Citation:

Luong, H, Nguyen, L and Yin, X 2015, 'When is a firm's information asymmetry priced? The role of institutional investors', International Review of Finance, vol. 15, no. 1, pp. 55-88.

Link to Published Version:

https://dx.doi.org/10.1111/irfi.12038

When Is A Firm's Information Asymmetry Priced?
The Role of Institutional Investors*

Hoang Luong Luong

Huong Giang (Lily) Nguyen

Xiangkang Yin

* Hoang Luong Luong is from the University of New South Wales, and Huong Giang (Lily) Nguyen and Xiangkang Yin are from La Trobe University. Corresponding author: Lily Nguyen, Tel: + 61 3 9479 3971, Email: lily.nguyen@latrobe.edu.au. We gratefully acknowledge the valuable comments from Jerry T. Parwada, Harald Scheule, Mark Humphery-Jenner, Darren Henry, Angela Low, Greg Jamieson, and the participants at the European FMA Annual Conference 2012 in Istanbul, FIRN Annual Conference 2011 in Queensland, Finance and Corporate Governance Conference 2012 in Melbourne, and Department Finance Seminar 2012 at La Trobe University.

Abstract

This study re-examines the competing claims that *PIN* [Easley, Hvidkjaer, and O'Hara, 2002, Is information risk a determinant of asset returns? Journal of Finance 57, 2185-2221] is priced in the cross-section of stock returns while *AdjPIN* [Duarte and Young, 2009, Why is *PIN* priced? Journal of Financial Economics 91, 119-138], the component of *PIN* related to information asymmetry, is not. We find that behind these seemingly contradicting conclusions is the role of institutional investors, and the pricing of *PIN* and *AdjPIN* depends on institutional ownership. Only for those stocks with low institutional ownership are both *PIN* and *AdjPIN* priced. Our findings imply that investors require compensation for information risk only from stocks with low institutional ownership.

JEL Classification Number*:* G3, G14, G32

Keywords: Institutional Ownership; Informed Trading; Information Asymmetry; Abnormal Returns

1 Introduction

This study examines the pricing of information asymmetry conditional on the equity ownership of institutional investors. Our study is motivated primarily by an influential set of papers by Easley, Kiefer, O'Hara and Paperman (1996), Easley, Hvidkjaer and O'Hara (2002), and Easley and O'Hara (2004), all of whom argue that information risk in the form of an asymmetric information distribution among traders of a risky asset is priced. Their key measure of information asymmetry is the probability of informed trading (*PIN*) and a number of studies have documented empirical evidence supporting the positive relation between expected returns and *PIN*.[1]

However, a small but growing empirical literature has casted doubt on whether PIN can capture information risk arising from information asymmetry.[2] Of particular note is Duarte and Young (2009), who examines whether *PIN* is priced because of information asymmetry or because of liquidity effects that are unrelated to information asymmetry. By decomposing *PIN* into two components, one related to information asymmetry and the other to illiquidity, they show that while the component related to illiquidity is priced, the component related to information asymmetry, which they call the "adjusted *PIN*" (*AdjPIN*), is not priced. In response to Duarte and Young (2009) as well as other critiques, however, Easley, Hvidkjaer, and O'Hara (2010) reconfirm that the *PIN* factor remains a significant determinant of stock returns even after

---

[1] For example, Easley, Hvidkjaer and O'Hara (2002), Easley, Hvidkjaer and O'Hara (2010), and Li, Wang, Wu and He (2009).

[2] Using a sample of mergers and acquisitions in Euronext Paris, Aktas, de Bodt, Declerck and Van Oppens (2007) find that the behavior of PIN is inconsistent with other evidence of information leakages during the pre-event period. Mohanram and Rajgopal (2009) find that the relation between PIN and stock returns is not significant when alternative model specifications and time periods are used.

controlling for the liquidity factor of Pastor and Stambaugh (2003) and the illiquidity factor of Amihud (2002).

In this paper, we also examine the relation between expected return and *PIN* or *AdjPIN* to see whether information risk that arises from information asymmetry is priced. Our innovation is that we predict and find the relation between information asymmetry and expected returns *conditional* on the equity shareholdings of institutional investors. We find that *PIN* and *AdjPIN* are priced only among the stocks that have low institutional ownership levels; this pricing effect disappears in stocks with high levels of institutional ownership. Our findings suggest that the question at issue is not as much about whether *PIN* and *AdjPIN* adequately capture information risk that originates from information asymmetry, but more about under which circumstances information asymmetry matters to asset pricing. The market appears to be less concerned about compensation for information risk in the event of high institutional ownership levels, that is, investors require compensation for information risk only among the stocks with low levels of institutional ownership.

Our focus on the roles of institutional investors in the pricing of information asymmetry is for at least two reasons. First, one potential role for large institutional investors to play is to provide a credible mechanism for transmitting information to the financial markets. Chidambaran and John (1998) find that large institutional investors convey private information obtained from management to other shareholders. Because institutional investors are likely to have informational superiority owning to the economy of scale in information acquisition and processing, as well as an easier access to and control of management, this suggests that the market may interpret the presence of an institutional shareholder in a firm as a credible signal regarding the firm's performance and future prospects.[3] Second, the presence of institutional investors in a firm may signal to the market that the severity of information asymmetry surrounding a firm is likely to be ameliorated by the monitoring activities of institutional investors. For instance, institutional investors can pressure a firm to disclose information in a timely manner,[4] which make stock price more informative about its fundamental value. Alternatively, institutional investors can threaten to sell their holdings upon receiving negative information about a firm,[5] which also causes stock price to reflect the firm's fundamental value more accurately. Therefore, we argue that the pricing of information asymmetry as measured by *PIN* or *AdjPIN* should depend on the presence of institutional investors. Our results confirm this prediction and imply that uninformed investors perceive institutional ownership as a positive, credible signal for mitigating the adverse effects of asymmetric information; hence, they require compensation for information risk only from stocks with low institutional ownership levels, that is, the pricing of information asymmetry matters only for firms with low institutional ownership.

