Rev Acc Stud (2007) 12:443–477
DOI 10.1007/s11142-007-9032-5

How disclosure quality affects the level of information asymmetry

Stephen Brown · Stephen A. Hillegeist

Published online: 4 April 2007

Abstract We examine two potential mechanisms through which disclosure quality is expected to reduce information asymmetry: (1) altering the trading incentives of informed and uninformed investors so that there is relatively less trading by privately informed investors, and (2) reducing the likelihood that investors discover and trade on private information. Our results indicate that the negative relation between disclosure quality and information asymmetry is primarily caused by the latter mechanism. While information asymmetry is negatively associated with the quality of the annual report and investor relations activities, it is positively associated with quarterly report disclosure quality. Additionally, we hypothesize and find that that the negative association between disclosure quality and information asymmetry is stronger in settings characterized by higher levels of firm-investor asymmetry.

Keywords Disclosure quality · Information asymmetry · Informed trading · Private information events

JEL Classifications M41 · D82 · G14

We examine how the quality of a firm's disclosures is related to the average level of information asymmetry among equity investors over a year. Information asymmetry occurs when one or more investors possess private information about the firm's

S. Brown
Department of Accounting, Goizueta Business School, Emory University,
Atlanta, GA, USA

S. A. Hillegeist (✉)
INSEAD, Accounting and Control Area, Boulevard de Constance,
77305 Fontainebleau Cedex, France
e-mail: Stephen.Hillegeist@INSEAD.edu

value while other uninformed investors only have access to public information. The presence of information asymmetry creates an adverse selection problem in the market when privately informed investors trade on the basis of their private information. Healy, Hutton, and Palepu (1999) Heflin, Shaw, and Wild (2005), and Welker (1995) find that there is a negative association between disclosure quality and spreads-based measures of information asymmetry. In this paper, we explore the precise mechanisms through which disclosure quality affects information asymmetry. Our findings provide some empirical support for regulators' beliefs that high quality disclosures make the capital markets more attractive to ''ordinary'' uninformed investors (FASB, 2001; FASC, 1998; Levitt, 1998).

We find that the association between disclosure quality and our proxy for information asymmetry is negative. Our empirical proxy, the probability of informed trade (PIN), is based on the imbalance between buy and sell orders among investors. Thus, we validate and strengthen prior analyses that utilize indirect, spreads-based proxies of information asymmetry. Using this measure is important because spread-based measures suffer from numerous econometric and interpretation difficulties (Callahan, Lee, & Yohn, 1997; Heflin et al., 2005; Lee, Mucklow, & Ready, 1994; O'Hara, 1995). For example, market makers protect themselves from information asymmetry by simultaneously manipulating both the quoted bid and ask prices along with the quoted depths associated with those prices. Therefore, analyses relying solely on spread-based measures are incomplete and difficult to interpret (Lee et al., 1994).[1]

We conduct three additional analyses to gain a deeper understanding of how disclosure quality affects information asymmetry. The first analysis decomposes our measure of information asymmetry and allows us to examine the underlying channels through which disclosure quality is related to information asymmetry. Our results indicate that disclosure quality primarily affects information asymmetry by reducing the likelihood that investors discover and trade on private information. If the negative association we document reflects a reduction in non-productive search activities, then higher quality disclosures could improve aggregate shareholder welfare by reducing search costs. We also find that disclosure quality is positively associated with the amount of trading by both uninformed *and* privately informed investors. However, these increases appear to offset one another such that we find no evidence of an association between disclosure quality and the *relative* amount of informed trading, which is consistent with Kyle (1985). Our results provide new evidence on how disclosure quality is negatively related to information asymmetry and are of interest to regulators and firms who wish to use disclosure policy to reduce the level of information asymmetry.

Our second extension examines whether disclosure quality in three different areas each has the same relation with information asymmetry as does an aggregate measure of disclosure quality. Specifically, we investigate disclosure quality related to (1) the annual report, (2) the quarterly reports, and (3) investor relations activities.

[1] In addition, prior studies commonly rely on the closing bid-ask spread. However, Madhavan, Richardson, and Roomans (1997) show that adverse selection costs decrease throughout the day, which suggests that the closing spread is a relatively weak proxy for information asymmetry.

Based on a regression that includes all three measures of disclosure quality, we find that the quality of the annual report and investor relations activities are negatively associated with the level of information asymmetry. Surprisingly, we find a positive association between information asymmetry and quarterly report disclosure quality. While inconsistent with our hypothesis, this latter finding is reminiscent of the unexpectedly positive association between quarterly report disclosure quality and the cost of capital documented in Botosan and Plumlee (2002).

Our final investigation examines whether the negative relation between disclosure quality and information asymmetry is stronger in settings characterized by higher levels of firm-investor asymmetry. We hypothesize and find that that in such cases, public disclosures are especially effective in reducing information asymmetry among investors. In particular, we find that the relation is significantly stronger in industry-years where market-to-book ratios are high. These findings indicate that the effects of disclosure quality on asymmetry are likely to vary systematically across firms.

We estimate our proxy for information asymmetry, the PIN, using an extended version of the popular EKO market microstructure model (Easley, Kiefer, & O'Hara, 1997).[2] The PIN is a firm-specific estimate of the probability that a trade originates from a privately informed investor; hence, it directly captures the extent of information asymmetry among investors in the secondary market. An important advantage of the EKO methodology over spreads-based proxies of information asymmetry is that we can disaggregate the PIN measure into its component parameters, each of which represents a different aspect of the firm's trading and information environments. Thus, it allows us to extend the analysis beyond simply *whether* disclosure quality and information asymmetry are related by examining *how* they are related.

We use analysts' evaluations of firms' disclosure activities compiled by the Association for Investment Management and Research (AIMR) as our proxy for disclosure quality. While the AIMR scores are imperfect measures of disclosure quality, they offer several advantages over alternative proxies. AIMR scores are based on a comprehensive evaluation of a firm's disclosure activities over an extended time period. Thus, our study generalizes and complements other studies that focus on just one type of disclosure (e.g. Brown et al., 2004; Coller & Yohn, 1997; Marquardt & Wiedman, 1998). In addition, the AIMR scores allow us to examine the effects of disclosure quality on a relatively large cross-section of firms, although one that is skewed toward larger firms with high analyst following. While not fully representative, using this sample makes it more difficult for us to reject the null hypotheses since there is likely less variation in disclosure quality and information asymmetry in our sample compared to the entire population of firms.

Understanding how disclosure quality affects information asymmetry is important because it provides insights into several fundamental issues that are of interest to managers, investors, academics and regulators. A growing body of literature reports a negative relation between various measures of disclosure quality and cost

[2] The PIN methodology is used frequently in the finance literature and has been used by Botosan and Plumlee (2004), Brown, Hillegeist, and Lo (2004) and Hillary (2006) in the accounting literature.

of capital estimates (Botosan, 1997; Botosan & Plumlee, 2002; Francis, LaFond, Olsson, & Schipper, 2005; Sengupta, 1998). Extant literature also documents a positive association between the level of information asymmetry and the cost of capital (Easley, Hvidkjaer, & O'Hara, 2002, 2004). Together, these findings suggest that disclosure quality is related to the cost of capital via its effect on information asymmetry. This link suggests that understanding how disclosure quality affects information asymmetry is an important step towards gaining a deeper understanding of why disclosure quality is related to the cost of capital.

We next discuss how we expect disclosure quality to affect the level of information asymmetry. We develop our empirical proxies in Sect. 2 and discuss our research methodology in Sect. 3. Section 4 discusses data sources and variable construction and provides descriptive statistics while Sect. 5 presents the results of our empirical analyses. We discuss the applicability of our results (based on 1985–1996 data) to more recent time periods in Sect. 6. Section 7 summarizes and concludes the paper.

1 The relation between disclosure quality and information asymmetry

Information asymmetry occurs when one or more investors possess private information about the firm's value. Asymmetry creates an adverse selection problem in the market as informed investors trade on the basis of their private information.[3] These trading activities manifest themselves as unusually large imbalances in the observed order flow; therefore the extent of information asymmetry between investors can be characterized as the probability that a particular buy or sell order comes from an investor with private information. In this section, we discuss how a firm's choice of disclosure quality potentially influences the level of information asymmetry.

One of the ways in which disclosure quality affects information asymmetry is by altering the trading behavior of uninformed investors. According to the Investor Recognition Hypothesis (Merton, 1987), such investors are more likely to invest and trade in firms that are well known or that they judge favorably. If higher disclosure quality increases a firm's visibility and/or reduces the costs of processing firm-specific public information, then higher disclosure quality will induce more trading in the firm's stock by uninformed investors. Fishman and Hagerty (1989) make a similar argument.

While a higher intensity of uninformed trading reduces the probability of trading against a privately informed investor, *ceteris paribus*, prior research indicates that greater uninformed trading attracts more informed trading. Kyle (1985) demonstrates that the amount of informed trading varies proportionately with the expected amount of uninformed, liquidity-based trading. The net result is that the *relative* amount of informed trading remains unchanged even as the expected amount of uninformed trading changes. However, to the extent that informed traders are risk

[3] The level of firm-investor information asymmetry is only relevant to the extent that it increases asymmetry among investors, such as through insider trading.

averse and capital constrained, we expect that the relative amount of informed trading will fall as uninformed trading increases. Accordingly, higher disclosure quality will be associated with relatively less informed trading, which in turn will reduce information asymmetry. Empirical evidence in Brown et al. (2004) supports this argument.

A second way disclosure quality affects information asymmetry is by altering the incentives to search for private information. Verrecchia (1982) examines a setting where public information disclosed by the firm is a perfect substitute for private information. He shows that the amount of costly private information that investors choose to acquire is generally decreasing in the amount of firm-disclosed public information. Diamond (1985) also finds that the incentives for investors to acquire private information are reduced when firms disclose information publicly.[4] Firms with high disclosure quality are more likely to publicly release material information promptly and provide forward-looking information. As such, we expect that higher disclosure quality reduces private information search incentives.

Prior empirical literature also suggests disclosure quality will be negatively related to the frequency of private information events. Gelb and Zarowin (2002) and Lundholm and Myers (2002) find that current stock returns reflect more information about future earnings when disclosure quality is higher. These results imply that by ''bringing the future forward,'' more informative disclosures reduce the total set of information about future earnings that can be privately discovered about a firm. Since there is less information available to be discovered, in addition to the reduced search incentives discussed above, we expect that the frequency of private information events will be declining in disclosure quality.

The above discussion suggests that public disclosure quality is negatively associated with information asymmetry because (1) although it is positively associated with the absolute amount of uninformed and informed trading, it is negatively associated with the relative amount of informed trading; and (2) it is negatively associated with the frequency with which informed investors discover and trade on private information. Figure 1 summarizes these hypothesized relations. The prior empirical literature has not examined these distinct mechanisms underlying the relation between disclosure quality and information asymmetry. We are able to do so using the PIN measure of information asymmetry, discussed below.

2 Variable measurement

2.1 Probability of informed trade

We rely on the EKO microstructure model (Easley et al., 1997) to proxy for the level of information asymmetry and the aspects of the firm's information and trading environments discussed above. The EKO model estimates the unconditional

[4] Also see Fama and Laffer (1971) and Hakansson (1977) for settings where public disclosure deters traders from expending resources on information collection for speculative purposes.

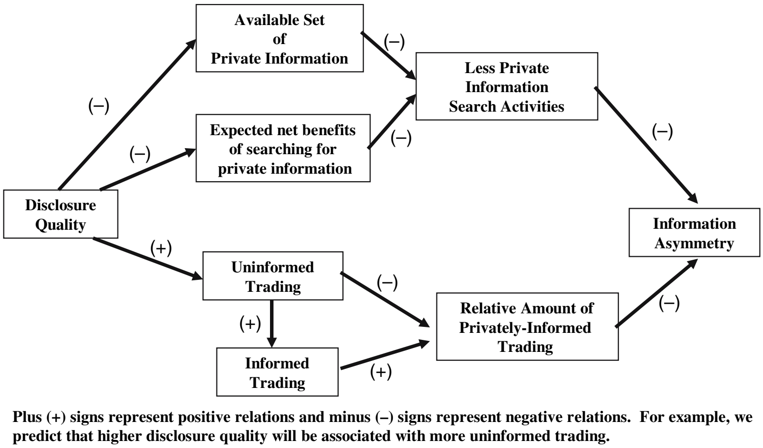

Fig. 1 Summary of predictions

probability of information-based trading (PIN) for a given stock based on the observed order flow. The model assumes that at the beginning of each trading day, a private information event occurs with probability α, whereby informed traders receive private information about the firm's value. The private information contains ''bad'' (''good'') news with probability δ $(1-\delta)$, where bad (good) news indicates that the profit maximizing trade is to sell (buy) the stock.[5] The news could either relate to private information of which the firm is aware, such as the outcomes of its own R&D projects or the terms of a new contract with a major supplier, or to private information unknown by the firm, such as a major customer's decision to defect to a rival or a rival's withdrawal from a contested market.

