity of our conclusions to the different choice of CoE proxies. Table 5 reports the SEM results when we proxy CoE only by either RFB, FVIX or VMB. We focus on changes in the direct effects, as indirect effects are unchanged because the interrelationship between the three information quality measures do not change. Panel A examines the contribution of each risk factor model to the common variation. The column 'Risk-Factor-Based (RFB)' shows significant loadings on rCAPM, rFF3 and rFF4, which suggest that all models contribute to the common variation. The results for FVIX proxies in column 'Risk Factor-Based + VIX (FVIX)' are similar.

Panel B models the relationship between information quality and CoE. For the column 'Risk-Factor-Based (RFB)', precision shows a positive association with CoE, and the coefficient on quantity is statistically zero. The results in the column 'Risk Factor-Based + VIX (FVIX)' show an insignificant relationship between precision and CoE, and asymmetry has a negative effect on CoE. Thus, using only risk factor–based CoE proxies produces evidence inconsistent with hypothetical predictions and our earlier conclusions.

The column 'Valuation-Model-Based (VMB)' repeats the SEM analysis when we use *rPE*, *rPEG* and *rAEGM* to capture CoE. Panel A reports significant loadings on *rPE*, *rPEG* and *rAEGM*, which suggests that all measures add to the common variation. Panel B shows that the associations between quantity, asymmetry and precision and CoE are similar to our main results.

The columns 'RFB and VMB' and 'FVIX and VMB' repeat the analysis when we combine the three risk factor proxies with the three VMB measures to capture CoE. Panel A factor analysis suggests that CoE has a significant positive effect on VMB but not risk factor measures. Panel B estimates for direct effects are similar to our main results. Overall, the Table 5 results suggest using only risk factor proxies to capture CoE will produce either statistically insignificant associations or those of the opposite sign. These results are consistent with the conclusion in Fama and French (1997) that risk factors produce imprecise CoE estimates. It also supports the arguments in Botosan and Plumlee (2005), Botosan et al. (2011) and Lee et al. (2010, 2015) that VMB outperforms RFB measures as a proxy for CoE. Using VMB models or a mix or VMB and risk factor models produce results consistent with theoretical associations.

(Table 5 about here)

6. Conclusions and Implications

We apply the SEM approach to reconcile and consolidate the extensive literature examining the association between information quality and a firm's costs of equity. Quantity, precision and asymmetry affect CoE both through direct and indirect channels, and we identify the direction of indirect channels. Specifically, we show that the direct effects of precision and asymmetry are equally important in explaining variation in CoE, while quantity has a negative direct effect on CoE. The positive association between quantity and CoE identified in previous studies is due to its indirect effect, which is mediated through asymmetry and precision. Further, we report that the strength of the effects varies according to firm size, maturity and profitability and depending on the choice of risk factor or valuation–based CoE proxies. Our results will help guide future research on the association between information quality and CoE.

Our research has implications for accounting literature and for practitioners. First, the evidence provides guidance on the strength and validity of most common empirical proxies for quantity, precision and asymmetry, and calibrates the effect the three information quality dimensions have on CoE. These insights will help steer future research design choices in this area among academics and practitioners.

Second, the study highlights that researchers should exercise caution when interpreting the direct associations between measures of information quantity and the variable of interest, e.g. CoE, audit quality or accuracy of analyst forecasts. This concern reflects the potential indirect effects of one dimension mediated through another that a researcher needs to model to ensure reliable conclusions.

Third, the study highlights the potential usefulness of structural estimation in other areas to address the concern of subjective selection of proxies for latent dimensions and the effect of errors-in-variables on empirical results. SEM acknowledges that a researcher may be interested in a dimension that is not directly observable. Thus, they will need to select from a range of proxy variables without knowing the true association between the proxy measures and the latent dimension. We recommend that future researchers consider applying SEM to other areas where the dimension of interest is unobservable and captured by several proxies, such as when relating information quality to audit quality or the accuracy of analyst research outputs.

References

Aboody, D., Hughes, J., Liu, J. (2005). Earnings Quality, Insider Trading, and Cost of Capital. *Journal of Accounting Research, 43*(5), 651-673.

Affleck-Graves, J., Callahan, C. M., Chipalkatti, N. (2002). Earnings Predictability, Information Asymmetry, and Market Liquidity. *Journal of Accounting Research, 40*(3), 561-583.

Akins, B. K., Jeffrey, N., Verdi, R. S. (2012). Investor Competition over Information and the Pricing of Information Asymmetry. *The Accounting Review, 87*(1), 35-58.

Amihud, Y., Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17*(2), 223-249.

Ang, A., Hodrick, R. J., Xing, Y., Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *Journal of Finance, 61*(1), 259-299.

Armstrong, C. S., Core, J. E., Taylor, D. J., Verrecchia, R. E. (2011). When Does Information Asymmetry Affect the Cost of Capital? *Journal of Accounting Research, 49*(1), 1-40.

Baginski, S. P., Rakow, K. C. (2012). Management earnings forecast disclosure policy and the cost of equity capital. *Review of Accounting Studies, 17*(2), 279-321.