Before addressing the central issue of the paper, we first examine and re-confirm the relation between institutional ownership and the *PIN* and *AdjPIN* measures.[6] Specifically, we find a negative effect of institutional ownership on both *PIN* and *AdjPIN*, which holds under

[3] For stock return information, see Gompers and Metrick (2001), Bennett, Sias and Stark (2003), Sias, Starks and Titman (2006), Yan and Zhang (2009), and Baik, Kang and Kim (2010). For corporate events, see Ke and Petroni (2004), Bushee and Goodman (2007), and Chiang, Qian and Sherman (2010).

[4] See, for example, Shleifer and Vishny (1986), Gillan and Starks (2000), Hartzell and Starks (2003), Ferreira and Matos (2008), and Aggarwal, Erel, Ferreira and Matos (2011).

[5] See, for example, Admati and Pfleiderer (2009), Edmans (2009), and Edmans and Manso (2011).

[6] There is limited empirical evidence on the relationship between information asymmetry and institutional shareholdings. Dennis and Weston (2001) finds that institutional ownership is negatively associated with *PIN*, but their study is restricted to a short period of 1997–1998 period. By contrast, Brockman and Yan (2009) find that both inside and outside block ownership levels are positively associated with *PIN*. To the best of our knowledge, there is to date no study that examines the relationship between institutional ownership and *AdjPIN*.

different estimation methods and model specifications. The fact that we document such a relationship substantiates the role of institutional investors in a market setting under which information is asymmetrically distributed over all traders. We then examine the competing claims that *PIN* is priced (Easley et al., 1996; Easley, Hvidkjaer, and O'Hara, 2002; Easley and O'Hara, 2004, and Easley, Hvidkjaer, and O'Hara., 2010), but *AdjPIN* is not (Duarte and Young, 2009). These claims are based on the use of different methodologies, which we closely follow but using our own sample data. In the portfolio approach, we extend the results of Easley, Hvidkjaer, and O'Hara. (2010), who focus solely on their *PIN* measure, and find that not only *PIN* but also *AdjPIN* are priced. The abnormal return of a large-*PIN* (large-*AdjPIN*) portfolio is significantly greater than that of a small-*PIN* (small-*AdjPIN*) portfolio; besides, the difference in abnormal returns between these two portfolios is statistically significant. On the other hand, in the Fama and Macbeth (1973) regressions we re-confirm the validity of Duarte and Young's (2009) claim that only *PIN* and not *AdjPIN* is priced.

To resolve these seemingly contradicting findings, we re-examine expected returns of our sample stocks in a setting where information asymmetry is most likely in evidence, together with a circumstance where information asymmetry is likely to exert the greatest effect on the expected returns. Given the significant monitoring role played by institutional investors, we introduce the level of institutional equity stock holdings of a firm into both the portfolio approach and the Fama and Macbeth (1973) regression framework. We find that both *PIN* and *AdjPIN* are priced only among those sample stocks that have low levels of institutional ownership; for stocks with high institutional ownership levels, the pricing effects of *PIN* and *AdjPIN* disappear. We make use of both the portfolio approach and the Fama-Macbeth regression framework to ensure that no fundamental difference originates from the choice of methodologies. In the portfolio approach, we double-sort stocks by institutional ownership into quintile portfolios and then by *PIN* (*AdjPIN*) into tercile portfolios. We form five hedge portfolios with long positions in the large-*PIN* (large-*AdjPIN*) tercile and short positions in the small-*PIN* (small-*AdjPIN*) tercile, and examine the abnormal returns of these hedge portfolios. We find that only the hedge portfolios with the lowest or second lowest levels of institutional ownership have a significantly positive abnormal return, while all the other hedge portfolios have insignificant alphas. These results still hold whether we use the 4-factor model of Carhart (1997) or whether we augment this model with the liquidity factor of Pastor and Stambaugh (2003). In the Fama-Macbeth framework, we create three dummy variables to index the institutional ownership groups to which a stock belongs. Again we find that only for stocks in the low or medium group of institutional shareholdings are *PIN* and *AdjPIN* priced, whereas for stocks in the high institution ownership group, *PIN* and *AdjPIN* are not priced. Taken together, these results suggest that the pricing of information asymmetry is subject to institutional ownership levels.

One study most closely related to ours is Armstrong, Core, Taylor and Verrecchia (2011), who examine when information asymmetry affects the cost of equity capital. Using the number of investors in a stock as a proxy for the level of competition for a firm's shares, they find that although information asymmetry is positively associated with the firm's cost of capital in excess of standard risk factors, this relationship holds only under imperfect market competition. Another closely related work is Akins, Ng, and Verdi (2011), who document similar results using the number and concentration of institutional investor ownership as proxies for competition among informed investors. This paper differs from these studies in that we focus on the governance role of institutional investors as large shareholders, while they concentrate on the degree of market competition. Further, we focus on equity ownership of institutional investors and relate it to widely used proxies for information asymmetry, i.e., *PIN* and *AdjPIN*.