Buy and sell orders from uninformed traders are randomly submitted each day according to independent Poisson processes at the daily rate ε. On days with good (bad) news, informed buy (sell) orders also arrive at a rate proportional to the amount of uninformed trading, $\mu = \nu\varepsilon$. Accordingly, the relative amount of trading by privately informed investors is equal to $\mu/\varepsilon = \nu$. On a no-news day, both buy and sell orders arrive at the daily rate ε. On bad-news days, buys continue to arrive at the rate ε while sells arrive at a rate equal to $(\varepsilon + \mu) = (\varepsilon + \nu\varepsilon) = \varepsilon(1 + \nu)$; vice versa on good-news days.[6]

An important assumption of the model is that the daily arrival rates of uninformed buy and sell orders are drawn from independent Poisson distributions with constant parameters; as such, the daily numbers of uninformed buys and sells

[5] In assuming that the type of news is unambiguous, the EKO model does not allow for both informed buying and selling on the same day, as in Kim and Verrecchia (1991); private information is one-sided, as in Kyle (1985). In addition, the precision of the private information (vis-à-vis the precision of the public information) does not matter since informed traders are assumed to be risk neutral.

[6] The model ignores the size of each trade order. While this simplifying assumption likely results in the loss of some information, the reduction may be minor as informed investors disguise their information by mimicking the trade sizes of uninformed traders (Barclay & Warner, 1993; Chakravarty, 2001). Also, see Jones, Kaul, and Lipson (1994).

are uncorrelated. However, in practice, public information events (such as the release of macroeconomic statistics and earnings announcements) often affect the trading intensity of all uninformed traders—both buyers and sellers—on a particular day so that the daily arrival rates of uninformed buy and sell orders are positively correlated. Evidence in Venter and de Jongh (2004) strongly supports this contention. In addition, we find that the correlation between the daily numbers of total (i.e. from informed and uninformed) buy and total sell orders is significantly positive in our sample (average correlation is 0.37). This finding strongly contrasts with the basic EKO model, where the implied correlation is negative.

To relax this restrictive assumption, Venter and de Jongh (2004) model the arrival of uninformed buy and sell orders as a bivariate Inverse Gaussian Poisson process. In this extension of the EKO model, the average trading intensities of uninformed investors, both buyers and sellers, are subject to a daily scaling factor W_t, where W_t is drawn from an inverse Gaussian distribution with parameter $\psi > 0$.[7] High (low) values of W_t reflect days on which trading is generally high (low)—such as might occur shortly after (before) an earnings announcement; ψ is a measure of the variation in the average level of trading between the high- and low-trade days. Hence, the extended model allows for a positive correlation between the daily number of buy and sells. We summarize the way in which order flow arises in Fig. 2 and provide a more detailed description of the extended model in the Appendix.

The extended model's parameters $(\alpha, \delta, \psi, \varepsilon, \mu)$ are estimated by maximizing the likelihood function given in Eq. A5 in the Appendix using the daily number of buys and sells over an annual period as inputs.[8] PIN is calculated as follows, where $v = \mu/\varepsilon$:

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} = \frac{\alpha\varepsilon\upsilon}{\alpha\varepsilon\upsilon + 2\varepsilon} = \frac{\alpha\upsilon}{\alpha\upsilon + 2}. \tag{1}$$

Equation 1 shows that information asymmetry increases when private information events happen more frequently (α) and when the absolute and relative intensity of informed trading increases (μ and v), and decreases with the trading intensity of uninformed investors (ε).

[7] The Inverse Gaussian distribution has mean $=[W_t] = 1$ and variance $= \mathrm{Var}[W_t] = (1/\psi^2)$. As the variance of the daily scaling factor W_t approaches zero (or equivalently, as ψ approaches infinity), W becomes a constant equal to one. When this occurs, the extended model simply reduces to the basic EKO model. In our sample, the median (95th percentile) value of ψ is 2.7 (4.1), and only four observations have values of $\psi > 8$. Thus, the extended model fits the trade order data significantly better than the basic EKO model.

[8] This estimation procedure assumes that the underlying information environment, and hence parameters, remains relatively stable over an annual period. Easley et al. (2002) conclude that this assumption is reasonable as individual stocks exhibit relatively low variability in PINs across years and the cross-sectional distribution of PIN is quite stable across time. In addition, they estimated PINs using rolling 60 day sample periods and found them to be quite similar to the annual estimates. Thus, we conclude that it is reasonable to estimate PINs over an annual period.

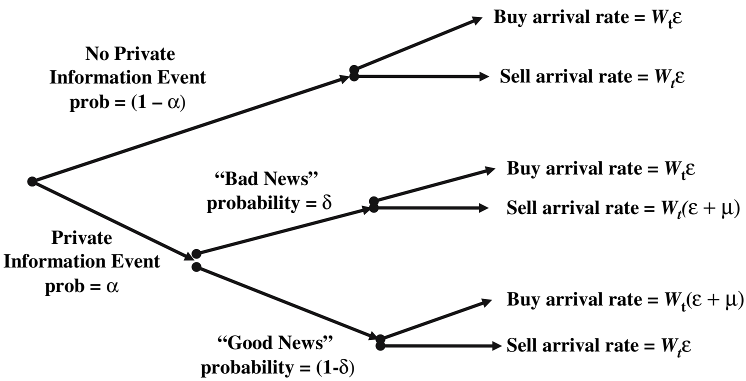

Fig. 2 Game tree of the Venter and de Jongh (2004) extension of the EKO model

2.2 Proxy for disclosure quality

We use the Association of Investment Management and Research (AIMR) total disclosure scores as our empirical proxy for a firm's disclosure quality.[9] The scores are intended to evaluate a ''firm's effectiveness in communicating with investors'' and the extent to which a firm's aggregate disclosures ensure that ''investors have the information necessary to make informed judgments.'' The AIMR formed industry-based committees composed of leading analysts to undertake a comprehensive evaluation of disclosure quality for a subset of firms in various industries.

The evaluation process typically results in a numerical score that represents the overall quality of the firm's disclosures throughout the year (*Total*). While the scores for a single industry-year are directly comparable, it is unclear to what extent that each analyst committee uses the same rating scale and criteria. Therefore, consistent with most of the prior literature, we restrict our analyses to examining intra-industry-year variation in disclosure quality and exclude what may be valid inter-industry-year variation.

The firms rated by AIMR tend to be larger, industry-leading firms with high analyst following and are generally thought to have higher (lower) and more uniform levels of disclosure quality (information asymmetry) compared to other firms. These characteristics reduce the variation in our sample as well as the size and significance of the estimated associations. Thus, while our sample is not fully representative of the entire population, we expect that the associations between disclosure quality and information asymmetry documented here are actually stronger for the general population of firms because we expect the variation in disclosure quality and information asymmetry to be much higher.

[9] Prior studies using the AIMR disclosure scores include Botosan and Plumlee (2002), Gelb and Zarowin (2002), Healy et al. (1999), Lang and Lundholm (1993), Lang and Lundholm (1996), Lundholm and Myers (2002), Sengupta (1998), and Welker (1995). Detailed discussions of the AIMR rating process and the disclosure scores can be found in Lang and Lundholm (1993) and Healy et al. (1999).

3 Methodology

We are interested in analyzing how disclosure quality is related to the level of information asymmetry. Economic theory and prior empirical evidence (Cohen, 2003; Leuz & Verrecchia, 2000; Marquardt & Wiedman,1998) indicate that these two variables are endogenously related. If better voluntary disclosure quality leads to less information asymmetry, then high asymmetry firms will have stronger incentives to choose higher disclosure quality to reduce the level of asymmetry, *ceteris paribus*. Failure to incorporate this endogeneity into our research design could result in misleading inferences. A common approach is to use 3 stage least squares (3SLS) in order to produce unbiased coefficient estimates. While this approach is theoretically appealing, it can be difficult to find appropriate exogenous variables with which to specify each equation. If the assumed exogenous variables are, in fact, correlated with both dependent variables, then the 3SLS results will be biased. To avoid this identification problem, we employ an alternative two-stage probit-based approach (Maddala, 1983; Wooldridge, 2002) used in Cohen (2003) and Leuz and Verrecchia (2000).[10]

In the first stage of the analysis, we use a probit estimation of disclosure quality where the dependent variable takes on a value of one if the firm's total disclosure quality score is above the median score for the industry-year, and equals zero otherwise. The independent variables consist of all the exogenous variables that affect either disclosure quality or information asymmetry. In the second stage, the fitted probabilities from the first-stage probit model, *PrTotal*, are included as an explanatory variable in the information asymmetry model. In effect, *PrTotal* acts as an instrumental variable for the actual disclosure quality score. Although this approach is less powerful, it avoids the identification issues of the 3SLS approach since the fitted probabilities are a non-linear function of the explanatory variables. Thus, our identifying variables do not have to be completely exogenous as is assumed in a 3SLS analysis (Wooldridge, 2002). In the second-stage estimation of the information asymmetry model, we obtain consistent and asymptotically efficient coefficient estimates using OLS.

Thus, our disclosure quality and information asymmetry models are as follows:

$$\begin{aligned}\text{Prob}(\mathit{Total}>\text{Industry-Year Median}) = \Phi(&\mathit{Size}, \mathit{Return}, \mathit{Surprise},\\ &\mathit{Correlation}, \mathit{Capital}, \mathit{InstOwn},\\ &\mathit{Analysts}, \mathit{Owners}, \mathit{EarnVol})\end{aligned} \tag{2}$$

$$\begin{aligned}\mathit{IAV} = \gamma_0 + \gamma_1 \mathit{Total} + \gamma_2 \mathit{Size} + \gamma_3 \mathit{InstOwn} + \gamma_4 \mathit{Analysts} + \gamma_5 \mathit{Dispersion}\\ + \gamma_6 \mathit{Leverage} + \gamma_7 \mathit{EarnVol} + \eta\end{aligned} \tag{3}$$

[10] Despite these potential difficulties, we also conduct a 3SLS analysis as an alternative methodological approach. The results from this approach are consistent with our main results and are discussed in Sect. 5.6.

Total represents the AIMR total disclosure quality score and *IAV* represents one of the information asymmetry variables $(PIN, \varepsilon, \mu, v, \alpha)$; firm and year subscripts are omitted. We discuss the independent variables in detail below.

3.1 Disclosure quality model

In addition to the level of information asymmetry, the previous literature identifies a number of other variables that are associated with firms' disclosure quality choices. Based on their survey of the literature, Lang and Lundholm (1993) identify the following variables as being related to disclosure quality: (1) *Size*—measured as the natural log of the firm's market value of equity as of the end of the firm's fiscal year. Bigger firms are expected to have higher disclosure quality because the benefits are expected to be higher while the costs are expected to be lower (Diamond, 1985). (2) *Return* is the absolute value of the market-adjusted stock return measured over the fiscal year and (3) *Surprise* is the absolute value of the difference between the firm's actual per share earnings and the consensus analyst forecast (scaled by price) measured eight months prior to fiscal year-end. To the extent that the level of firm-investor information asymmetry is increasing with performance variability, then the Expectations Adjustment Hypothesis (Ajinkya & Gift, 1984) predicts that firms with high performance variability will have higher disclosure incentives. Thus, we expect positive coefficients on *Return* and *Surprise*.[11] (4) *Correlation* is the correlation between annual stock returns and annual earnings measured over the previous ten years. Lang and Lundholm (1993) find a negative relation between disclosure quality and *Correlation*, inferring that a high correlation represents low levels of firm-investor asymmetry, and hence, lower incentives to disclose. (5) *Capital* is an indicator variable that equals one if the firm issues public debt or equity during the current or following 2 years, and zero otherwise. Firms have incentives to increase disclosure prior to raising capital in order to reduce the level of information asymmetry, and hence, the cost of capital (Frankel, McNichols, & Wilson, 1995).