Bandalos, D. L., & Boehm-Kaufman, M. R. (2009). Common misconceptions in exploratory factor analysis. *Statistical and methodological myths and urban legends: Where pray tell did they get this idea*, 63-88.

Barinov, A. (2013). Analyst disagreement and aggregate volatility risk. *Journal of Financial and Quantitative Analysis, 48*(06), 1877-1900.

Barron, O. E., Harris, D. G., Mary, S. (2005). Evidence That Investors Trade on Private Event-Period Information around Earnings Announcements. *The Accounting Review, 80*(2), 403-421.

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

Barron, O. E., Sheng, X. S., Thevenot, M. (2012). *The Information Environment and Cost of Capital*. Available at SSRN: https://ssrn.com/abstract=2099825.

Barron, O. E., Stuerke, P. S. (1998). Dispersion in Analysts' Earnings Forecasts as a Measure of Uncertainty. *Journal of Accounting, Auditing & Finance, 13*(3), 245-270.

Barry, C. B., Brown, S. J. (1984). Differential information and the small firm effect. *Journal of Financial Economics, 13*(2), 283-294.

Barry, C. B., Brown, S. J. (1985). Differential Information and Security Market Equilibrium. *Journal of Financial and Quantitative Analysis, 20*(4), 407-422.

Barth, M. E., Konchitchki, Y., Landsman, W. R. (2013). Cost of capital and earnings transparency. *Journal of Accounting and Economics, 55*(2–3), 206-224.

Beattie, V., McInnes, B., Fearnley, S. (2004). A methodology for analysing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes. *Accounting Forum, 28*(3), 205-236.

Beretta, S., Bozzolan, S. (2008). Quality versus Quantity: The Case of Forward-Looking Disclosure. *Journal of Accounting, Auditing & Finance, 23*(3), 333-376.

Beyer, A., Cohen, D. A., Lys, T. Z., Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics, 50*(2–3), 296-343.

Bhattacharya, N., Desai, H., Venkataraman, K. (2013). Does Earnings Quality Affect Information Asymmetry? Evidence from Trading Costs*. *Contemporary Accounting Research, 30*(2), 482-516.

Bhattacharya, N., Ecker, F., Olsson, P. M., Schipper, K. (2012). Direct and Mediated Associations among Earnings Quality, Information Asymmetry, and the Cost of Equity. *The Accounting Review, 87*(2), 449-482.

Biddle, G. C., Seow, G. S., Siegel, A. F. (1995). Relative versus Incremental Information Content. *Contemporary Accounting Research, 12*(1), 1-23.

Bloomfield, R., Fischer, P. E. (2011). Disagreement and the Cost of Capital. *Journal of Accounting Research, 49*(1), 41-68.

Bollen, K. A. (2014). *Structural equations with latent variables* (Vol. 210). John Wiley & Sons.

Botosan, C. A. (1997). Disclosure Level and the Cost of Equity Capital. *The Accounting Review, 72*(3), 323-349.

Botosan, C. A., Plumlee, M. A. (2005). Assessing Alternative Proxies for the Expected Risk Premium. *The Accounting Review, 80*(1), 21-53.

Botosan, C. A., Plumlee, M. A. (2013). Are Information Attributes Priced? *Journal of Business Finance & Accounting, 40*(9-10), 1045-1067.

Botosan, C. A., Plumlee, M. A., Wen, H. E. (2011). The Relation between Expected Returns, Realized Returns, and Firm Risk Characteristics. *Contemporary Accounting Research, 28*(4), 1085-1122.

Brennan, M. J., Huh, S. W., Subrahmanyam, A. (2016). Asymmetric Effects of Informed Trading on the Cost of Equity Capital. *Management Science, 62*(9), 2460-2480.

Brown, S., Hillegeist, S. (2007). How disclosure quality affects the level of information asymmetry. *Review of Accounting Studies, 12*(2-3), 443-477.

Brown, S., Hillegeist, S. A., Lo, K. (2004). Conference calls and information asymmetry. *Journal of Accounting and Economics, 37*(3), 343-366.

Busemeyer, J. R., & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin, 93*(3), 549.

Byard, D., Shaw, K. W. (2003). Corporate Disclosure Quality and Properties of Analysts' Information Environment. *Journal of Accounting, Auditing & Finance, 18*(3), 355-378.

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H.-m., Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies, 19*(1), 396-455.

Cao, Y., Myers, L. A., Tsang, A., Yang, Y. G. (2017). Management forecasts and the cost of equity capital: international evidence. *Review of Accounting Studies*, 1-48.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance, 52*(1), 57-82.

Caskey, J., Hughes, J. S., Liu, J. (2015). Strategic informed trades, diversification, and expected returns. *The Accounting Review, 90*(5), 1811-1837.

Clarkson, P., Guedes, J., Thompson, R. (1996). On the diversification, observability, and measurement of estimation risk. *Journal of Financial and Quantitative Analysis, 31*(01), 69-84.

Clarkson, P. M., Satterly, A. (1997). Australian evidence on the pricing of estimation risk. *Pacific-Basin Finance Journal, 5*(3), 281-299.