Overall, our study contributes to the literature in three ways. First, we document a negative relationship between institutional ownership and *PIN* or *AdjPIN* measure for a large sample over a sufficiently long period. Second, we show that the contradictory results documented in prior literature depend on the methods of analysis that the authors have adopted. Third, we show that both *PIN* and *AdjPIN* are priced among only those stocks that have low institutional ownership levels, which suggests that institutional ownership is likely to affect the pricing of information asymmetry.

The remainder of the paper is organized as follows. Section 2 describes the sample data, variables and methodologies. Section 3 presents empirical results and analysis. Section 4 concludes.

2 Data, Variables, and Methodologies

2.1 Data

Our sample starts with all publicly traded firms on the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX) for the 1993–2004 period. This period overlaps largely with those in Duarte and Young (2009) and Easley, Hvidkjaer, and O'Hara (2010).[7] We obtain stock return, stock price, shares outstanding, and firm age from the Center for Research in Security Prices (CRSP), insider transactions and institutional holdings from the Thomson Reuters ownership database, accounting data from COMPUSTAT, and analyst

[7] This makes our analysis more relevant for the debate on the empirical controversy over the pricing effect of *PIN* and *AdjPIN*. Mohanram and Rajgopal (2009) find that the pricing effect of *PIN* is restricted to a particular period.

coverage from the Institutional Brokers Estimate System (I/B/E/S). To estimate *PIN* and *AdjPIN*, we obtain trades and quotes from the Trade and Automated Quote (TAQ) database. We exclude financial companies (SIC 6000–6999), utilities (SIC 4900–4999), American Depository Receipts (ADRs), Real Estate Investment Trusts (REITs), stocks of companies incorporated outside of the U.S., and closed-end funds. We also exclude stocks that have fewer than 60 days of quotes or trades in any sample year because it is impossible to estimate the *PIN* and *AdjPIN* models reliably for such stocks. Also excluded are trades and quotes that occur before and at the open, and at and after the close, quotes that have zero bid and/or ask prices, and trades that have zero prices. Finally, we eliminate observations with missing stock returns or accounting data. We winsorize all variables at the 1% and 99% tails to eliminate the effects of outliers. The final sample consists of 3,782 firms with 16,866 firm-year observations.

2.2 Variables

2.2.1 PIN and AdjPIN

As the most widely used proxy for information asymmetry, *PIN* originates from the theoretical model of Easley et al. (1996). To date, a large body of research has adopted *PIN* as the principal measure to examine the pricing of information asymmetry.[8] The *PIN* model of Easley et al. (1996) is based on the sequential trade models of Glosten and Milgrom (1985) and Easley and O'Hara (1987) in which orders come from either informed traders who trade for speculative purposes based on private information, or uninformed (noise) traders whose reasons

[8] See, for example, Easley and O'Hara (2004), Hughes, Liu, and Liu (2007), Easley et al. (2002), Hail and Leuz (2006), Easley, Hvidkjaer, and O'Hara (2010), Duarte, Han, Harford, and Young (2008), Duarte and Young (2009), Chen, Goldstein and Jiang (2007), Ferreira, Ferreira, and Raposo (2011), Aktas et al. (2007), Benos and Jochec (2007), among others.

for trading are exogenous. The model assumes that there is an uninformed liquidity provider who sets the bid and ask quotes by observing the flows of buy and sell orders and assessing the probability that the orders come from informed traders. The bid-ask spread compensates the liquidity provider for the possibility of trading with informed traders. At the beginning of each trading day, the arrival rate of buy (sell) orders follows the independent Poisson distribution, and the likelihood function of the Easley et al. (1996) model is as follows:

$$L(a,d,u,\varepsilon_b,\varepsilon_s|B,S) = (1-a)e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} + ade^{-(u+\varepsilon_b)}\frac{(u+\varepsilon_b)^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} + a(1-d)e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-(u+\varepsilon_s)}\frac{(u+\varepsilon_s)^S}{S!}, \quad (1)$$

where B (S) is the number of buys (sells) for a given trading day, a is the probability of a private information event occurring before the day, d and $(1-d)$ are the probabilities of good news and bad news, respectively, if the information event occurs, u is the arrival rate of buy or sell orders submitted by informed traders, and ε_b (ε_s) is the arrival rate of buy (sell) orders submitted by the uninformed traders. With the structural parameters estimated by maximizing the log-likelihood function based on (1),[9] Easley, Hvidkajaer, and O'Hara (2002) computes *PIN* as a fraction of orders that arises from informed traders relative to the overall total order flow, as follows:

$$PIN = \frac{au}{au+\varepsilon_b+\varepsilon_s} \quad (2)$$

We estimate *PIN* for a sample of all ordinary common stocks listed on the NYSE and AMEX (CRSP exchange codes 1 and 2, and share codes 10 and 11) for the years 1993–2004,

[9] We follow Easley, Hvidkjaer, and O'Hara (2010) to factor out the common term $e^{-\varepsilon_b-\varepsilon_s}(u+\varepsilon_b)^B(u+\varepsilon_s)^S/(B!S!)$ from (1), because computating the factorial and exponential of a large number of buy and sell orders will likely cause numerical overflows.

using intraday trades and quotes of stocks collected from the TAQ database. We focus on the NYSE and AMEX stocks because these exchanges possess the market microstructure that conforms most closely to the *PIN* structural model. We use the Lee and Ready (1991) algorithm to classify buy-initiated trades (buys) and sell-initiated trades (sells). Trades with a price above the midpoint of the bid-ask spread five seconds before the trades are classified as "buys" and those below the midpoint as "sells". Trades that occur at the mid-point of the bid and ask prices are classified as buyer- or seller-initiated according to a tick test, which classifies a trade as buyer-initiated (sell-initiated) if the price was above (below) that of the previous trade. If there are no quotes posted during the trading day, we use the tick test to sign any trades made during the day. Then, for each day, we aggregate the total number of buys and sells for each stock.