We include four additional variables that we expect to be associated with disclosure quality based on our review of the more recent literature. *InstOwn* is the percentage of shares owned by institutional shareholders as of the end of the calendar year. *Analysts* is measured as the monthly average number of analysts in the annual consensus IBES forecast over the twelve month period starting eight months prior to the fiscal year end. *Owners* is the natural log of the number of registered shareholders as of the end of the fiscal year.[12] These three variables capture differences in shareholders' demands for disclosure quality and we expect them to have positive coefficients (Bushee, Matsumoto, & Miller, 2003). *EarnVol* is the standard deviation of earnings scaled by assets over the previous ten years. Firms with more volatile earnings face a greater risk of inaccurate forecasts and their associated litigation and reputation costs. Evidence in Brown, Hillegeist, and

[11] Lang and Lundholm (1993) use the signed values of *Return* and *Surprise* rather than the absolute values. Using signed values does not alter our results qualitatively.

[12] We measure *Size, InstOwn*, and *Owners* at the end of the fiscal year as an estimate of the average level of these variables during the annual measurement periods for *PIN* and the AIMR scores.

Lo (2005) and Waymire (1985) indicates firms make fewer forecasts when the volatility of earnings is higher. Therefore, we expect a negative coefficient on *EarnVol*. We also include *Dispersion* and *Leverage* in Eq. 2 because they are included as control variables in Eq. 3.

3.2 Information asymmetry model

In addition to disclosure quality, we expect several variables to be associated with the information asymmetry variables based on a review of the prior literature. Except where noted, we expect α, μ, and v (ε) to have the same (opposite) directional relation with the control variables as PIN does. Previous research indicates that stock prices incorporate information about large firms earlier than for small firms. Based on the results in Atiase (1985), Brown, Hillegeist, and Lo (2006), and Diamond and Verrecchia (1991), we expect a negative coefficient on *Size*.

Certain institutional investors undoubtedly trade based on private information (Bollen & Busse, 2005; Jiambalvo, Rajgopal, and Venkatachalam, 2002). To the extent that these institutional investors are present, we expect institutional ownership to be positively associated with higher values for α and μ. However, other types of institutional investors are unlikely to trade on private information. For example, S&P 500 index funds behave as uninformed investors by definition. To the extent uninformed institutions are present, we expect institutional ownership to be positively associated with ε. Thus, the expected associations between *PIN* or v and *InstOwn* are unclear.[13]

The relation between analyst following and information asymmetry is also complex. The results in Ayers and Freeman (2001) and Piotroski and Roulstone (2005) suggest that higher analyst following is associated with more trading by privately informed investors (and thus higher values for α and μ).[14] On the other hand, evidence in Brown et al. (2004) and Easley, O'Hara, and Paperman (1998) indicates that analyst coverage is positively associated with the amount of uninformed trading (ε). Thus, the expected associations between *PIN* and *Analyst* and v and *Analyst* are unclear.

Dispersion is a measure of uncertainty based on analyst forecasts and is measured as ln((standard deviation of forecast earnings per share in the 4th month of the fiscal year/stock price) + 0.001). When there is greater uncertainty regarding future earnings, more potential private information can be discovered and traded upon. However, a potentially offsetting effect is that the increased uncertainty makes it more costly to discover and profit from private information (Jiang, Lee, & Zhang, 2005). Therefore, we do not make a directional prediction.

Boot and Thakor (1993) argue that the incentives for private information acquisition are increasing with *Leverage*, the firm's debt-to-assets ratio measured at the end of the fiscal year. For a fixed amount of private information about the value

[13] Consistent with heterogeneity among institutional investors, Brown et al. (2004) find that the association between institutional ownership and PIN changes from quarter to quarter.

[14] However, Jiambalvo et al. (2002) find that the number of analysts is negatively related to the extent that prices reflect future earnings, which suggests that analyst following is negatively related to the amount of informed trading.

of a firm's assets, the expected profits from trading on that information in the equity market increase with the firm's leverage, *ceteris paribus*, which implies a positive association between *Leverage* and *PIN*. On the other hand, the Pecking Order theory of capital structure implies that there is a negative association between leverage and the amount of firm-investor information asymmetry. Therefore, the Pecking Order theory suggests a negative association between *PIN* and *Leverage*. Since these two arguments suggest different associations between *Leverage* and *PIN*, we do not make a directional prediction.

Finally, Zhang (2001) demonstrates that private information production increases with the volatility of earnings, *EarnVol*, because higher volatility increases the expected profits from trading on private information. In this case, we expect α and μ to increase with *EarnVol*. However, other arguments suggest that the expected benefits of private information may be decreasing with earnings volatility, causing the relation between *EarnVol* and *PIN* (α, μ) to be indeterminate. For example, firms with highly volatile earnings tend to have lower earnings response coefficients (due to less persistent earnings) and hence, the expected price effects per unit of earnings surprise are lower.

4 Sample description

Our sample is based on firms evaluated by the AIMR between 1986 and 1996, the last year for which the evaluations were published. Our sample consists of 2,204 firm-year observations representing 423 individual firms across 34 industries that have the required data. For each firm-year observation, we collect trade data from either the ISSM Transactions File or the Trades and Quotes database over the 12-month period beginning 8 months prior to the firm's fiscal year-end. We classify every trade as either buyer- or seller-initiated using the standard Lee–Ready algorithm (Lee & Ready, 1991). Based on the number of daily buys and sells for each trading day, we use Eq. A5 to compute the maximum likelihood estimates for the PIN parameters (α, δ, ψ, ε, μ). *PIN* is then calculated for each firm-year observation using Eq. 1.

Data for the control variables come from a variety of sources. Accounting data are obtained from COMPUSTAT and market prices and return data come from CRSP. Institutional ownership data are derived from the CDA/Spectrum 13F Institutional Holdings database, and SDC Platinum is the source for capital raising data. Analyst forecast data come from IBES.

Table 1 provides descriptive statistics for our sample. The mean (median) *PIN* is 18.6 (18.2), which indicates a roughly 18% chance that a trade is based on private information. The mean and median values of α indicate that private information events occur on just over half of all trading days. These values are generally consistent with those reported in prior literature that uses the basic EKO model. The mean value of ε indicates that the average number of uninformed trades (buys and sells) is almost 73 per day while there is an average of 28.1 trades by informed investors on private information event days. The average value of ν is 89%. This value indicates that informed trades are almost equal in intensity to the amount of

Table 1 Descriptive statistics for regression variables used in tests of the association between disclosure quality and information asymmetry

	Mean	Std Dev	Median	5%	95%
PIN model parameters					
PIN	19%	5%	18%	11%	28%
α	53%	12%	52%	36%	75%
ν	89%	30%	85%	49%	147%
ε	38.3	36.2	26.5	6.3	119.0
μ	28.1	20.3	22.3	6.6	72.7
ψ	38.9	850	2.7	1.6	4.1
Disclosure Scores					
Total	73%	13%	75%	49%	92%
Annual	75%	13%	77%	51%	93%
Quarterly	73%	15%	75%	46%	93%
IR	75%	16%	77%	45%	98%
Control variables					
Size ($m)	4,940	8,309	2,324	305	19,607
InstOwn	54%	15%	56%	25%	76%
Analysts	20	9	20	7	36
Dispersion	1.10	1.36	0.69	0.21	3.25
Leverage	25%	15%	24%	4%	50%
EarnVol	−3.8	1.0	−3.6	−5.7	−2.4
Return	18%	16%	14%	1%	50%
Surprise	0.04	0.08	0.01	0.00	0.16
Correlation	0.14	0.33	0.16	−0.45	0.68
Capital	0.6	0.5	1.0	0.0	1.0
Owners	41.3	74.5	18.5	2.6	166.0

Sample is based on 2,206 firm-year observations that have AIMR disclosure quality scores between 1986 and 1996. (1,776 firm-years for the sub-scores, *Annual, Quarterly* and *IR*.) *PIN* is the Probability of Informed Trade based on the Venter and de Jongh (2004) extension of the EKO model, and measured over the annual period beginning 8 months before the firm's fiscal year end and expressed as a percentage; α is the percentage of days on which private information events occur; ε (unlogged) is the average daily trading intensity of uninformed investors; μ (unlogged) is the average daily trading intensity of informed investors on private information event days; ν is the ratio of μ to ε; ψ is the variance parameter for the trading scale factor *W*. Disclosure scores are expressed as a percentage of the maximum score for the industry-year; *Total* is the overall disclosure score from AIMR; *Annual* is the score for 10-K related disclosures; *Quarterly* is the AIMR score for quarterly reports and other published information; *IR* is the AIMR score for investor relations activities; *Size* is the market value of the firm's equity at the end of the fiscal year (in $ millions); *InstOwn* is the percentage of shares owned by institutional shareholders at the end of the fiscal year; *Analysts* is the average number of analysts covering the firm from 8 months before fiscal year end to 4 months after fiscal year end; *Dispersion (unlogged)* is the standard deviation of analyst forecast earnings per share (measured 8 months before the fiscal year end) scaled by share price; *Leverage* is the firm's debt-to-assets ratio at the end of the fiscal year; *EarnVol* is the log of the standard deviation of earnings (scaled by assets) measured over the previous 10 fiscal years; *Return* is the market-adjusted stock return of the firm's equity measured over the fiscal year; *Surprise* is the difference between the firm's actual earnings per share and the consensus forecast measured eight months prior to the fiscal year end scaled by price; *Correlation* is the correlation between annual stock returns and annual earnings measured over the 10 years prior to the current fiscal period; *Capital* is an indicator variable equal to 1 if the firm issues public debt or equity during the current and following two-year period, and 0 otherwise; *Owners (unlogged)* is the number of registered shareholders (in thousands) at the end of the fiscal year uninformed trades in the same direction and represent about 31% of total trades on private information event days.[15] The median value of ψ is only 2.7 and more importantly, there are only four firm-years for which ψ is greater than 8.[16] Therefore, we conclude that in general, the extended model fits the data significantly better than the basic EKO model.

The AIMR scores presented in Table 1 represent the reported score as a percentage of the maximum possible score in each industry-year. The mean *Total* score is 73% and considerable variation occurs; the standard deviation is 13 and the 5th (95th) percentile value is 49% (92%). The three subscores, *Annual*, *Quarterly*, and *IR*, have similar averages and standard deviations. Table 1 indicates that the firms rated by the AIMR tend to be larger firms with significant analyst following (median = 20) and in which institutions typically hold over half of all outstanding shares. Ownership in these firms also tends to be widespread, with an average of over 41,000 registered shareholders.

Table 2 presents the Spearman correlations for the sample. As expected, α and v are both positively correlated with *PIN* (0.45 and 0.67, respectively) while ε is negatively correlated with *PIN* (−0.59). Somewhat surprisingly, there is a negative correlation between μ and *PIN* (−0.38). However, we expect that this correlation is caused by the high positive correlation between μ and ε (0.91). The correlations between *PIN* and the disclosure scores are significantly negative, although somewhat moderate in magnitude (between −0.11 and −0.14). The relatively low magnitude is expected because endogeneity will cause the cross-sectional correlation between them to be less negative.

5 Analysis and results

In this section, we report the results of cross-sectional analyses that investigate how disclosure quality is related to the level of information asymmetry. The first three sections discuss the relation between the *Total* disclosure quality score and *PIN* and the PIN parameters (ε, μ, v, and α). Next, we analyze whether the relation between disclosure quality and information asymmetry are the same across three different types of disclosure quality. We then examine the role of the market-to-book ratio and analyst coverage in determining the association between disclosure quality and information asymmetry. Finally, we discuss the results of a 3SLS specification that confirms and strengthens our main findings.

5.1 Disclosure quality and information asymmetry

We present the results from estimating the disclosure quality model in the left side of Table 3 where the explanatory variables include all of the variables in the disclosure quality and information asymmetry models (Eqs. 2, 3). The explanatory power of the

[15] The fraction of privately informed trades on information event days = $v/(v+2) = 0.89/2.89 \approx 0.31$.