Clarkson, P. M., Thompson, R. E. X. (1990). Empirical Estimates of Beta When Investors Face Estimation Risk. *Journal of Finance, 45*(2), 431-453.

Copeland, T. E., Galai, D. (1983). Information Effects on the Bid-Ask Spread. *Journal of Finance, 38*(5), 1457-1469.

Core, J., Guay, W., Van Buskirk, A. (2003). Market valuations in the new economy: An investigation of what has changed. *Journal of Accounting and Economics, 34*(1-3), 43–67.

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

Corwin, S. A., Schultz, P. (2012). A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *Journal of Finance, 67*(2), 719-760.

Cox, C. T. (1985). Further evidence on the representativeness of management earnings forecasts. *The Accounting Review*, 692-701.

Dahlquist, M., & Robertsson, G. (2001). Direct foreign ownership, institutional investors, and firm characteristics. *Journal of Financial Economics, 59*(3), 413-440.

Dechow, P. M., Dichev, I. D. (2002). The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors. *The Accounting Review, 77*(s-1), 35-59.

Diamond, D. W., Verrecchia, R. E. (1991). Disclosure, Liquidity, and the Cost of Capital. *Journal of Finance, 46*(4), 1325-1359.

Duarte, J., Han, X., Harford, J., Young, L. (2008). Information asymmetry, information dissemination and the effect of regulation FD on the cost of capital. *Journal of Financial Economics, 87*(1), 24-44.

Easley, D., Hvidkjaer, S., O'Hara, M. (2010). Factoring Information into Returns. *Journal of Financial and Quantitative Analysis, 45*(2), 293-309.

Easley, D., Hvidkjaer, S., O'Hara, M. (2002). Is Information Risk a Determinant of Asset Returns? *Journal of Finance, 57*(5), 2185-2221.

Easley, D., Kiefer, N. M., O'Hara, M. (1997). One day in the life of a very common stock. *Review of Financial Studies, 10*(3), 805-835.

Easley, D., Kiefer, N. M., O'Hara, M., Paperman, J. B. (1996). Liquidity, Information, and Infrequently Traded Stocks. *Journal of Finance, 51*(4), 1405-1436.

Easley, D., O'Hara, M. (2004). Information and the Cost of Capital. *Journal of Finance, 59*(4), 1553-1583.

Easton, P. D. (2004). PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital. *The Accounting Review, 79*(1), 73-95.

Easton, P. D., Monahan, S. J. (2005). An Evaluation of Accounting-Based Measures of Expected Returns. *The Accounting Review, 80*(2), 501-538.

Easton, P. D., Monahan, S. J. (2016). Review of Recent Research on Improving Earnings Forecasts and Evaluating Accounting-based Estimates of the Expected Rate of Return on Equity Capital. *Abacus, 52*(1), 35-58.

Ecker, F. (2014). Information Precision and Long-Run Performance of Initial Public Offerings. *Contemporary Accounting Research, 31*(3), 876-910.

Ecker, F., Francis, J., Kim, I., Olsson, P. M., Schipper, K. (2006). A Returns-Based Representation of Earnings Quality. *The Accounting Review, 81*(4), 749-780.

Ertugrul, M., Lei, J., Qiu, J., & Wan, C. (2017). Annual report readability, tone ambiguity, and the cost of borrowing. *Journal of Financial and Quantitative Analysis, 52*(2), 811-836.

Falkenstein, E. G. (1996). Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *The Journal of Finance, 51*(1), 111-135.

Fama, E. F., French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.

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

Ferson, W. E., Locke, D. H. (1998). Estimating the Cost of Capital Through Time: An Analysis of the Sources of Error. *Management Science*, *44*(4), 485-500.

Francis, J., LaFond, R., Olsson, P. M., Schipper, K. (2004). Costs of Equity and Earnings Attributes. *The Accounting Review, 79*(4), 967-1010.

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

Francis, J., Nanda, D., Olsson, P. E. R. (2008). Voluntary Disclosure, Earnings Quality, and Cost of Capital. *Journal of Accounting Research, 46*(1), 53-99.

Francis, J., Schipper, K. (1999). Have financial statements lost their relevance?. Journal of accounting Research, 37(2), 319-352

Fu, R., Kraft, A., Zhang, H. (2012). Financial reporting frequency, information asymmetry, and the cost of equity. *Journal of Accounting and Economics*, *54*(2–3), 132-149.

Giglio, S., & Xiu, D. (2017). *Inference on risk premia in the presence of omitted factors* (No. w23527). National Bureau of Economic Research.

Giovannelli, A., Massacci, D., & Soccorsi, S. (2018). Forecasting stock returns with large dimensional factor models. *Available at SSRN 2958491*.

Glosten, L. R., Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics, 14*(1), 71-100.

Gow, I. D., Larcker, D. F., & Reiss, P. C. (2016). Causal inference in accounting research. *Journal of Accounting Research*, *54*(2), 477-523.