While *PIN* is commonly adopted as a measure of information asymmetry, a small but growing stream of research has casted doubt on whether it even captures information risk that arises from information asymmetry. In particular, Duarte and Young (2009) argue that when *PIN* is decomposed into illiquidity and asymmetric information, the component related to asymmetric information, i.e., *AdjPIN*, is not priced in the cross section of stock returns. Duarte and Young (2009) design the *AdjPIN* measure to better cope with the prevalent positive contemporaneous correlation between buys and sells as actually observed in the data. In this study, we use *AdjPIN* as the second proxy for information asymmetry, which is computed based on Duarte and Young (2009), as follows:

$$AdjPIN = \frac{a(du_b + (1-d)u_s)}{a(du_b + (1-d)u_s) + (\Delta_b + \Delta_s)(a\theta' + (1-a)\theta) + \varepsilon_b + \varepsilon_s}, \quad (3)$$

where the parameters in (3) are derived from Duarte and Young's (2009) AdjPIN-model,[10] θ (θ') is the probability of a symmetric order-flow shock conditional on the absence (arrival) of private information, Δ_b (Δ_s) is the arrival rate of buys (sells) caused by symmetric order-flow shocks, and u_b (u_s) is the arrival rate of buy (sell) orders submitted by informed traders if the information event occurs. The definitions of ε_b, ε_s, a, and d are the same as in equation (1).

The *AdjPIN* measure differs from *PIN* in several aspects. First, *AdjPIN* allows for the arrival rate of informed buyers, u_b, to be different from the arrival rate of informed sellers, ε_s, which enables the model to account for the fact that buy order flow has a greater variance than sell order flow for virtually all firms in the data. The more important difference is that the *AdjPIN* model allows for a new type of arrival rates of buys (sells) in the event of symmetric order-flow shocks, Δ_b (Δ_s).

We follow Duarte and Young (2009) and estimate *AdjPIN* for each firm-year over the 1993–2004 period by setting $\theta = \theta'$. To avoid numerical overflows, we use $exp(-\lambda + Xln(\lambda) - \sum_{i=1}^{X} i)$ for the Poisson density function of the form $e^{-\lambda}\frac{\lambda^X}{X!}$. As with *PIN*, we factored out the common term in the joint probability density function.

2.2.2 Institutional Ownership

Institutional ownership (*IO*): This variable is defined as the sum of shares held by institutional investors as a fraction of the firm's total shares outstanding measured at the end of each year. Following prior work (e.g., Gompers and Metrick, 2001; Yan and Zhang, 2009), we exclude observations with *IO* greater than 100%, and set *IO* to zero if a stock is not held by any institutions. A negative and significant correlation between this variable and *PIN* or *AdjPIN* will

[10] The likelihood function of Duarte and Young's (2009) *AdjPIN* model is presented in the Appendix.

provide support for the hypothesis that institutional investors can mitigate asymmetric information on a stock.

Institutional ownership concentration (*TOP5* or *IOHHI*): Following the standard approach (cite one or two papers here if possible), we use two measures to proxy for the concentration of institutional ownership of a firm. The first measure, *TOP5*, is the ratio of the shares owned by the five largest institutional investors to total shares outstanding. The other proxy is the Herfindahl index of institutional ownership (*IOHHI*) that is calculated as the sum of squares of the proportions of the firm's shares held by institutional investors.

Change of institutional ownership (ΔIO): This variable is defined as the difference in *IO* between two consecutive years. Because prior work (e.g., Sias, Starks, and Titman, 2006; Bushee and Goodman, 2007) suggests that changes in ownership by institutional investors may proxy for institutional trading associated with informed trading, we expect to find a significant and positive relation between this variable and the information asymmetry measures.

2.2.3 Abnormal Stock Returns

We closely follow prior work (e.g., Easley, Hvidkjaer, and O'Hara, 2010; Armstrong et al., 2011) to compute abnormal returns (alpha) by estimating the four-factor model of Carhart (1997). The next subsection discusses the details of how to construct a hedge portfolio and how its abnormal return is estimated.

2.2.4 Other Control Variables

To disentangle the effect of institutional ownership on information asymmetry from other effects, we follow prior work (e.g., Piotroski and Toulstone, 2004; Ferreira and Laux, 2007) and use the following control variables:

- Firm size (*SIZE*): We use the natural logarithm of market capitalization (*MCAP*) to proxy for firm size, which is calculated as share price times total shares outstanding at the end of each year. We use CRSP cumulative adjustment factors to correct both share price and shares outstanding for stock splits and dividends.
- Firm age (*LAGE*): Firm age is defined as the natural logarithm of the number of years (*AGE*) since the first return appears in CRSP.
- Market-to-book ratio (*MTB*): This ratio is defined as the natural logarithm of market value over book value of equity at the end of each year.
- Turnover (*TURN*): This variable is defined as the natural logarithm of average monthly turnover (monthly trading volume divided by shares outstanding) for the past twelve months.
- Transaction costs (*SPREAD*): This variable is measured by the average daily quoted spread over the previous calendar year.
- Leverage (*LEV*): Leverage is defined as the ratio of long-term debt to total assets at the end of the year.
- Return on equity (*ROE*): This ratio is defined as the most recent earnings before extraordinary items during the year, divided by the book value of equity at the end of the year.
- Volatility of return on equity (*VOL*): The volatility of *ROE* is the sample variance of monthly *ROE*s over the last three years.