[16] There are four extreme values where the estimated value of ψ is around 2,000—a corner solution that occurs when the basic EKO model describes the underlying trade data reasonably well. However, diagnostic tests confirm that these cases do not result in extreme estimates of PIN or the PIN parameters.

Table 2 Spearman rank correlation coefficients

	PIN	α	ν	ψ	ε	μ	*Total*	*Annual*	*Quarterly*	*IR*	*Size*	*InstOwn*	*Analysts*	*Dispersion*	*Leverage*	*EarnVol*	*Return*	*Surprise*	*Correlation*	*Capital*
α	0.45																			
ν	0.67	−0.26																		
ψ	−0.25	0.13	−0.38																	
ε	−0.59	−0.16	−0.54	0.23																
μ	−0.38	−0.32	−0.18	0.07	0.91															
Total	−0.13	0.02	−0.17	0.15	0.13	0.08														
Annual	−0.11	0.04	−0.16	0.13	0.14	0.09	0.85													
Quarterly	−0.11	0.01	−0.14	0.10	0.09	0.04	0.78	0.61												
IR	−0.14	0.01	−0.17	0.16	0.13	0.08	0.76	0.51	0.47											
Size	−0.51	−0.01	−0.57	0.44	0.70	0.55	0.25	0.24	0.14	0.28										
InstOwn	−0.12	0.07	−0.19	0.16	0.05	−0.04	0.11	0.08	0.13	0.12	0.12									
Analysts	−0.43	−0.01	−0.48	0.39	0.62	0.49	0.29	0.26	0.22	0.28	0.69	0.21								
Dispersion	0.12	−0.07	0.19	−0.25	−0.05	0.03	−0.18	−0.15	−0.10	−0.18	−0.32	−0.10	−0.19							
Leverage	−0.06	−0.07	−0.02	−0.08	0.11	0.13	0.01	−0.02	0.02	0.02	−0.08	0.05	0.07	0.27						
EarnVol	0.11	−0.06	0.18	−0.21	−0.07	0.01	−0.15	−0.16	−0.10	−0.14	−0.26	0.04	−0.20	0.27	0.03					
Return	0.06	−0.04	0.10	−0.14	0.00	0.05	−0.06	−0.04	−0.04	−0.05	−0.14	−0.07	−0.10	0.15	0.15	0.14				
Surprise	0.08	−0.04	0.14	−0.10	0.00	0.06	−0.15	−0.12	−0.09	−0.16	−0.25	−0.09	−0.14	0.42	0.23	0.18	0.23			

Table 2 continued

	PIN	α	v	ψ	ε	μ	*Total*	*Annual*	*Quarterly*	*IR*	*Size*	*InstOwn*	*Analysts*	*Dispersion*	*Leverage*	*EarnVol*	*Return*	*Surprise*	*Correlation*	*Capital*
Correlation	0.10	0.03	0.07	−0.04	−0.04	−0.01	−0.10	−0.05	−0.06	−0.13	−0.07	−0.03	−0.12	0.12	0.07	0.13	0.12	0.11		
Capital	−0.18	−0.06	−0.17	0.11	0.19	0.16	0.16	0.12	0.12	0.17	0.20	0.07	0.20	0.03	0.23	−0.11	0.03	0.01	0.00	
Owners	−0.34	−0.02	−0.39	0.31	0.64	0.57	0.15	0.15	0.09	0.18	0.65	−0.09	0.55	−0.02	0.05	−0.14	−0.05	0.01	−0.03	0.21

The sample consists of 1,775 firm-year observations with available data for all variables between 1986 and 1996. (Sample size is limited by the need for sub-scores *Annual*, *Quarterly* and *IR*. Correlations for the remaining variables in the full sample of 2,206 observations are not materially different.) Variable definitions are in Table 1 except that here and in subsequent tables, (1) ε and μ are the natural logs of the average daily trading intensity of uninformed trades and information-based trades, respectively; (2) *Size* is the natural log of market capitalization (in $million), *Owners* is the natural log of the number of shareholders (in thousands) and *EarnVol* is the natural log of the standard deviation of earnings (scaled by assets) measured over the previous 10 years; (3) *Dispersion* is the natural log of ((standard deviation of forecast earnings per share in the 4th month of fiscal year/stock price) +.001); and (4) all variables are industry-year adjusted by subtracting the mean value for the corresponding industry-year. Coefficients with absolute values greater than 0.05 are significant at the 0.01 level

Table 3 Coefficient estimates, *t*-statistics and *p*-values for tests of the endogenous association between the probability of informed trade and the probability that the *Total* disclosure score is above the industry-year median score

$$\text{Prob}(\textit{Total}{>}\text{Industry-Year Median}) = \Phi(\textit{Size}, \textit{Return}, \textit{Surprise}, \textit{Correlation}, \textit{Capital}, \textit{InstOwn}, \textit{Analysts}, \textit{Owners}, \textit{EarnVol}, \textit{Constant})$$

$$\textit{PIN} = \delta_0 + \delta_1 \textit{PrTotal} + \delta_2 \textit{Size} + \delta_3 \textit{InstOwn} + \delta_4 \textit{Analysts} + \delta_5 \textit{Dispersion} + \delta_6 \textit{Leverage} + \delta_7 \textit{EarnVol} + \zeta$$

Disclosure Quality Equation				*PIN* Equation			
Variable	Coefficient	*z*-stat	*p*-value	Variable	Coefficient	*t*-stat	*p*-value
Size (+)	−0.02	−0.38	*0.70*	*PrTotal* (−)	−2.80	−2.3	*<0.01*
Return (+)	−0.01	−0.03	*0.97*	*Size* (−)	−2.35	−16.9	*<0.01*
Surprise (+)	0.00	0.00	*1.00*	*InstOwn*(+/−)	−1.38	−2.0	*0.05*
Correlation (−)	−0.20	−1.96	*0.05*	*Analysts* (+/−)	−0.02	−1.0	*0.34*
Capital (+)	0.32	4.84	*<0.01*	*Dispersion* (+/−)	−0.43	−3.1	*<0.01*
InstOwn (+)	0.66	2.72	*<0.01*	*Leverage* (+/−)	−2.20	−3.5	*<0.01*
Analysts (+)	0.04	6.41	*<0.01*	*EarnVol* (+/−)	−0.47	−2.6	*0.01*
Owners (+)	0.08	1.79	*0.07*				
EarnVol (−)	−0.20	−3.81	*<0.01*	Adj.-R^2	41.2%		
Dispersion	−0.15	−2.56	*0.01*				
Leverage	−0.02	−0.09	0.93				
psuedo-R^2	8.2%						

PIN is the probability of informed trade based on the Venter and de Jongh (2004) extension of the EKO model, and measured over the annual period beginning 8 months before the firm's fiscal year end and expressed as a percentage. Other variable definitions are in Tables 1 and 2. Φ represents the standard normal distribution. The Disclosure Quality equation is a probit regression of the probability that the *Total* disclosure score is above the industry-year median *Total* disclosure score. *PrTotal* is the fitted probability from estimating the Disclosure Quality equation. *p*-Values are based on one-sided or two-sided values where appropriate. Coefficients on the constants are not reported. Number of observations after elimination of outlier observations with abs(dfits) > 0.1 = 1,987

model is somewhat modest as the pseudo-R^2 is 8.2%. Seven of the eleven coefficients are significant at the 7% level or better. The coefficients on *Capital, InstOwn*, *Analysts*, and *Owners* are positive and significant, as expected. Also consistent with expectations, *EarnVol* has a significantly negative coefficient. In addition, the *Dispersion* coefficient from the information asymmetry model is negative and significant. The *Size, Return, Surprise*, and *Leverage* coefficients are insignificant. The lack of significance for *Size* and *Return* is surprising given that Lang and Lundholm (1993) find that both variables are significantly associated with the AIMR disclosure quality scores. Untabulated analyses indicate that the associations are affected by our inclusion of *Analysts* in the regression, which Lang and Lundholm do not include. When *Analysts* is excluded from the disclosure quality model, both the *Size* and *Return* coefficients are significantly positive, as expected.

In the second step, we regress *PIN* on *PrTotal* and the control variables in the information asymmetry model (Eq. 3), where *PrTotal* is the fitted probability that the firm's *Total* disclosure quality score is greater than the median industry-year score based on the estimated coefficients from the disclosure quality model. The results from estimating this model are presented in the right side of Table 3. The adjusted-R^2 for the regression is 41.2%. The *PrTotal* coefficient is negative and significant at the 1% level.[17] This finding supports our hypothesized negative relation between disclosure quality and the level of information asymmetry among investors after controlling for the potentially endogenous relation between the two variables. The magnitude of the *PrTotal* coefficient (−2.80) indicates that an increase in the probability of the firm having an above-median total disclosure score from 0.25 to 0.75 will lead to a decrease in PIN of 2.80/2 = 1.4 percentage points. This decline represents an economically significant decrease in PIN of 7.4% (7.8%) for the mean (median) firm in our sample. Combined with the findings in Easley et al. (2002) on the association between PIN and the cost of equity capital, a 1.4 percentage point reduction in PIN is associated with a 35 basis point reduction in the cost of capital.

Examining the results for the control variables, Table 3 shows that the *Size* coefficient has the predicted negative sign and is highly significant (t-statistic = −16.9). The coefficients on *Leverage* and *EarnVol* are also negative and significant. The negative coefficient on *Leverage* is consistent with the predictions of the Pecking Order theory and is inconsistent with the arguments in Boot and Thakor (1993). The *InstOwn* coefficient is significantly negative (t-statistic = −2.0).[18] This finding is inconsistent with the popular notion that all institutions are sophisticated investors who frequently trade on the basis of private information. We discuss the role of institutional ownership in more detail below in the context of the PIN parameters.

Table 3 shows that *Analysts* is insignificant (t-statistic = −1.0). One possible explanation for the lack of significance is that the effect of *Analysts* is subsumed into that of *PrTotal* since Lang and Lundholm (1996) find that analyst following is increasing in disclosure quality. Untabulated results show that the correlation between *PrTotal* and *Analysts* is 0.70.[19] We find that when *PrTotal* is excluded from the *PIN* equation, there is a strong negative relation between *PIN* and *Analysts*. In addition, the statistical significance of the *PrTotal* coefficient increases when *Analysts* is excluded from the regression.

5.2 Disclosure quality and trading behavior

To gain a deeper understanding about why there is a negative association between disclosure quality and information asymmetry, we exploit the EKO model to examine the relation between disclosure quality and the absolute and relative trading behavior of informed and uninformed investors. These analyses involve the

[17] All reported t-statistics are based on clustered standard errors where outliers based on abs(dfits) >0.1 are eliminated (Belsley, Kuh, & Welsch, 1980).

[18] While Brown et al. (2004) report a significantly positive coefficient on institutional ownership in their pooled sample where PIN is estimated using the basic EKO model, institutional ownership has a significantly negative coefficient in several of their quarterly regressions.

[19] The correlation between *Analysts* and *Size* is also quite large (0.69), and this multi-collinearity could also be responsible for the lack of significance for *Analysts*.

following EKO model parameters: ε–the average daily trading intensity of uninformed buyers and sellers; μ–the average daily trading intensity of informed investors on private information event days; and v–the relative amount of informed trading. For each dependent variable, we use the same two-step estimation procedure as we used for the *PIN* analysis presented in Table 3.