Gray, P., Koh, P.-S., Tong, Y. H. (2009). Accruals Quality, Information Risk and Cost of Capital: Evidence from Australia. *Journal of Business Finance & Accounting, 36*(1-2), 51-72.

Guay, W., Kothari, S., Shu, S. (2011). Properties of implied cost of capital using analysts' forecasts. *Australian Journal of Management, 36*(2), 125-149.

Hartzell, J. C., Starks, L. T. (2003). Institutional investors and executive compensation. *The Journal of Finance*, *58*(6), 2351-2374.

Healy, P. M., Hutton, A. P., Palepu, K. G. (1999). Stock Performance and Intermediation Changes Surrounding Sustained Increases in Disclosure. *Contemporary Accounting Research, 16*(3), 485-520.

Heflin, F. L., Shaw, K. W., Wild, J. J. (2005). Disclosure Policy and Market Liquidity: Impact of Depth Quotes and Order Sizes. *Contemporary Accounting Research, 22*(4), 829-865.

Hinson, L. A., & Utke, S. (2018). Structural Equation Modeling in Archival Capital Markets Research: An Empirical Application to Disclosure and Cost of Capital. *Available at SSRN 2614085*.

Hu, L. t., Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, *28*(3), 791-837.

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

Imhoff, E. A. (1978). The representativeness of management earnings forecasts. *The Accounting Review*, 836-850.

Jayaraman, S. (2008). Earnings Volatility, Cash Flow Volatility, and Informed Trading. *Journal of Accounting Research, 46*(4), 809-851.

Kim, D., Qi, Y. (2010). Accruals quality, stock returns, and macroeconomic conditions. *The Accounting Review, 85*(3), 937-978.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Kothari, S. P., Li, X., Short, J. E. (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.

Kross, W., Schroeder, D. A. (1989). Firm Prominence and the Differential Information Content of Quarterly Earnings Announcements. *Journal of Business Finance & Accounting, 16*(1), 55-74.

Kumar, P., Sorescu, S. M., Boehme, R. D., Danielsen, B. R. (2008). Estimation Risk, Information, and the Conditional CAPM: Theory and Evidence. *Review of Financial Studies, 21*(3), 1037-1075.

Lai, S., Ng, L., & Zhang, B. (2014). Does PIN affect equity prices around the world?. *Journal of Financial Economics*, *114*(1), 178-195.

Lambert, R., Leuz, C., Verrecchia, R. E. (2007). Accounting Information, Disclosure, and the Cost of Capital. *Journal of Accounting Research, 45*(2), 385-420.

Lambert, R. A., Leuz, C., Verrecchia, R. E. (2012). Information Asymmetry, Information Precision, and the Cost of Capital. *Review of Finance*, *16*(1), 1-29.

Lang, M., Lundholm, R. (1993). Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research*, 246-271.

Lang, M. H., Lundholm, R. J. (1996). Corporate Disclosure Policy and Analyst Behavior. *The Accounting Review, 71*(4), 467-492.

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, *56*(1), 130-147.

LeBreton, J. M., Wu, J., & Bing, M. N. (2009). The truth (s) on testing for mediation in the social and organizational sciences. *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences*, 107-141.

Lee, C., So, E. C., Wang, C. (2010). Evaluating Implied Cost of Capital Estimates. *SSRN eLibrary*.

Lee, C., So, E. C., Wang, C. C. (2015). Evaluating firm-level expected-return proxies. *Harvard Business School Accounting & Management Unit Working Paper*(15-022), 15-57.

Lee, P., Stokes, D., Taylor, S., Walter, T. (2003). The association between audit quality, accounting disclosures and firm-specific risk: Evidence from initial public offerings. *Journal of Accounting and Public Policy, 22*(5), 377-400.

Lev, B., Zarowin, P. (1999). The boundaries of financial reporting and how to extend them. *Journal of Accounting research, 37*(2), 353-385.

Levi, S., Zhang, X. J. (2015). Do Temporary Increases in Information Asymmetry Affect the Cost of Equity? *Management Science, 61*(2), 354-371.

Lewellen, J., Shanken, J. (2002). Learning, Asset-Pricing Tests, and Market Efficiency. *Journal of Finance, 57*(3), 1113-1145.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*(2-3), 221-247.