- Dividend dummy (*DD*): This variable equals one if a firm pays dividends during a year and zero otherwise.
- Diversification dummy (*DIVER*): This dummy variable equals one if a firm operates in multi-business segments and zero otherwise.
- *SP500*: A dummy variable for membership in the S&P 500 index.

Aside from these firm characteristics, we also include two control variables that proxy for private and public information in a stock, respectively.

- Insider trading (*INSIDE*): Insiders transmit private information to market participants through their trading activity (Piotroski and Roulstone, 2004; Chen, Goldstein, and Zhang, 2007; Sias and Whidbee, 2010). We measure insider trading as the difference between the number of shares purchased and the number of shares sold in a year by insiders, scaled by the firm's total number of shares outstanding at the end of each year. We expect to find a significant and positive relationship between information asymmetry and insider trading.
- Analyst coverage (*ALYST*): Analysts disseminate private information through their earnings forecasts, revisions, and stock recommendations. Easley, O'Hara, and Paperman (1998) find that the presence of analysts reduces the content of private information in the stock price. Piotroski and Roulstone (2004) document that analysts can increase the relative amount of market- and industry-level information in stock prices, which suggests that they are likely to reduce private information in stock prices. We use the log of the number of analysts following a firm in each year as a proxy for analyst coverage, set to

zero if there is no information on analyst coverage. We expect to find a negative and significant effect of analyst coverage on information asymmetry.

2.3 Methodology

The focus of our study is to examine how institutional ownership affects the relationship between information asymmetry and expected returns. Before addressing this question, we first examine the relationship between institutional ownership and information asymmetry by estimating the following regression:

$$INFO_{i,t} = \beta_0 + \beta_1 IO_{i,t-1} + \gamma' CONTROL_{i,t-1} + e_{i,t}, \tag{4}$$

where the dependent variable is proxied by either *PIN* or *AdjPIN*, both measured in year *t*, *CONTROL* is a vector of firm characteristics as discussed in Subsection 2.2.4, subscripts *i* and *t* index stock and year, respectively. The key variable is institutional ownership, *IO*.

To examine the pricing effect of information asymmetry, we use both the portfolio approach and the Fama-MacBeth framework. In the portfolio approach, we follow Easley, Hvidkjaer, and O'Hara (2010) to compare the abnormal return of a portfolio of large-*PIN* (large-*AdjPIN*) stocks with that of a portfolio of small-*PIN* (small-*AdjPIN*) stocks. We use this approach because of the advantage that it frees us from the assumption of linearity in the variable of interest (e.g., the sort variable). Besides, this approach collapses the cross section of returns into a single time series of observations and thus alleviates concerns over the cross-sectional dependence. At the end of each year, we sort sample stocks into terciles based on the *PIN* (*AdjPIN*) measure estimated over the year. Then for each of these portfolios, we compute both monthly equally weighted or value-weighted portfolio returns, with the value-weighted returns calculated based on the market capitalization in the previous month. Next, we form a hedge portfolio by taking a long position in the large-*PIN* (large-*AdjPIN*) portfolio and a short position in the small-*PIN* (small-*AdjPIN*) portfolio. Finally, we estimate the abnormal return of the hedge portfolio by running the following four-factor regression of Carhart (1997):

$$R_{hedge,t} = \alpha_{INFO} + \beta_1 MKTRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + e_t, \quad (5)$$

where the dependent variable is the monthly return of the hedge portfolio, *MKTRF*, *SMB*, and *HML* are the three factors of Fama and French (1993), and *UMD* is the momentum factor of Carhart (1997). Our interest is in α_{INFO}, which refers to either α_{PIN} or α_{AdjPIN} depending on whether the portfolio is sorted by *PIN* or *AdjPIN*. A positive and statistically significant α_{INFO} suggests that a trading strategy of buying large-*PIN* (large-*AdjPIN*) stocks and selling small-*PIN* (small-*AdjPIN*) stocks is profitable and thus implies that investors generally require compensation for information risk that arises from information asymmetry.

As an alternative to the portfolio approach, we follow Duarte and Yong (2009) to use the Fama and MacBeth's (1973) regression method to examine the pricing of information asymmetry, as follows:

$$R_{i,t} = \beta_0 + \beta_1 INFO_{i,t-1} + \beta_2 Beta_{i,t-1} + \beta_3 SIZE_{i,t-1} + \beta_4 BTM_{i,t-1} + e_{i,t}, \quad (6)$$

where the dependent variable is the monthly return of stock *i*. *INFO* is measured by either *PIN* or *AdjPIN*. We estimate the beta, *Beta*, of a stock using the data of the past 60 months. *SIZE* is the natural logarithm of market capitalization, and *BTM* is the natural logarithm of book to market ratio.