In Table 4A, we present the results for the information asymmetry model where (the natural log of) ε is the dependent variable. The *PrTotal* coefficient is significant (t-statistic = 2.6) and positive (0.40), as expected. In addition, all of the control variables are highly significant with p-values less than 0.01. The results indicate that firms with higher disclosure quality experience more trading by uninformed investors. The estimated effect is economically significant: an increase in the probability of the firm having an above-the-median total disclosure score from 0.25 to 0.75 increases the number of uninformed trades by approximately 22%

Table 4 Coefficient estimates, t-statistics and p-values for tests of the endogenous association between the PIN parameters and the probability that the *Total* disclosure score is above the industry-year median score

$$IAV = \gamma_0 + \gamma_1 PrTotal + \gamma_2 Size + \gamma_3 InstOwn + \gamma_4 Analysts + \gamma_5 Dispersion + \gamma_6 Leverage + \gamma_7 EarnVol + \eta$$

Variable	*PrTotal*	*Size*	*InstOwn*	*Analysts*	*Dispersion*	*Leverage*	*EarnVol*
ε **Equation**:	(+)						
Coefficient	0.40	0.56	−0.46	0.02	0.23	0.49	0.15
t-statistic	2.6	26.4	−4.9	5.7	12.0	5.5	6.6
p-value	*0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*
μ **Equation**:	(+)						
Coefficient	0.41	0.32	−0.61	0.01	0.24	0.32	0.15
t-statistic	*2.6*	*19.3*	*−6.8*	*4.4*	*12.4*	*3.8*	*6.8*
p-value	*0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*
v **Equation**:	(−)						
Coefficient	−5.42	−12.78	−19.44	−0.33	−0.28	−10.02	0.41
t-statistic	*−0.8*	*−19.6*	*−5.6*	*−2.3*	*−0.4*	*−3.0*	*0.4*
p-value	*0.21*	*<0.01*	*<0.01*	*0.03*	*0.68*	*<0.01*	*0.68*
α **Equation:**	(−)						
Coefficient	−8.22	−0.03	6.06	0.10	−1.22	−3.17	−1.65
t-statistic	−2.5	−0.1	3.7	1.6	−3.4	−2.1	−4.0
p-value	*0.01*	*0.94*	*<0.01*	*0.11*	*<0.01*	*0.04*	*<0.01*

α is the percentage of days on which private information events occur; ε is the natural log of the average daily trading intensity of uninformed investors; μ is the natural log of the average daily trading intensity of informed investors on private information event days; v is the ratio of μ (unlogged) to ε(unlogged). Other variable definitions are in Tables 1 and 2. *PrTotal* is the fitted probability that the total disclosure score is above the industry-year median total disclosure score based on the results of estimating the Disclosure Quality equation presented in Table 3. *IAV* refers to one of the four PIN parameters, ε, μ, v, and α, depending on the specification. p-values are based on one-sided or two-sided values where appropriate. The constant is not tabulated. Number of observations after elimination of outlier observations with abs(dfits) >0.1 vary from 1,968 to 1,991 (exp(0.40* (0.75 ,– 0.25)) = 1.22). These findings are consistent with the arguments in Fishman and Hagerty (1989) and Merton (1987) that higher disclosure quality reduces the costs of processing public information about the firm, resulting in more non-privately informed investors. One possible explanation is that uninformed investors are attracted to and have higher confidence in firms that consistently provide high quality disclosures, which reduces, *ceteris paribus*, the risk of trading against a privately informed investor. Our findings also provide some support for regulators' beliefs that high quality disclosures make the capital markets more attractive to ''ordinary'' uninformed investors (FASB, 2001; FASC, 1998; Levitt, 1998).

The results for where (the natural log of) μ is the dependent variable in the information asymmetry model are presented in Table 4B. The μ coefficient is positive (0.41) and significant (t-statistic = 2.6), as expected. The magnitude of the coefficient indicates that if *PrTotal* changes from 0.25 to 0.75, the average daily number of informed trades increases by 23% (exp(0.41*(0.75 − 0.25)) = 1.23). This finding suggests that informed investors increase their trading intensity when the level of uninformed trading increases, which is consistent with Kyle (1985). In addition, all of the control variables are significant with p-values less than 0.01 and have the same signs as in the ε regression in Panel A.

Table 4A, B show that, somewhat surprisingly, institutional ownership is negatively associated with both the amount of uninformed and informed trading. Recall that the PIN parameters are estimated based on the number of trades, not the volume of shares traded. Thus, one possible explanation for our results is that while institutions generally trade larger blocks of shares than individuals, they engage in relatively fewer transactions overall. Consistent with this conjecture, Bushee (1998) finds that certain classes of institutional investors engage in longer-term buy-and-hold strategies which result in low rates of trading. Together, these observations likely explain the negative association between *InstOwn* and ε (μ).

Recall that v equals μ/ε and hence, measures the relative amount of trading by informed investors on private information event days. The results for where v is the dependent variable in the information asymmetry equation are presented in Table 4C. The arguments in Sect. 1 suggest that to the extent informed investors are capital constrained and/or risk averse, there will be a negative relation between disclosure quality and v; otherwise, there will be no association between them. Table 4 shows that while the *PrTotal* coefficient is negative (−5.42), it is not significantly different from zero (p-value = 0.21). We infer from the combined results in Panels A, B, and C that although the amount of trading by both informed and uninformed investors increases in disclosure quality, the ratio of the two is unchanged, resulting in no significant association between the relative amount of informed trading and disclosure quality.

The results in Panel C also show that the coefficients on *Size* and *InstOwn* are negative and highly significant (t-statistics = −19.6 and −5.6, respectively). The *Size* coefficient indicates that despite the high absolute level of trading noise that disguises privately informed trade in large firms, there is actually relatively less informed trading in large stocks. The negative *InstOwn* coefficient indicates that on average, the relative amount of informed trading decreases in the level of institutional ownership. While contrary to the notion that higher institutional ownership leads to more informed trading, our finding is consistent with a substantial fraction of institutional ownership consisting of institutions such as index funds that do not trade on the basis of short-term private information.[20] The findings in Ke, Ramalingegowda, and Yu (2006) and Yan and Zhang (2006) also support this interpretation. They find no evidence to suggest that ''buy-and-hold'' institutional investors with long investment horizons earn positive abnormal returns. Together, this evidence suggests that the negative relation between *InstOwn* and the relative amount of informed trading is due to the uninformed trading of long-term institutional investors. It also provides an explanation for the negative coefficient on *InstOwn* in the *PIN* equation reported in Table 3.

The coefficient on *Analysts* in Panel C is significantly negative (p-value = 0.03), indicating that the relative amount of informed trading is decreasing with analyst coverage. Collectively, the results in Panels A, B, and C indicate that while analyst following is positively associated with both informed and uninformed trading intensities, the increase in uninformed trading dominates, resulting in relatively less privately informed trading.[21]

5.3 Disclosure quality and the frequency of private information events

Our second analysis of how disclosure quality is related to information asymmetry examines the association between disclosure quality and the frequency of private information events. As with the analyses of the other PIN parameters, we use the same two-step estimation procedure that we used for the PIN analysis in Table 3. The results when α is the dependent variable in the information asymmetry model are presented in Table 4D. The results show that the coefficient on *PrTotal* is negative and significant (t-statistic = -2.5), as expected. The magnitude of the coefficient (-8.2) indicates that if *PrTotal* increases by 50 percentage points, the daily probability of a private information event occurring falls by 4.1%. This amount is economically significant and represents about a 7.7% (7.9%) decrease for the mean (median) firm; for a typical firm, it implies that there will be approximately 10 fewer days per year on which privately informed trading occurs. This result suggests that firms can reduce the frequency of private information events by pursuing high quality disclosure policies. While our analyses are based on differences in voluntary disclosure quality, they may also be applicable to regulators and exchanges contemplating mandatory changes in disclosure quality. Assuming that the frequency of private information events corresponds to the amount of (non-productive) private information search activities, then higher quality disclosures can improve aggregate social welfare by reducing socially wasteful search costs (before considering the costs of disclosure).

[20] Malkiel and Radisich (2001) report that throughout the 1990s, 30% of institutionally managed assets were indexed, suggesting that a substantial amount of institutional trading was not based on private information.

[21] In comparing a small group of high- and low-analyst firms, Easley et al. (1998) provides univariate evidence that both uninformed and informed trading is higher in the high-analyst group than in the low-analyst group. However, they do not discuss or analyze the ratio of the two variables.

Examining the control variables, we find that *Size* is not significantly associated with the frequency of private information events, which is somewhat surprising given *Size*'s significance in the prior results. However, it is consistent with the univariate results in Table 2 where the correlation between α and *Size* is insignificant. In contrast, we find a significantly positive relation between *InstOwn* and α. This association is consistent with at least some proportion of institutional investors trading on private information. For example, we expect that ''transient'' institutional investors (as identified in Bushee and Noe (2000)) actively trade on private information. In addition, evidence in Ke et al. (2006) and Yan and Zhang (2006) suggests that only actively trading institutional investors with short investment horizons are able to generate positive abnormal returns, which is consistent with their trading on the basis of private information that is quickly impounded into prices. Accordingly, one interpretation of our finding that *InstOwn* is positively associated with the frequency of private information events is that it results from short-term institutions trading on private information.

5.4 Information asymmetry and different types of disclosure quality

As discussed in the AIMR reports, the total disclosure quality score aggregates the evaluation of three distinct types of disclosures made by firms: (1) The *Annual* score reflects the quality of the 10-K and other annual published information; (2) The *Quarterly* score reflects the quality of the firm's quarterly reports and other published information, such as proxy statements and press releases; and (3) The *IR* score reflects the investor relations activities and is primarily based on the firm's interactions with analysts. In this section, we jointly analyze how each of the three subscores, *Annual, Quarterly*, and *IR*, are associated with the information asymmetry variables.

Disclosure quality depends on several attributes of the information being disclosed, each of which is likely to be related to the level of information asymmetry. While there is no widely-accepted definition of disclosure quality, we believe that important attributes of disclosure quality include the quantity of value-relevant information that is conveyed, its timeliness, precision, credibility, and how widely it is disclosed. As discussed below, we do not expect that any of the three types of disclosures will rank higher than the others with respect to all of these attributes. Thus, we do not predict which type of disclosure quality will have the strongest (weakest) association with the information asymmetry variables.

While the total quantity of information disclosed in the annual report is quite large, our empirical tests rely on differences in disclosure quality. Since many of the items contained in the annual reports are mandatory, it may be that the real differences in quality, which reflect voluntary differences in disclosure quality, are too small to generate significant differences in information asymmetry. Yet the emphasis on the annual report by the AIMR evaluation committees argues against this interpretation and the committee reports document substantial intra-industry differences in the extent of disclosures contained in the annual reports. Annual reports also rank high in terms of credibility (since they are audited and subject to litigation) and precision (due to the detailed, quantitative nature of many of the disclosures). In addition, annual reports are broadly disseminated among the public. Despite these positive attributes, annual reports are often criticized for their lack of timeliness since by the time they are publicly released; much of their information has already been conveyed through more timely channels. Combined with their historical emphasis, this lack of timeliness will reduce (differences in) the ability of annual report quality to affect the level of information asymmetry.

In many respects, the attributes of the investor relations activities reflected in the *IR* score contrast sharply with those of the annual report. IR activities are purely voluntary, exhibit a high degree of timeliness, and often take a forward-looking perspective. The importance attached to them by analysts indicates that they are an important source of information. However, IR activities have two attributes that will limit their association with information asymmetry. First, these disclosures are less credible because they are often disclosed verbally and represent non-quantifiable and non-verifiable information (such as the degree of optimism held by executives). Second, many, if not most, of these disclosures were made through private communications with analysts during our pre-Regulation FD sample period. As such, higher quality IR disclosures will temporarily increase the level of information asymmetry if the selective disclosures create private information events. However, if private disclosures to analysts (singly or in groups) concentrate the release of private information, then they could still lead to an overall decrease in information asymmetry.

The disclosures represented by the *Quarterly* scores fall between the other two types of disclosures along most of the attributes. They are timelier than the annual reports but will generally be less timely than the IR disclosures. While not audited, their credibility is still quite high given that they are official public documents that are subject to review by the auditors and litigation risk. While both the quarterly and annual reports are broadly disseminated, the quantity of information disclosed and its precision is likely lower for the quarterly reports since there is less supplementary and supporting material.

Our sample size is reduced from 2,206 to 1,775 observations since the AIMR did not provide the three subscores for all industry-years. Table 2 shows that the subscores are highly correlated with each other with the correlations ranging between 0.47 (*Quarterly* and *IR*) and 0.61 (*Annual* and *Quarterly*). The high correlations suggest that firms choose the quality of their disclosures in a consistent manner. *PrAnnual, PrQuarterly*, and *PrIR* are the fitted values from unreported probit regressions corresponding to Eq. 2. We replicate each of the analyses in Tables 3 and 4 substituting in the three predicted subscores in place of *PrTotal*. Since the fitted subscores are very highly correlated—greater than 0.86—it is important to include all three variables in the same regression; otherwise, the reported coefficients will be biased. We report the results of these analyses in Table 5, along with the corresponding *PrTotal* results from Tables 3 and 4 for comparison purposes. For brevity, we report only the coefficients and test statistics for the disclosure variables.