Li, G. (2005). Information Quality, Learning, and Stock Market Returns. *Journal of Financial and Quantitative Analysis, 40*(3), 595-620.
Lipe, R. C. (1986). The information contained in the components of earnings. *Journal of Accounting Research*, 37-64.
Little, T. D., Lindenberger, U., Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods, 4*(2), 192-211.
Liu, J., Nissim, D., Thomas, J. (2002). Equity Valuation Using Multiples. *Journal of Accounting Research, 40*(1), 135-172.
McNichols, M. F. (2002). The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors: Discussion. *The Accounting Review*, *77*, 61-69.
Mohanram, P., Rajgopal, S. (2009). Is PIN priced risk? *Journal of Accounting and Economics, 47*(3), 226-243.
Nezlobin, A., Sloan, R. G., Zha Giedt, J. (2019). Measuring Accruals Quality: A Theoretical and Empirical Evaluation. Available at SSRN: https://ssrn.com/abstract=3301083
Ogneva, M. (2012). Accrual Quality, Realized Returns, and Expected Returns: The Importance of Controlling for Cash Flow Shocks. *The Accounting Review*, *87*(4), 1415-1444.
Ramakrishnan, R. T., & Thomas, J. K. (1998). Valuation of permanent, transitory, and price-irrelevant components of reported earnings. *Journal of Accounting, Auditing & Finance, 13*(3), 301-336.
Rao, P. (1973). Some Notes on the Errors-in-Variables Model. *The American Statistician, 27*(5), 217-218.
Shevlin, T. (2013). Some personal observations on the debate on the link between financial reporting quality and the cost of equity capital. *Australian Journal of Management, 38*(3), 447-473.
Stoll, H. R. (1978). The Pricing of Security Dealer Services: An Empirical Study of Nasdaq Stocks. *Journal of Finance, 33*(4), 1153-1172.
Tasker, S. C. (1998). Bridging the information gap: Quarterly conference calls as a medium for voluntary disclosure. *Review of Accounting Studies, 3*(1-2), 137-167.
Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139-1168.
Van Ness, B. F., Van Ness, R. A., & Warr, R. S. (2001). How well do adverse selection components measure adverse selection?. *Financial Management*, 77-98.
Waymire, G. (1985). Earnings volatility and voluntary management forecast disclosure. *Journal of Accounting Research*, 268-295.
Welker, M. (1995). Disclosure Policy, Information Asymmetry, and Liquidity in Equity Markets. *Contemporary Accounting Research, 11*(2), 801-827.
Wild, J. J. (1992). Stock price informativeness of accounting numbers: evidence on earnings, book values, and their components. *Journal of Accounting and Public Policy*, *11*(2), 119-154.
Wysocki, P. (2009). Assessing earnings and accruals quality: US and international evidence. *Unpublished working paper. Cambridge: MIT Sloan School of Management*.
Zhang, X. F. (2006). Information Uncertainty and Stock Returns. *Journal of Finance, 61*(1), 105-137.

Appendix A. Information quality and CoE: sample period till 2015 excluding the PIN measure

Panel A: Effects from Quantity to CoE	Estimate	T-statistic	p-value (two-tailed)
Indirect + Direct effect=Total effect	-0.157	-6.710	0.000
Total Indirect effect	-0.281	-7.282	0.000
Indirect effect through			
Quant-Prec-CoE	-0.138	-4.269	0.000
Quant-Asym-CoE	-0.104	-6.941	0.000
Quant-Prec-Asym-Coe	-0.039	-4.385	0.000
Direct effect	0.124	5.470	0.000
Panel B: Effects from Precision to CoE			
Indirect + Direct effect=Total effect	-0.581	-6.557	0.000
Total Indirect Effect	-0.127	-5.279	0.000
Indirect effect through			
Prec-Asym-CoE	-0.127	-5.279	0.000
Direct effect	-0.453	-4.423	0.000
Panel C: Effects from Asymmetry to CoE			
Direct Effect = Total Effect	0.565	6.954	0.000
Panel D: Model Fit Statistics			
Obs.	85,262		
χ^2	191.0		
χ^2 p-value	0.000		
CoE R-Square	0.614		

Notes: The table reports total, indirect, and direct CoE effects for the information quality dimensions. The sample period is 1989-2015 and we exclude PIN from the measures of information asymmetry. Panel D shows model fit statistics and explained variance for CoE. ***, **, * denote significance at 1%, 5% and 10% level.

Table 1. Variables

Measure	Abbreviation	Association with CoE	Source
Quantity			
Period of listing (IPO)	Listing	pos.	SDC Platinum, Osiris & CRSP
Firm age (incorporation)	Age	pos.	Osiris
Relative media coverage	RMC	pos.	Factiva
Precision			
Accrual quality	AQ	pos.	Compustat
Earnings value relevance	VR	pos.	Compustat & CRSP
Total analyst forecast precision	AFP_Total	pos.	I/B/E/S
Asymmetry			
Probability of an informed trade	PIN	pos.	Stephen Brown
Bid/Ask spread	SPREAD	pos.	CRSP
Investor concentration	INV_Conc	pos.	CDA/Spectrum (s12)
CoE – Risk Factor-Based (RFB)			
Capital Asset Pricing Model	rCAPM	pos.	
Fama-French 3-factor model	rFF3	pos.	Kenneth French & CRSP
Carhart's 4-factor model	rFF4	pos.	
CoE – Risk Factor-Based + VIX (FVIX)			
Capital Asset Pricing Model + FVIX	rFVIX	pos.	
Fama-French 3-factor model + FVIX	rFVIX3	pos.	Kenneth French, Alexander Barinov & CRSP
Carhart's 4-factor model + FVIX	rFVIX4	pos.	
CoE – Valuation Model-Based (VMB)			
Price-Earnings-Ratio	rPE	pos.	
Price-Earnings-Growth	rPEG	pos.	I/B/E/S & Compustat
Abnormal Earnings Growth Model	rAEGM	pos.	