To address the main research question of whether and under which setting information asymmetry as captured by *PIN* or *AdjPIN* is priced in the cross section of stock returns, we extend both (5) and (6) by accommodating institutional ownership in our analysis. In the portfolio approach, we double-sort stocks by *IO* and *PIN* (*AdjPIN*). Double-sorting makes it possible to see whether there are significant differences in the abnormal returns of the hedge portfolios across different *IO*-sorted groups. Specifically, we sort stocks into quintiles based on *IO* measure at the end of the year. The stocks in each quintile portfolio are then sorted into terciles according to *PIN* (*AdjPIN*) estimated over the year. Double-sorting thus produces fifteen *IO-PIN* (*IO-AdjPIN*) portfolios shown as follows:

IO \ *PIN* (*AdjPIN*)	Small (A)	Medium (B)	Large (C)	Hedge Portfolio
Lowest (1)	(1, A)	(1, B)	(1, C)	(1, C) – (1, A)
(2)	(2, A)	(2, B)	(2, C)	(2, C) – (2, A)
(3)	(3, A)	(3, B)	(3, C)	(3, C) – (3, A)
(4)	(4, A)	(4, B)	(4, C)	(4, C) – (4, A)
Highest (5)	(5, A)	(5, B)	(5, C)	(5, C) – (5, A)

We construct five hedge portfolios by longing the large-*PIN* (large-*AdjPIN*) portfolio and shorting the small-*PIN* (small-*AdjPIN*) portfolio. We then compute the abnormal returns (α_{INFO}) of these five hedge portfolios by estimating equation (5). Checking the statistical and economic significance of α_{INFO} across the *IO* quintile groups will reveal a setting in which institutional ownership affects the pricing of information asymmetry.

We also identify this setting by extending equation (6) to include interaction terms between information asymmetry proxy *PIN* (*AdjPIN*) and a dummy variable that indexes to which IO group a stock belongs. We first sort stocks into tercile group based on *IO* measure at the end of each year, and then use three dummy variables to index the IO group to which a stock belongs. We interact these variables with an information asymmetry measure and estimate the following regression:

$$R_{i,t} = \beta_0 + \beta_1(INFO_{i,t-1} \times DL) + \beta_2(INFO_{i,t-1} \times DM) + \beta_3(INFO_{i,t-1} \times DH)$$
$$+\beta_4 DM + \beta_5 DH + \beta_6 Beta_{i,t-1} + \beta_7 SIZE_{i,t-1} + \beta_8 BTM_{i,t-1} + e_{i,t}. \quad (6a)$$

Our variables of interest include three interaction terms, $(INFO_{i,t-1} \times DL)$, $(INFO_{i,t-1} \times DM)$, and $(INFO_{i,t-1} \times DH)$, which capture the effects of information asymmetry on expected returns for firms belonging to the low-, medium- and high-*IO* groups, respectively.

3 Empirical Results

3.1 Descriptive Statistics

We compute mean cross-sectional measures of information asymmetry, institutional ownership, and other control variables for each year over the 1993–2004 period. Panel A of Table 1 reports the time-series summary statistics of these twelve cross-sectional averages. The mean estimate of *PIN* is 0.22, compared to 0.18 for *AdjPIN*, which means that about 20% of all observed trades in our sample period originates from informed investors. Both the mean and median estimates of *AdjPIN* are smaller than are those of *PIN*, which is consistent with the argument in Duarte and Young (2009) that *AdjPIN* captures only one component of *PIN*. On average, a firm in our sample has about 48% of its total shares held by institutional investors, with the top five largest institutional investors holding nearly 23%. It has an average market capitalization of USD 267 million. The average firm age is about 12 years.

Insert Table 1 here

Panel B of Table 1 presents a correlation matrix for the time-series averages of cross-sectional correlations between the information asymmetry measures and institutional ownership, as well as other control variables. *PIN* and *AdjPIN* are positively correlated, with a high significant coefficient of 0.73. Institutional ownership is negatively correlated with both *PIN* and *AdjPIN*, with a correlation coefficient comparable in magnitude (-0.52 for *PIN* and -0.48 for *AdjPIN*), suggesting that there is possibly a relationship between institutional ownership and information asymmetry. With regard to control variables, analyst coverage (*ALYST*) is negatively correlated with both *PIN* and *AdjPIN* as expected, consistent with earlier work which shows that analyst coverage reduces the amount of private information in stock prices. Insider trading (*INSIDE*) is positively correlated with *PIN* and *AdjPIN*. Both firm size and age are negatively correlated with information asymmetry measures, which is intuitive because older and larger firms tend to be associated with a lower probability of informed trading as documented in prior literature.

Table 2 shows key characteristics of the fifteen portfolios double-sorted by *IO* and the information asymmetry measures. The key characteristics include the average of *PIN*, *AdjPIN*, *IO*, *MCAP*, and monthly returns. Several remarks are in order. First, both *PIN* and *AdjPIN* decrease monotonically from the lowest- to the highest-*IO* stocks. Second, *IO* monotonically fall from small- to large-*PIN* (large-*AdjPIN*) stocks within each of the *IO* quintile groups. These results suggest that there may be an inverse relationship between institutional ownership and information asymmetry captured by *PIN* and *AdjPIN*. Firm size increases when we move from the lowest- to the highest-*IO* stocks, regardless of *PIN* (*AdjPIN*) tercile groups. Within each *IO* group, firm size decreases monotonically as we move from the small- to the large-*PIN* (large-*AdjPIN*) stocks, consistent with the conjecture that the trading of small stocks is associated with a greater probability of private information. While monthly returns apparently depend on both information asymmetry and *IO*, the difference in monthly returns between the large-*PIN* (large-*AdjPIN*) and small-*PIN* (small-*AdjPIN*) portfolios is positive and significant only for the lowest- and second lowest-*IO* quintiles.