Examining the results where *PIN* is the dependent variable, we find that the coefficients on *PrAnnual* and *PrIR* are significantly negative (t-statistics = −2.6 and −2.9, respectively) as expected; the negative coefficients are consistent with the

Table 5 Coefficient estimates, *t*-statistics and *p*-values for tests of the endogenous association between the PIN and PIN parameters and the probability that each *SubScore (Annual*, *Quarterly*, or *IR)* is above the industry-year median *SubScore*

$$IAV = \theta_0 + \theta_1 PrAnnual + \theta_2 PrQuarterly + \theta_3 PrIR + \theta_4 Size + \theta_5 InstOwn + \theta_6 Analysts + \theta_7 Dispersion + \theta_8 Leverage + \theta_9 EarnVol + \xi$$

Variable	*PrTotal*[a]	*PrAnnual*	*PrQuarterly*	*PrIR*	*F test that* $\theta_1 = \theta_2 = \theta_3 = 0$	*F test that* $\theta_1 = \theta_2 = \theta_3$
PIN **Equation:**	(−)	(−)	(−)	(−)		
Coefficient	−2.80	−9.41	11.41	−6.10		
t-statistic	−2.3	−2.6	2.6	−2.9	6.82	4.78
p-value	*0.01*	*0.01*	*0.01*	*<0.01*	*<0.01*	*0.01*
ε **Equation:**	(+)	(+)	(+)	(+)		
Coefficient	0.40	2.11	−0.40	−0.57		
t-statistic	2.6	4.0	−0.7	−2.4	8.04	11.85
p-value	*<0.01*	*<0.01*	*0.50*	*0.02*	*<0.01*	*<0.01*
μ **Equation:**	(+)	(+)	(+)	(+)		
Coefficient	0.41	2.19	−0.23	−0.84		
t-statistic	2.6	4.5	−0.4	−3.5	10.67	15.97
p-value	*<0.01*	*<0.01*	*0.66*	*<0.01*	*<0.01*	*<0.01*
v **Equation:**	(−)	(−)	(−)	(−)		
Coefficient	−5.42	24.9	−26.45	−6.40		
t-statistic	−0.8	1.3	−1.3	−0.6	2.28	1.22
p-value	*0.21*	*0.21*	*0.10*	*0.52*	*0.08*	*0.30*
α **Equation:**	(−)	(−)	(−)	(−)		
Coefficient	−8.22	−49.01	39.76	−10.89		
t-statistic	−2.5	−5.1	3.5	−2.0	12.01	12.25
p-value	*0.01*	*<0.01*	*<0.01*	*0.02*	*<0.01*	*<0.01*

PIN is the probability of informed trade based on the Venter and de Jongh (2004) extension of the EKO model, and measured over the annual period beginning 8 months before the firm's fiscal year end and expressed as a percentage; α is the percentage of days on which private information events occur; ε is the natural log of the average daily trading intensity of uninformed investors; μ is the natural log of the average daily trading intensity of informed investors on private information event days; *v* is the ratio of μ (unlogged) to ε (unlogged). Other variable definitions are in Table 1. *Score* refers to one of the three AIMR disclosure subscores, *Annual*, *Quarterly*, and *IR*, depending on the specification. *PrScore* is the fitted probability that *Score* is above the industry-year median *Score* based on the results of estimating the disclosure quality equation where *Score* is the dependent variable. *IAV* refers to one of the information asymmetry variables, *PIN*, ε, μ, *v*, and α, depending on the specification. The results for the control variables are not reported. *p*-values are based on one-sided or two-sided values where appropriate. Before weeding of outliers, *N* = 1,775; after weeding outliers identified as abs(dfits) > 0.1, *N* is between 1,486 and 1,531

[a] Coefficients on *Total* are repeated from Tables 3 and 4 for ease of comparison

PrTotal results. However, the *PrQuarterly* coefficient is unexpectedly positive and significant (*t*-statistic = 2.6). *F*-tests reject the null hypotheses that all three subscore coefficients are equal to zero and that the three coefficients are equal to each other.

The unexpectedly positive coefficient on *PrQuarterly* suggests that after controlling for the quality of the annual report and investor relations activities, higher quality quarterly reports actually result in higher levels of information asymmetry. While inconsistent with our hypothesis, it is consistent with managers' claims that higher quality disclosures result in increased stock price volatility to the extent higher volatility is driven by more frequent private information events (also see Bushee & Noe, 2000). In addition, Botosan and Plumlee (2002) find an unexpectedly negative association between quarterly report quality and the estimated cost of equity capital. This result is consistent with ours to the extent that information asymmetry is a priced risk factor.[22]

The *PrAnnual* coefficient (−9.41) is larger in magnitude than the *PrIR* coefficient (−6.10). However, an unreported *F*-test indicates that the two coefficients are not significantly different from each other. We conjecture that the importance of annual report quality is due to the large quantity of information contained in the report and its high level of credibility. Our results also suggest criticisms that annual reports are too ''boilerplate'' to reflect meaningful differences in quality are unjustified. The importance of firms' interactions with analysts likely arises from the broad range and timeliness of the information disclosed. While these communications are typically informal and not subject to litigation concerns, reputational concerns of managers serve to enhance their credibility.

Examining the results where the PIN parameters are the dependent variables reveals a similarly ambiguous pattern for the subscore coefficients compared to the *PrTotal* results. While the *PrAnnual* coefficients have the same signs and similar significance levels as the corresponding *PrTotal* coefficients in the ε and μ equations, the *PrQuarterly* and *PrIR* coefficients are either insignificant or are significant in the opposite direction. However, consistent with the results for *PrTotal*, none of the subscore coefficients are significantly different from zero in the v equation, consistent with the two effects offsetting each other.[23] One possible interpretation of the negative *PrIR* coefficients in the ε and μ equations is that in the pre-Regulation FD period, higher quality but selective disclosures to analysts were perceived by uninformed investors as potentially disadvantaging them, and consequently, they traded less frequently in these stocks.

The results in Panel E show that the frequency of private information events is significantly and negatively associated with both annual report and investor relations disclosure quality, but is positively associated with quarterly disclosure quality. These results are robust to alternate methods of elimination of influential

[22] Botosan and Plumlee (2002) also find a negative association between annual report quality and the estimated cost of equity capital, consistent with the negative *PrAnnual* coefficient we report here. However, they do not find a significant association with the IR score, in contrast with the significantly negative association we find here.

[23] The levels of statistical significance that we find on the *PrQuarterly* and *PrIR* coefficients are somewhat sensitive to deletion of outliers. However, in all cases the coefficients are negative in the ε and μ equation and in no case were they significantly different from zero (or from each other) in the v equation.

observations. Together with the insignificant results in the v equation, our findings for α suggest that the associations between the subscores and *PIN* documented in Panel A are driven primarily by the associations between the subscores and α.

5.5 Role of market-to-book on the association between disclosure quality and information asymmetry

In this section, we examine the role of the market-to-book (M/B) ratio on the association between disclosure quality and information asymmetry. The M/B ratio is increasing in the amount of unrecorded intangible assets and the amount of growth opportunities. Firms with these characteristics typically have higher amounts of information asymmetry. For example, Aboody and Lev (2000) find that trades by insiders at R&D intensive firms, which will have higher M/B, *ceteris paribus*, are substantially more profitable than insider trades at non-R&D intensive firms. Barth, Kasznik, and McNichols (2001) find that analyst coverage is higher for firms in industries with high levels of intangible assets and that analysts expend more effort in analyzing firms with more intangible assets. Both of these findings imply that the benefits to producing information about firms is increasing in the level of firm-investor information asymmetry and in how much inherent uncertainty there is about firm value. Accordingly, we expect that firms in industries characterized by high levels of M/B will exhibit a stronger negative association between disclosure quality and information asymmetry than firms in other industries.

We calculate the average M/B ratio for each industry-year in our sample and define the following indicator variables: *Hi_M/B* (*Lo_M/B*) equals one when the average value of *M/B* for the firm's industry-year is greater than (less than) the sample median value, and zero otherwise. We analyze whether the association between disclosure quality and information asymmetry varies across industries based on their market-to-book ratio using the following regression model, where *IAV* represents one of the information asymmetry variables, *PIN*, ε, μ, v, or α, depending on the specification:

$$\begin{aligned} IAV = {} & \gamma_0 + \gamma_1 PrTotal^{*} Hi\_M/B + \gamma_2 PrTotal^{*} Lo\_M/B + \gamma_3 Size + \gamma_4 InstOwn \\ & + \gamma_5 Analysts + \gamma_6 Dispersion + \gamma_7 Leverage + \gamma_8 EarnVol + \eta \end{aligned} \tag{4}$$

The estimated γ_1 and γ_2 coefficients are presented in Table 6. For each of the five dependent variables, the coefficients have the expected signs and 8 of 10 are significantly different from zero at the 6% level or better. When *PIN* is the dependent variable, the *PrTotal*Hi_M/B* coefficient is -3.54 (p-value < 0.01) while the coefficient for *PrTotal*Lo_M/B* is -2.09 (p-value $= 0.05$). An F-test rejects the null hypothesis that both coefficients are equal at the 8% level. This finding indicates that disclosure quality is more negatively related to the level of information asymmetry in settings where the usefulness of firms' disclosures in reducing information asymmetry between investors is expected to be higher.

Examining the results for when ε is the dependent variable, Table 6 shows that while the *PrTotal*Hi_M/B* coefficient is larger than the *PrTotal*Lo_M/B* coefficient (0.45 vs. 0.33), an F-test indicates that the difference is not significant

Table 6 Coefficient estimates, *t*-statistics and *p*-values for tests of the endogenous association between the information asymmetry variables and the probability that disclosure quality is above the industry-year median level conditional on the market-to-book ratio

$$IAV = \gamma_0 + \gamma_1 PrTotal^{*} Hi\_M/B + \gamma_2 PrTotal * Lo\_M/B + \gamma_3 Size + \gamma_3 InstOwn + \gamma_5 Analysts + \gamma_6 Dispersion + \gamma_7 Leverage + \gamma_8 EarnVol + \eta$$

Variable	*PrTotal* * *Hi_M/B*	*PrTotal* * *Lo_M/B*	*F test that* $\gamma_1 = \gamma_2 = 0$	*F test that* $\gamma_1 = \gamma_2$
PIN **Equation:**	(−)	(−)		
Coefficient	−3.54	−2.09		
t-statistic	−2.9	−1.7	4.54	2.99
p-value	*<0.01*	*0.05*	*0.01*	*0.08*
ε **Equation:**	(+)	(+)		
Coefficient	0.45	0.33		
t-statistic	2.8	2.2	21.85	3.81
p-value	*<0.01*	*0.01*	*0.02*	*0.31*
μ **Equation:**	(+)	(+)		
Coefficient	0.45	0.23		
t-statistic	3.0	1.5	5.02	4.66
p-value	*<0.01*	*0.06*	*0.01*	*0.03*
v **Equation:**	(−)	(−)		
Coefficient	−0.06	−5.6		
t-statistic	−0.0	−0.8	0.89	1.64
p-value	*0.45*	*0.2*	*0.41*	*0.20*
α **Equation:**	(−)	(−)		
Coefficient	−12.62	−6.70		
t-statistic	−3.7	−2.0	8.68	8.90
p-value	*<0.01*	*0.02*	*<0.01*	*<0.01*

Variable definitions are in Tables 1 and 2. *PrTotal* is the fitted probability that the total disclosure score is above the industry-year median total disclosure score that is based on the results of estimating the disclosure quality equation presented in Table 3. *Hi_M/B* is an indicator variable set equal to 1 if the ratio of the aggregate market value of firms in the industry-year to book value of those firms is above the median value for all industry-years within the sample and zero otherwise. *Lo_M/B* is defined similarly for firms in industry-years that are below the median M/B for all industry-years within the sample. *IAV* refers to PIN or one of the four underlying parameters (ε, μ, v and α) from the EKO microstructure model, depending on the specification. *p*-values are based on one-sided values. Parameter estimates for control variables are not tabulated; they are not materially different from those reported in Tables 3 and 4. Number of observations after elimination of outlier observations with abs(dfits) >0.1 vary from 1,884 to 1,938

(*p*-value = 0.31). In contrast, the results for the μ regression show that the association between disclosure quality and informed trading is significantly stronger (*p*-value = 0.03) in high M/B industries. However, Table 6 shows that when v is the dependent variable, the *PrTotal**Lo_M/B* and *PrTotal**Hi_M/B* coefficients are not significantly different from zero and they are not significantly different from each other (p-value = 0.20). Thus, we do not find evidence suggesting that the association between disclosure quality and the relative amount of informed trading varies with the industry-year market-to-book ratio.