Notes: The table reports variables used in the study to measure Quantity, Precision, Asymmetry and cost of equity capital.

Table 2. Summary Statistics

	Mean	Std. Dev.	25%	Median	75%	Cronbach's Alpha
Quantity						
Listing	14.668	12.614	7.000	11.000	17.000	
Age	27.862	26.232	11.000	18.000	35.000	0.66
RMC	-10.134	1663.926	-535.663	-245.481	111.810	
Precision						
AQ	-0.061	0.067	-0.073	-0.040	-0.023	
VR	-0.410	0.269	-0.523	-0.348	-0.225	0.43
AFP_Total	1059.000	1440.074	105.310	412.833	1414.583	
Asymmetry						
PIN	0.205	0.107	0.123	0.184	0.268	
SPREAD	0.019	0.027	0.003	0.010	0.024	0.65
INV_Conc	0.195	0.224	0.048	0.103	0.248	
CoE-RFB						
rCAPM	0.084	0.106	0.006	0.063	0.138	
rFF3	0.115	0.127	0.038	0.104	0.181	0.77
rFF4	0.103	0.208	0.026	0.093	0.174	
CoE-FVIX						
rFVIX	0.079	0.127	0.005	0.063	0.140	
rFVIX3	0.115	0.148	0.040	0.110	0.187	0.75
rFVIX4	0.102	0.238	0.022	0.100	0.184	
CoE-VMB						
rPE	0.069	0.060	0.047	0.064	0.084	
rPEG	0.124	0.096	0.084	0.103	0.137	0.33
rAEGM	0.139	0.310	0.093	0.113	0.146	

Notes. The table reports summary statistics for the variables used in the SEM analysis. Variable names as defined in Table 1.

Table 3. Indirect and Total CoE Effects

Panel A: Effects from Quantity to CoE	Estimate	T-statistic	p-value (two-tailed)
Indirect + Direct effect=Total effect	-0.105	-5.347	0.000
Total Indirect effect	-0.250	-6.805	0.000
Indirect effect through			
Quant-Prec-CoE	-0.130	-3.763	0.000
Quant-Asym-CoE	-0.107	-5.447	0.000
Quant-Prec-Asym-Coe	-0.014	-4.442	0.000
Direct effect	0.145	5.568	0.000
Panel B: Effects from Precision to CoE			
Indirect + Direct effect=Total effect	-0.520	-5.184	0.000
Total Indirect Effect	-0.050	-4.154	0.000
Indirect effect through			
Prec-Asym-CoE	-0.050	-4.154	0.000
Direct effect	-0.470	-4.521	0.000
Panel C: Effects from Asymmetry to CoE			
Direct Effect = Total Effect	0.441	6.385	0.000
Panel D: Model Fit Statistics			
Obs.	60,995		
χ^2	413.0		
χ^2 p-value	0.000		
CoE R-Square	0.437		

Notes: The table reports total, indirect, and direct CoE effects for the information quality dimensions. Panel D shows model fit statistics and explained variance for CoE. ***, **, * denote significance at 1%, 5% and 10% level.