Insert Table 2 here

3.2 Institutional Shareholding and Information Asymmetry

To examine the relation between institutional ownership and information asymmetry visually, we plot the average *IO* against the *PIN* quintile (left panel) and *AdjPIN* quintile (right) in Figure 1. Figure 1 shows that the smallest-*PIN* quintile (Q1) has a greater *IO* relative to Q2 and this relationship is monotonic across all *PIN* quintiles. A qualitatively similar result is also documented for *AdjPIN*.

Insert Figure I here

To examine the relationship between institutional ownership and information asymmetry closely in a regression framework, we estimate the regression equation (4) for 16,866 firm-year observations over the 1994–2004 sample period and present the results in Table 3. Following Petersen (2009), we adjust standard errors for clustering at firm and year levels to account for possible residual correlations among different firms in the same year and across different years for the same firm. To control for possible industry and time fixed effects, we include a set of dummies for one-digit standard industrial classification (SIC) as well as for years in all regressions.

Insert Table 3 here

As we can see in both panels, the *IO* coefficient is negative and highly significant in all specifications. Column 1 shows that the coefficient estimate of *IO* is negative and significant when we control for firm characteristics, analyst coverage, and insider trading. In columns 2 and 3, we add institutional ownership concentration (*TOP5* or *IOHHI*) and find that its coefficient estimates are positive and statistically significant, which is consistent with Heflin and Shaw (2000) and Aslan et al. (2011) that institutional ownership concentration increases the probability of informed trading. However, the coefficient estimates of *IO* remain negative and significant in both columns 2 and 3. In column 4, we control for transaction costs (*SPREAD*) and find that stocks with higher transaction costs are associated with a higher probability of informed trading, consistent with prior work (e.g., Falkenstein, 1996; Gompers and Metrick, 2001). The coefficient estimates of *IO* remain negative and significant, though. In columns 5–6, where we control for both institutional ownership concentration and transaction costs simultaneously, the coefficient estimates of *IO* remain negative and statistically significant. Collectively, these results suggest that there exists a negative relationship between institutional ownership and *PIN* or *AdjPIN*, which is robust to the use of controls for institutional ownership concentration, transaction costs, and other firm characteristics. These results support the hypothesis that the presence of institutional investors through share ownership sends a (more) credible signal that proper monitoring is in place to ameliorate a firm’s information environment, consistent with Dennis and Weston (2001) and Aslan et al. (2011), who document a negative relationship between institutional ownership and *PIN*.

With regard to other control variables, we find results that are largely consistent with the literature. Larger or older firms are negatively associated with *PIN* or *AdjPIN*, as are firms with greater stock liquidity, higher market-to-book ratios, and greater analyst coverage. Firms with higher leverage or in the S&P 500 index are positively associated with information asymmetry, as are firms with greater insider trading activities.

To check the robustness of our baseline results and address endogeneity concerns, we also use alternative estimation methods. We use the Fama-MacBeth (1973) regression method to account for possible cross-sectional correlations between firms in a given year. In addition, we use the firm fixed effects regressions to control for unobserved time-invariant sources of firm heterogeneity. Another major concern is the reverse causality between information asymmetry and *IO*, that is, institutional investors may avoid stocks with a greater probability of private information in stock prices. To address the reverse causality issue, we add a one-year lag of the dependent variable (see Chung and Zhang, 2011) to the regressions as an explanatory variable. Columns 1–3 in Table 4 present the results based on these estimation methods. The *IO* coefficient remains negative and highly significant in all specifications regardless of the use of proxies for information asymmetry. We also estimate a system of simultaneous equations in which the dependent variables are *PIN*/*AdjPIN* and *IO* (columns 4–5). We add stock prices (*PRC*) and cumulative stock returns over the past twelve months (*MOM*) to the *IO* equation (Gompers and Mettrick, 2001). The *IO* coefficient remains negative and significant in this alternative specification (column 4), while the coefficients on lagged *PIN* and lagged *AdjPIN* in the *IO* regression (column 5) are significantly negative. These results reinforce the evidence reported earlier that an inverse relationship exists between *IO* and information asymmetry.

Insert Table 4 here

Overall, this subsection provides evidence that supports a negative and significant relationship between institutional ownership and information asymmetry, which is robust to the inclusion of institutional ownership concentration, transaction costs, and other controls for firm characteristics, as well as to the use of different estimation methods.

3.3 Are *PIN* and *AdjPIN* Priced?

To examine the pricing effect of information asymmetry, we first run the regression equation (5) based on three portfolios single-sorted on the basis of *PIN* (*AdjPIN*) estimates and a hedge portfolio with a long position in the large-*PIN* (large-*AdjPIN*) stocks and a short position in the small-*PIN* (small-*AdjPIN*) stocks. Table 5 reports the coefficient estimates with robust standard errors. Panel A displays the estimates for equally weighted portfolios wile Panel B presents the estimates for value-weighted portfolios. As shown in both panels, the estimated abnormal return of each portfolio (α_{PIN} and α_{AdjPIN}) is positive and significant, increasing monotonically from the small-*PIN* (small-*AdjPIN*) to the large-*PIN* (large-*AdjPIN*) stocks. More importantly, the abnormal return of the hedge portfolio is positive and statistically significant. The estimates of α_{PIN} are 0.5% and 0.4% per month for equally weighted and value-weighted hedge portfolios, respectively, while the estimates of α_{AdjPIN} are smaller (0.4% and 0.3% per month, respectively). These results show that abnormal returns are earned on stocks with greater information asymmetry in excess of standard risk factors, implying that investors appear to require compensation for information risk that arises from information asymmetry. The results are consistent with findings in Easley, Hvidkjaer, and O'Hara (2002, 2010) that investors do require compensation for information disadvantages, but in this study, we use both *PIN* and *AdjPIN* as proxies for information asymmetry.