The results for the α equation in Table 6 show that the *PrTotal*Hi_M/B* and *PrTotal*Lo_M/B* coefficients are both negative and significant (p-values < 0.01 and = 0.02, respectively), as predicted. The *PrTotal*Hi_M/B* coefficient is almost twice as large (in absolute magnitude) as the *PrTotal*Lo_M/B* coefficient (−12.62 vs. −6.70) and an F-test shows that they are significantly different from each other at better than the 1% level. Collectively, these results indicate that disclosure quality is more strongly associated with a lower frequency of private information events in industry-years with high market-to-book ratios, and suggests that it is this association that drives the stronger association between disclosure quality and information asymmetry for these firms.

5.6 Alternative approach to modeling endogeneity

The results discussed above are based on a two-stage approach that uses the fitted probability a firm's disclosure score is above its industry-year median as an instrumental variable for the disclosure score. One disadvantage of this approach is that it only utilizes a small amount of the information contained in the disclosure scores. As an alternative approach, we use three stage least squares (3SLS) regressions that more fully utilize the information contained in the disclosure scores while still accounting for the endogeneity between disclosure quality and information asymmetry (Maddala, 1983).

Consistent with the prior studies, we standardize the AIMR scores by subtracting the industry-year mean and, in addition, divide by the industry-year standard deviation. We rely on the same disclosure quality and information asymmetry models as before (Eqs. 2, 3). Untabulated Hausman (1978) tests reject the null hypothesis of ''no simultaneity'' at the 0.01 level for all the models.[24] In the information asymmetry equation, the coefficient on *Total* is negative (−1.84) and highly significant (t-statistic = −2.9). The coefficient indicates that a one standard deviation increase in the *Total* is associated with a 1.84% percentage point decline in *PIN*. The results for the PIN parameters are also consistent with our expectations and each of the coefficients is significant at the 3% level or better (one-tailed tests). These results are similar to the corresponding results in Tables 3 and 4, except that we also find that the association between disclosure quality and v is significantly negative (p-value = 0.02) in the 3SLS specification whereas it is not significantly different from zero in Table 4. One explanation for this difference in results is that the 3SLS approach incorporates more information about disclosure quality, and hence, generates more powerful tests. One could then infer that higher disclosure quality also reduces information asymmetry by lowering the relative trading intensity of informed trading. However, an alternative explanation is that the various

[24] This supports our decision to treat the relation between them as endogenous. Untabulated results indicate that the *PIN* coefficient is significantly positive in the disclosure quality regression. Accordingly, we infer that managers take the level of information asymmetry into account when they make their disclosure quality choices.

methodological and measurement issues associated with 3SLS are leading to spurious inferences. See Wooldridge (2002) for a discussion of the benefits and potential limitations of 3SLS.

6 Applicability to the post-sample period

The AIMR discontinued its disclosure quality evaluations after 1996. Since that time, there have been numerous changes in the disclosure legal environment (e.g., Private Securities Litigation Reform Act, Regulation FD, and the Sarbanes-Oxley reforms), and disclosure practices, such as the dramatic increase in conference calls and management earnings forecasts (Brown, Hillegeist, & Lo, 2003). These changes call into question the generalizability of our results to the post-sample period. These questions are particularly important with respect to the changes in firms' investor relations activities as many of the IR disclosures were made selectively to analysts, and as such, are no longer allowed under Regulation FD. However, we believe that this suggests that the negative association between IR disclosure quality and information asymmetry is even stronger post-Regulation FD when all voluntary disclosures are (supposed to be) made simultaneously to the entire market. Consistent with this belief, a 2002 working paper version of Brown et al. (2004) reports that both in their pooled sample and in each of the 12 sample quarters, the association between information asymmetry and the number of open conference calls is more negative (and with higher statistical significance) than the association between information asymmetry and the number of closed conference calls (limited to analysts and large institutional investors).[25] Thus, widely disseminated disclosures appear to be more effective at reducing information asymmetry than more selective types of disclosures, such as those captured by the IR score. Therefore, we expect the same types of disclosures to be more strongly associated with information asymmetry during the post-Regulation FD period as compared to our pre-Regulation FD sample period.

However, the results in Brown et al. (2004) on the association between conference calls and information asymmetry are not entirely consistent with our results. Specifically, while they find a negative association between conference call frequency and both PINs and $\ln(\mu/\varepsilon)$, they find an unexpectedly positive association between conference calls and α, the probability of a private information event. This later association contrasts sharply with the negative coefficients on *Total* and *IR* in the α regressions documented above. One possible explanation for the differences between the two sets of results could be due to differences in PIN estimation. While this paper relies on the Venter and de Jongh (2004) extension of the EKO model, Brown et al. (2004) employ the basic EKO model. Accordingly, we replicate the analyses in Brown et al. (2004) but use the extended EKO model to estimate PIN and the PIN parameters. Consistent with Easley et al. (2002), we find no time trends in the average values of PIN within and across the two samples.

[25] They also find that the negative association between the number of management forecasts and information asymmetry is significantly stronger in the post-Regulation FD period.

In untabulated results, we find that the negative association between conference call frequency and information asymmetry is robust to the PIN model employed. In addition, the results for when $\ln(\mu/\varepsilon)$ is the dependent variable do not vary materially depending on which model is used to estimate the PIN parameters. However, the results for α are quite different. Brown et al. (2004) (Table 3B) report that in a pooled regression where α is the dependent variable, the *Calls* coefficient is positive and highly significant (t-statistic = 5.70). In contrast, the *Calls* coefficient is negative and marginally significant (t-statistic = -1.65) when α is estimated using the extended EKO model. These analyses suggest that intervening changes in the disclosure environment, which have generally broadened access to information, are unlikely to have invalidated the associations documented here.

7 Summary and conclusions

This study examines how disclosure quality is related to the level of information asymmetry. Our information asymmetry measure is based on an extended version of the EKO microstructure model and we use analysts' evaluations of disclosure quality as our proxy for disclosure quality. Our cross-sectional analyses take into account the potential endogeneity between disclosure quality and information asymmetry using a two-stage, probit-based methodology; we obtain similar, but slightly stronger, results in an alternative 3SLS specification.

Our main results are as follows: we find that the overall quality of a firm's disclosures is negatively associated with the average level of information asymmetry. Our analyses indicate this relation is primarily caused by a negative association between disclosure quality and the frequency of private information events. This finding indirectly suggests that disclosure quality reduces the incentives to search for private information, which in turn, results in fewer private information events. As such, our evidence suggests that high quality disclosures crowd out or dampen the incentives to engage in costly private information search activities, consistent with Diamond (1985) and Verrecchia (1982). We leave a direct examination of this conclusion to future research. In addition, we find no evidence of a significant association between disclosure quality and the relative amount of trading by privately informed investors. While we find a positive association between disclosure quality and uninformed trading, this association is offset by a positive association between disclosure quality and the level of informed trading.

We conduct two additional investigations to gain additional insights into the relation between disclosure quality and information asymmetry. The first examines whether three different types of disclosure quality have the same relation with information asymmetry as the aggregate measure of disclosure quality does. While we find that the quality of the annual report and investor relation activities are negatively associated with the level of information asymmetry, there is a surprisingly positive association between information asymmetry and the quality of the quarterly reports. Together, our findings indicate that the effects of disclosure quality are unlikely to be the same across all firms or, for the same firm, across different types of disclosure quality. The second investigation examines whether the relation is stronger in settings characterized by high levels of firm-investor asymmetry where public disclosures may be especially effective in reducing information asymmetry among investors. Consistent with our expectations, we find that the negative association between disclosure quality and information asymmetry is significantly stronger in industry-years with above median market-to-book ratios.

Our results should be interpreted in light of the limitations of our empirical methodology. Our analysis only allows for the relation between disclosure quality and the information asymmetry variables to be endogenous. However, other variables, such as *Analysts* and *InstOwn*, may also be endogenous. To the extent that this is true, we would have to model each endogenous variable in a simultaneous equations framework, necessitating the difficult task of finding an exogenous variable that uniquely identifies each such equation. Such a task is beyond the scope of this paper.

Acknowledgements This paper has benefited from the comments and suggestions of Eli Bartov, Sudipta Basu, George Benston, Tarun Chordia, Yonca Ertimur, Paul Fischer, Simon Gervais, Wayne Guay, Frank Heflin, Ole-Kristian Hope, Ravi Jagannathan, Stephen Monahan, Joseph Paperman, Gideon Sarr, Yong-Chul Shin, Sri Sridhar, Beverly Walther, Greg Waymire, and seminar participants at the University of Chicago, Emory University, Georgia State University, the University of Illinois at Chicago, the University of Michigan, the University of Minnesota, New York University, Northwestern University, and conference participants at the 4th Winter Accounting Conference (University of Utah) and the 2006 Review of Accounting Studies Conference (INSEAD) and anonymous referees. The second author gratefully acknowledges the financial support of the Accounting Research Center at Northwestern University. We thank Mark Finn for his efforts on an earlier version of this paper; Christine Botosan, Russell Lundholm, Marlene Plumlee, and Mark Soszek for supplying the AIMR scores; IBES for making available the analyst forecast data; and Hennie Venter for assistance in implementing the extension of the EKO model.

Appendix: Venter and de Jongh (2004) Extension of EKO model

The extended EKO model allows for the daily level of trading intensity to vary with a daily trading intensity factor, W_t. The distribution of buys (B) and sells (S) on day t is given by

$(B_t, S_t) \mid$ no-news, $W_t \sim$ Independent Bivariate Poisson $(\varepsilon W_t, \varepsilon W_t)$
$(B_t, S_t) \mid$ bad-news, $W_t \sim$ Independent Bivariate Poisson $(\varepsilon W_t, \varepsilon(1 + v)W_t)$
$(B_t, S_t) \mid$ good-news, $W_t \sim$ Independent Bivariate Poisson $(\varepsilon(1 + v)W_t, \varepsilon W_t)$.

The likelihood function induced by the model for a trading day, conditional on the Poisson trading intensities λ_{Bt} and λ_{St} for buys and sells, respectively, is given by:

$$L_t(B_t, S_t|\lambda_{Bt}, \lambda_{St}) = f_{POISS}(B_t, S_t|\lambda_{Bt}, \lambda_{St}) = \frac{(\lambda_{Bt})^{B_t}}{B_t!} * \frac{(\lambda_{St})^{S_t}}{S_t!} * e^{-\lambda_{Bt}-\lambda_{St}}. \quad (A1)$$

The overall likelihood function is a ‘‘mixture’’ model where the weights on the three components (no news, bad news, and good news) reflect the probabilities of their occurrence in the data. Denote the trading intensity associated with a no-news day (uninformed traders only) by $\lambda_{Nt} = \varepsilon W_t$ and the joint informed and uninformed trading intensity by $\lambda_{It} = \varepsilon(1 + v)W_t$. Thus:

$$
\begin{aligned}
&L_t(B_t, S_t|\lambda_{Nt}, \lambda_{It}) = L_t(B_t, S_t|\varepsilon, \mu, W_t) \\
&= (1-\alpha)f_{POISS}(B_t, S_t|\lambda_{Nt}, \lambda_{Nt}) + \alpha\delta f_{POISS}(B_t, S_t|\lambda_{Nt}, \lambda_{It}) \\
&\quad + \alpha(1-\delta)f_{POISS}(B_t, S_t|\lambda_{It}, \lambda_{Nt}) \\
&= (1-\alpha)\frac{\lambda_{Nt}^{B_t}}{B_t!} * \frac{\lambda_{Nt}^{S_t}}{S_t!} * e^{(-2\lambda_{Nt})} + \alpha\delta\frac{\lambda_{Nt}^{B_t}}{B_t!} * \frac{\lambda_{It}^{S_t}}{S_t!} * e^{(-\lambda_{Nt}-\lambda_{It})} \\
&\quad + \alpha(1-\delta)\frac{\lambda_{It}^{B_t}}{B_t!} * \frac{\lambda_{Nt}^{S_t}}{S_t!} * e^{(-\lambda_{It}-\lambda_{Nt})}. \qquad \text{(A2)}
\end{aligned}
$$

The random variable W is assumed to have a unit inverse Gaussian distribution with parameter $\psi > 0$. The density function of W is given by

$$
f_{UIG}(w;\psi) = \frac{\psi\exp(\psi^2)}{\sqrt{2\pi}} w^{-\frac{3}{2}}\exp\left(-\frac{1}{2}\psi^2(w^{-1}+w)\right), \quad w>0. \qquad \text{(A3)}
$$

The expected value of this distribution is equal to one and the variance is equal to $(1/\psi^2)$. Thus, as $\psi \to \infty$, the variance in daily trading intensities induced by general market conditions goes to zero and the extended model reduces to the basic EKO model.