Table 4. Variations with estimates based on characteristics

	Exchange			Size			Profitability			Listing			Firm Age			Subperiods	
Panel A: Effects from Quantity to CoE	AMEX	NASDAQ	NYSE	Small	Medium	Large	Low	Medium	High	Short	Medium	Long	Young	Medium	Old	1993/99	2000/10
Indirect + Direct effect=Total effect	-0.091	-0.042**	-0.051**	0.110	0.044	0.032	0.035	-0.059**	-0.550	n/a	0.002	-0.137***	0.021*	0.005	-0.132***	-0.153**	-0.098***
Total Indirect effect	-0.297	-0.226***	-0.147***	-0.195***	-0.125***	-0.143***	-0.166***	-0.200***	-0.711	n/a	-0.049***	-0.255***	0.019	-0.020	-0.206***	-0.204***	-0.231***
Indirect effect through																	
Quant-Prec-CoE	-0.174	-0.118***	-0.073***	-0.192***	-0.119***	-0.116***	-0.091***	-0.124***	-0.397	n/a	-0.032**	-0.087	0.020*	-0.021	-0.030	-0.107**	-0.130***
Quant-Asym-CoE	-0.122	-0.123***	-0.062***	-0.008	-0.010	-0.037***	-0.077***	-0.064***	-0.273	n/a	-0.015**	-0.124***	-0.006	0.006	-0.132**	-0.076***	-0.072***
Quant-Prec-Asym-Coe	-0.001	0.014**	-0.012***	0.005	0.004	0.011***	0.002	-0.012***	-0.041	n/a	-0.003**	-0.044**	0.004*	-0.006*	-0.044**	-0.022***	-0.029***
Direct effect	0.206	0.185***	0.096***	0.304***	0.169***	0.175***	0.201***	0.140***	0.162	n/a	0.051**	0.117***	0.002	0.025	0.074**	0.051*	0.133***
Panel B: Effects from Precision to CoE																	
Indirect + Direct effect=Total effect	-0.633***	-0.400***	-0.391***	-0.766***	-0.584***	-0.351***	-0.448***	-0.611***	-1.357	-0.471***	-0.392***	-0.392*	-0.224*	-0.241*	-0.205**	-0.524***	-0.618***
Total Indirect Effect	-0.004	0.055***	-0.057***	0.022	0.021	0.035	0.011	-0.054***	-0.128	0.003	-0.030***	-0.132**	-0.038**	-0.052*	-0.122**	-0.088***	-0.113***
Indirect effect through																	
Prec-Asym-CoE	-0.004	0.055***	-0.057***	0.022	0.021	0.035	0.011	-0.054***	-0.128	0.003	-0.030***	-0.132**	-0.038**	-0.052*	-0.122**	-0.088***	-0.113***
Direct effect	-0.629***	-0.455***	-0.334***	-0.788***	-0.605***	-0.386***	-0.459***	-0.557***	-1.229	-0.475***	-0.361***	-0.260	-0.186*	-0.189*	-0.083	-0.436***	-0.505***
Panel C: Effects from Asymmetry to CoE																	
Direct Effect = Total Effect	0.490**	0.450***	0.206***	0.053	0.047	0.126***	0.333***	0.301***	1.337	0.363***	0.263***	0.598***	0.148*	0.168*	0.486***	0.365***	0.660***
Panel D: Model Fit Statistics																	
Obs.	4,018	27,612	29,365	20,308	20,308	20,308	18,187	18,187	18,187	11,793	9,203	8,866	7,895	6,959	7,414	21,475	39,520
χ^2	104.3	284.6	517.4	1,859.1	1,463.3	1,199.1	471.1	299.4	410.7	1,358.6	764.5	2,462.8	529.1	1,165.3	1,643.0	271.3	459.0
χ^2 p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CoE R-Square	0.604	0.339	0.182	0.566	0.332	0.127	0.299	0.441	UD	0.354	0.220	0.470	0.071	0.083	0.247	0.396	0.761
Panel E: Differences in Structural Coefficients																	
Direct Effects	AMEX-NYSE	NASDAQ-NYSE	AMEX-NASDAQ	Small-Large	Medium-Large	Small-Medium	Low-High	Medium-High	Low-Medium	Short-Long	Medium-Long	Short-Medium	Young-Old	Medium-Old	Young-Medium	1993/99-2000/10	
Quant. → CoE	0.110	0.089***	0.021	0.129**	-0.006	0.135*	0.039***	-0.022***	0.061***	n/a	-0.066***	n/a	-0.072***	-0.049	-0.023*	-0.082	
Prec. → CoE	-0.295*	-0.120	0.295***	-0.402***	-0.219**	-0.183*	0.770**	0.672***	0.089	-0.197*	-0.101***	-0.114	-0.103	-0.106	0.003	0.069***	
Asym. → CoE	0.284**	0.244*	0.040*	-0.073	-0.079	0.006*	-1.004***	-1.036**	0.032**	-0.235***	-0.335**	0.100	-0.338**	-0.318**	-0.020	-0.295***	

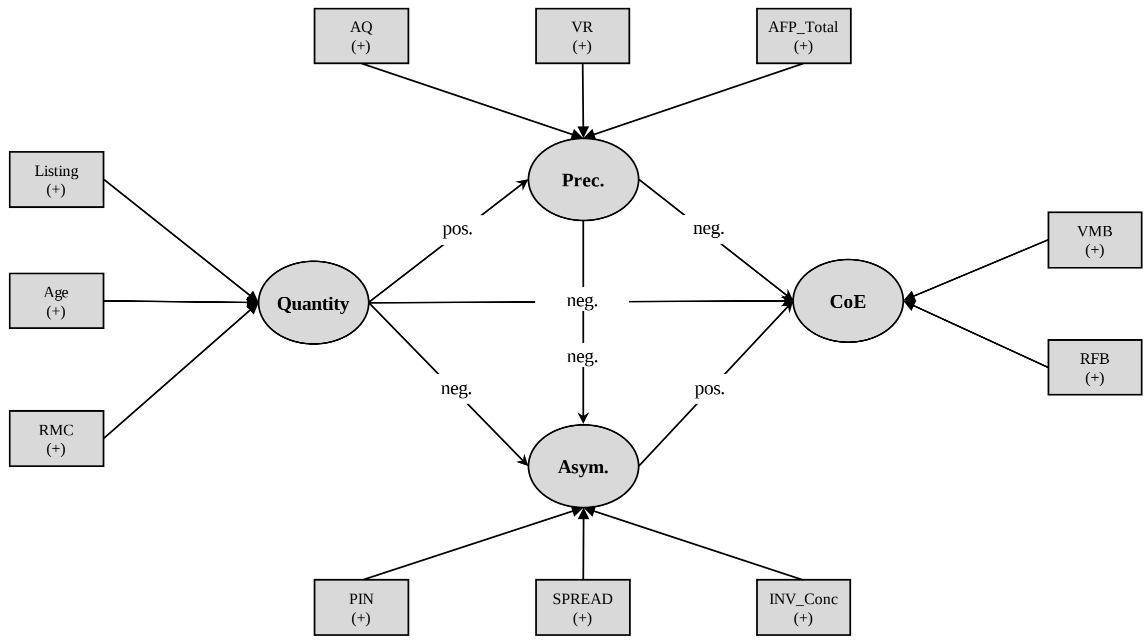
Notes: The table reports total, indirect, and direct CoE effects for the information quality dimensions for samples split by the listing exchange, size (market value), profitability (ROA), listing (years after IPO), firm age (year of incorporation) and subperiods; n/a indicates standardisation of coefficients not possible due to negative residual variance (Heywood case). Panel D shows model fit statistics and explained variance for CoE; UD denotes undefined R-square due to a standardised coefficient of above 1. Panel E tests for statistically significant differences in structure coefficients using the Satorra-Bentler Scaled Chi-Square. ***, **, * denote significance at 1%, 5% and 10% level..