The distributional assumption for W implies that the joint distribution of B_t and S_t is given by a multivariate Poisson inverse Gaussian distribution (Stein, Zucchini, & Juritz, 1987). If λ_1 (λ_2) is the base level of trading intensity for buys (sells) on a particular day (i.e., $\lambda_{Bt} = W_t\lambda_1$ and $\lambda_{St} = W_t\lambda_2$), then the likelihood function for observing the mixed Poisson distribution of B_t buys and S_t sells is:

$$
\begin{aligned}
f_{PIG} &= f_{PIG}(B_t, S_t|\lambda_1, \lambda_2, \psi) \\
&= \frac{(\lambda_1)^{B_t}}{B_t!} * \frac{(\lambda_2)^{S_t}}{S_t!} * \left[\frac{\psi^2}{\psi^2 + 2(\lambda_1+\lambda_2)}\right]^{\frac{B_t+S_t}{2}} * e(\psi^2 - \psi\sqrt{(\psi^2 + 2(\lambda_1+\lambda_2))} \\
&\quad * \hat{K}_{\left(B_t+S_t-\frac{1}{2}\right)}\left(\psi\sqrt{(\psi^2+2(\lambda_1+\lambda_2))}\right) \qquad \text{(A4)}
\end{aligned}
$$

where $\hat{K}_n(z) = K_n(z)/K_{0.5}(z)$ and $K_n(z)$ is the modified Bessel function of the second kind. Then, the expectation of B_t is given by $\mathrm{E}[B_t] = \mathrm{E}[B_t|W_t] = \mathrm{E}[\lambda_1 W_t] = \lambda_1$ and $\mathrm{Var}(B_t) = \lambda_1 + (\lambda_1/\psi^2)$; similarly for S_t. The covariance of B_t and S_t is given by $\mathrm{Cov}(B_t, S_t) = (\lambda_1\lambda_2)/\psi^2$. Therefore, the expected values of B_t and S_t are given by λ_1 and λ_2—as in the basic EKO model. However, in the extended model, if $\psi \neq \infty$, then the dispersions of B_t and S_t are higher than those in the EKO model and the daily values of buys and sells are positively correlated. Therefore, the full likelihood function is given by

$$
\begin{aligned}
L_t(B_t, S_t|\alpha, \delta, \psi, \varepsilon, v) = {} & (1-\alpha)f_{PIG}(B_t, S_t|\varepsilon, \varepsilon, \psi) \\
& + \alpha\delta f_{PIG}(B_t, S_t|\varepsilon, \varepsilon(1+v), \psi) \\
& + \alpha(1-\delta)f_{PIG}(B_t, S_t|\varepsilon(1+v), \varepsilon, \psi). \qquad \text{(A5)}
\end{aligned}
$$

References

Aboody, D., & Lev, B. (2000). Information asymmetry, R&D, and insider gains. *Journal of Finance, 55*, 2747–2766.

Ajinkya, B. B., & Gift, M. J. (1984). Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research, 22*, 425–444.

Atiase, R. K. (1985). Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research, 23*, 21–36.

Ayers, B. C., & Freeman, R. N. (2001). Evidence that price leads of earnings increase with analyst following and institutional ownership. University of Georgia working paper.

Barclay, M. J., & Warner, J. B. (1993). Stealth trading and volatility: Which trades move prices. *Journal of Financial Economics, 34*, 281–305.

Barth, M. E., Kasznik, R., & McNichols, M. F. (2001). Analyst coverage and intangible assets. *Journal of Accounting Research, 39*, 1–34.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: John Wiley.

Bollen, N. P. B., & Busse, J. (2005). Short-term persistence in mutual fund performance. *Review of Financial Studies, 18*, 569–597.

Boot, A. W. A., & Thakor, A. V. (1993). Security design. *Journal of Finance, 48*, 1349–1378.

Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *Accounting Review, 72*, 323–349.

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

Botosan, C. A., & Plumlee, M. A. (2004). Are information attributes priced? University of Utah working paper.

Brown, S., Hillegeist, S. A., & Lo, K. (2003). Regulation FD and voluntary disclosure practices. Northwestern University working paper.

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

Brown, S., Hillegeist, S. A., & Lo, K. (2005) Management forecasts and litigation risk. INSEAD working paper.

Brown, S., Hillegeist, S. A., Lo, K. (2006) The effect of meeting or missing earnings expectations of information asymmetry. University of British Columbia working paper.

Bushee, B. J. (1998). The influence of institutional investors on myopic R&D investment behavior. *Accounting Review, 73*, 305–333.

Bushee, B. J., Matsumoto, D. A., & Miller, G. S. (2003). Open versus closed conference calls: The determinants and effects of broadening access to disclosure. *Journal of Accounting & Economics, 34*, 149–180.

Bushee, B., & Noe, C. (2000). Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research, 38*, 171–202.

Callahan, C. M., Lee, C. M. C., & Yohn, T. L. (1997). Accounting information and bid-ask spreads. *Accounting Horizons, 11*, 50–60.

Chakravarty, S. (2001). Stealth-trading: Which traders' trades move stock prices. *Journal of Financial Economics, 61*, 289–307.

Cohen, D. (2003). Quality of financial reporting choice: Determinants and economic consequences. Northwestern University working paper.

Coller, M., & Yohn, T. L. (1997). Management forecasts and information asymmetry: An examination of bid-ask spreads. *Journal of Accounting Research, 35*, 181–191.

Diamond, D. W. (1985). Optimal release of information by firms. *Journal of Finance, 40*, 828–862.

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *Journal of Finance, 46*, 1325–1360.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is information risk a determinant of asset returns? *Journal of Finance, 57*, 2185–2221.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2004). Factoring information into returns. Cornell University working paper.

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

Easley, D., O'Hara, M., & Paperman, J. B. (1998). Financial analysts and information-based trade. *Journal of Financial Markets, 1*, 175–201.

Fama, E. F., & Laffer, A. B. (1971). Information and capital markets. *Journal of Business, 44*, 289–298.

FASC. (1998). Criteria for assessing the quality of an accounting standard. *Accounting Horizons, 12*, 161–162.

FASB. (2001). Improving business reporting: Insights into enhancing voluntary disclosures. Financial Accounting Standards Board's Business Reporting Research Project. Norwalk, CT.

Fishman, M. J., & Hagerty, K. M. (1989). Disclosure decisions by firms and the competition for price efficiency. *Journal of Finance, 44*, 633–646.

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

Frankel, R., McNichols, M., & Wilson, G. P. (1995). Discretionary disclosure and external financing. *Accounting Review, 70*, 135–150.

Gelb, D. S., & Zarowin, P. (2002). Corporate disclosure policy and the informativeness of stock prices. *Review of Accounting Studies, 7*, 33–52.

Hakansson, N. (1977). Interim disclosure and public forecasts: An economic analysis and framework for choice. *Accounting Review, 52*, 396–416.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*, 1251–1271.

Healy, P., Hutton, A., & Palepu, K. (1999). Stock performance and intermediation changes surrounding sustained increases in disclosures. *Contemporary Accounting Research, 16*, 485–520.

Heflin, F., Shaw, K. W., & Wild, J. J. (2005). Disclosure policy and market liquidity: Impact of depth quotes and order sizes. *Contemporary Accounting Research, 22*, 829–866.

Hillary, G. (2006). Organized labor and information asymmetry in the financial markets. *Review of Accounting Studies, 11*, 525–548.

Jiambalvo, J., Rajgopal, S., & Venkatachalam, M. (2002). Institutional ownership and the extent to which stock prices reflect future earnings. *Contemporary Accounting Research, 19*, 117–145.

Jiang, G., Lee, C. M. C., & Zhang, Y. (2005). Information uncertainty and expected returns. *Review of Accounting Studies, 10*, 185–222.

Jones, C. M., Kaul, G., & Lipson, M. L. (1994). Transactions, volume, and volatility. *Review of Financial Studies, 7*, 631–651.

Ke, B., Ramalingegowda, S., & Yu, Y. (2006). The effect of investment horizon on institutional investors' incentive to acquire private information on long-term earnings. Pennsylvania State University working paper.

Kim, O., & Verrecchia, R. E. (1991). Market reaction to anticipated announcements. *Journal of Financial Economics, 30*, 273–309.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica, 53*, 1315–1335.

Lang, M. H., & Lundholm, R. J. (1993) Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research, 31*, 246–271.

Lang, M. H., & Lundholm, R. J. (1996). Corporate disclosure policy and analyst behavior. *Accounting Review, 71*, 467–492.

Lee, C. M. C., Mucklow, B., & Ready, M. (1994). Spreads, depths, and the impact of earnings information: An intraday analysis. *Review of Financial Studies, 6*, 345–374.

Lee, C. M. C., & Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance, 46*, 733–747.

Leuz, C., & Verrecchia, R. (2000). The economic consequences of increased disclosure. *Journal of Accounting Research, 38*, 91–124.

Levitt, A. (1998). The importance of high quality accounting standards. *Accounting Horizons, 12*, 79–82.

Lundholm, R. J., & Myers, L. (2002). Bringing the future forward: The effect of disclosure on the returns-earnings relation. *Journal of Accounting Research, 40*, 809–839.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. New York, NY: Cambridge University Press.

Madhavan, A., Richardson, M., & Roomans, M. (1997). Why do securities prices change? A transaction-level analysis of NYSE stocks. *Review of Financial Studies, 10*, 1035–1064.

Malkiel, B., & Radisich, A. (2001). The growth of index funds and the pricing of equity securities. *Journal of Portfolio Management, 27*(2), 8–21.

Marquardt, C. A., & Wiedman, C. I. (1998). Voluntary disclosure, information asymmetry, and insider selling through secondary equity offerings. *Contemporary Accounting Research, 15*, 505–537.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance, 42*, 483–510.

O'Hara, M. (1995). *Market microstructure theory*. Malden, MA: Blackwell Publishers, Ltd.

Piotroski, J., & Roulstone, D. (2005). The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *Accounting Review, 79*, 1119–1151.

Sengupta, P. (1998). Corporate disclosure quality and the cost of debt. *Accounting Review, 73*, 459–474.

Stein, G., Zucchini, W., & Juritz, M. (1987). Parameter estimation for the Sichel distribution and its multivariate extension. *Journal of the American Statistical Association, 82*, 938–944.

Venter, J. H., & de Jongh, D. C. J. (2004). Extending the EKOP model to estimate the probability of informed trading. *Proceedings of first african finance conference*.

Verrecchia, R. E. (1982). The use of mathematical models in financial accounting. *Journal of Accounting Research, 20*, 1–46.

Waymire, G. (1985). Earnings volatility and voluntary management forecast disclosure. *Journal of Accounting Research, 23*, 268–295.

Welker, M. (1995). Disclosure policy, information asymmetry, and liquidity in equity markets. *Contemporary Accounting Research, 11*, 801–827.

Wooldridge, J. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Yan, X., & Zhang, Z. (2006). Institutional investors and equity returns: Are short-term institutions better informed? *Review of Financial Studies*, forthcoming.

Zhang, G. (2001). Private information production, public disclosure, and the cost of capital: Theory and implications. *Contemporary Accounting Research, 18*, 363–384.