Table 5. CoE Measurement Variations

	Risk-Factor-Based (RFB)	Risk-Factor-Based + VIX (FVIX)	Valuation-Model-Based (VMB)	RFB and VMB	FVIX and VMB
Panel A: CoE Factor Loadings					
rCAPM	0.455***			-0.087	
rFF3	1.313***			-0.104	
rFF4	0.567***			-0.087	
rFVIX		0.416***			-0.128
rFVIX3		1.253***			-0.154
rFVIX4		0.593***			-0.125
rPE			0.694***	0.728***	0.744***
rPEG			1.044***	0.996***	0.968***
rAEGM			0.320***	0.349***	0.364***
Panel B: Direct Effects					
Quant. → CoE	-0.036	-0.035*	0.102***	0.112***	0.116***
Asym. → CoE	0.016**	-0.035***	0.323***	0.358***	0.366***
Prec. → CoE	0.046***	0.010	-0.353***	-0.340***	-0.340***
Panel C: Model Fit Statistics					
Obs.	60,995	60,995	60,660	60,995	60,995
χ^2	991.5	1038.9	197.3	842.7	615.0
χ^2 p-value	0.000	0.000	0.000	0.000	0.000
CoE R-Square	0.003	0.002	0.244	0.259	0.264

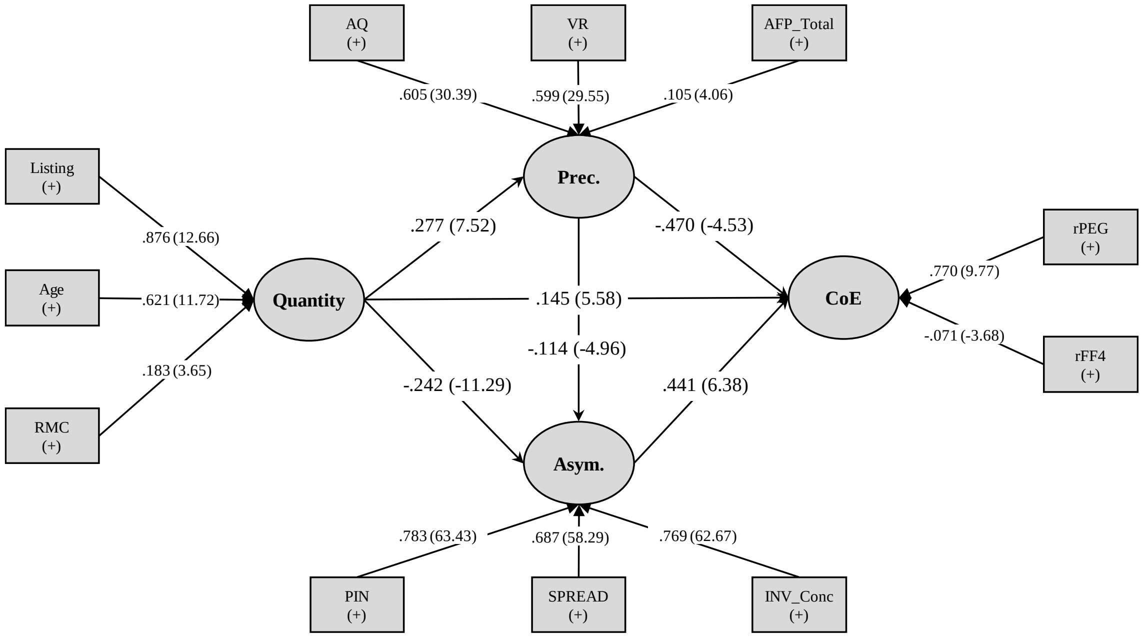
Notes: The table reports factor loadings (Panel A) and structural coefficients (Panel B) for several CoE measurement variations. Panel C shows model fit statistics and explained variance for CoE. ***, **, * denote significance at 1%, 5% and 10% level.

Figure 1. Conceptual Model



Notes: Oval figures indicate latent constructs (i.e. information attributes) which are only indirectly measurable through their impact on observable indicator variables/proxies (rectangular figures).

Figure 2. SEM estimation results



Notes. t-statistics based on standard errors clustered by firm are reported in parentheses. Variable names as defined in Table 1.