Insert Table 5 here

As mentioned above, α_{AdjPIN} is smaller than α_{PIN} for both equally weighted and value-weighted hedge portfolios. Using the Fama-MacBeth regression equation (6), Duarte and Young (2009) argue that *AdjPIN*, which captures the asymmetric information component of *PIN*, cannot explain the cross-sectional differences in stock returns. We estimate the regression model (6) based on our sample data and report the results in Table 6. The coefficient estimate of *PIN* is positive and significant (column 1), but the coefficient on *AdjPIN* (column 3) is insignificant, which is similar to Duarte and Young's (2009) results that information asymmetry proxied by *AdjPIN* is not priced in cross-sectional return estimations.

Insert Table 6 here

Overall, this subsection, on the one hand, presents evidence that supports and extends the findings of Easley, Hvidkjaer, and O'Hara (2002, 2010) that both *PIN* and *AdjPIN* is priced, and on the other hand, documents evidence similar to Duarte and Young (2009) that shows *AdjPIN* is not priced in the cross section of stock returns. It is then natural to ask whether there is some fundamental economic reason behind this contradiction. We provide analysis to answer this question in the next subsection.

3.4 When are *PIN* and *AdjPIN* Priced?

A main purpose of this study is to identify a circumstance under which information asymmetry is likely to exhibit the greatest effect on expected returns, coupled with a setting where information asymmetry is most likely in evidence. Given the baseline results that institutional ownership has a negative effect on information asymmetry, together with the evidence that information asymmetry is positively associated with expected return, the questions to ask are whether the pricing effect of information asymmetry is affected by institutional ownership and whether institutional ownership explains the empirical controversy that we noted in the preceding subsection.

To answer these questions, we first re-estimate equation (5) but based on five hedge portfolios double-sorted by *IO* and *PIN* (*AdjPIN*) to account for different levels of institutional ownership. Table 7 presents the results with robust standard errors, where Panel A uses equally weighted portfolios while Panel B uses value-weighted portfolios. Both panels show that the abnormal return (α_{PIN} and α_{AdjPIN}) of the hedge portfolios is decreasing in *IO*. In particular, only α_{PIN} and α_{AdjPIN} estimated from the regressions based on the lowest- and the second lowest-*IO* groups (columns 1 and 2) are significantly positive, ranging from between 0.4% to 0.9% per month. The α_{PIN} and α_{AdjPIN} estimates from other *IO* groups are all statistically insignificant and their economic significance also is much smaller (one of them is close to zero). These results show that the pricing effect of information asymmetry is restricted to low-*IO* stocks, which implies that a substitution effect may operate between institutional ownership and information asymmetry. For instance, uninformed investors, such as retail traders, demand a higher expected return on stocks that contain substantial private information as compensation for information risk that arises from their informational disadvantages; however, this adverse effect is likely to be mitigated in firms by high institutional ownership.

Insert Table 7 here

Consolidating Tables 5 and 7 suggests that the results documented for the aggregate-*IO* hedge portfolio are driven largely by a subset of stocks, i.e., stocks with low levels of institutional ownership. Although the aggregate hedge portfolio yields a significant and positive abnormal return as shown in the last column of Table 5, Table 7 shows that this result is driven by stocks with low institutional ownership levels. High-IO stocks do not contribute to abnormal returns in any meaningful sense. Hence, these results support the hypothesis that the pricing effect of information asymmetry is significant only for stocks with low levels of institutional ownership.

An interesting debate is on whether PIN is priced because of asymmetric information in stock trading or because of stock illiquidity (e.g. Easley, Hvidkjaer, and O'Hara, 2010; Duarte and Young, 2009). We augment the regression equation (5) by adding the liquidity factor of Pastor and Stambaugh (2003) to see if our main results are still valid, and present the results in Table 8. Again, the results are qualitatively similar.

Insert Table 8 here

As a robustness check, we use an alternative sorting procedure with the portfolio approach in which stocks are sorted into terciles, rather than quintiles, based on IO, and then within each IO tercile, stocks are further sorted into tercile groups based on PIN ($AdjPIN$), resulting in a total of nine portfolios. A reverse sorting, i.e., stocks are sorted first by PIN ($AdjPIN$) and then by IO, is also used. None of these procedures alters the main results qualitatively. For brevity, we do not report these results in this paper.

Because different approaches may lead to contradicting results as we mentioned in the preceding section, we use our sample data and estimate the regression equation (6a) to examine further the effect of institutional ownership on expected returns. If the results are consistent with those obtained from the portfolio approach, the coefficient estimate of $(INFO \times DL)$ should be positive and significant, whereas the coefficient estimate of $(INFO \times DH)$ should be insignificant. Columns 2 and 4 of Table 6 report the results from estimating equation (6a). As we can see in column 2, the coefficient estimate of $(PIN \times DL)$ is significant and positive, while the coefficient on $(PIN \times DH)$ is insignificant. Similar results are observed for $AdjPIN$ in column 4. These results reinforce our results reported earlier that information asymmetry as captured by *PIN* or $AdjPIN$ is priced only among stocks with low levels of institutional stock ownership.

Overall, this subsection shows that the pricing effect of information asymmetry proxied by *PIN* or *AdjPIN* is significant only among stocks with low institutional ownership levels. The effect disappears among the stocks with high levels of institutional ownership.

4 Conclusions

This paper examines the pricing effect of information asymmetry measured by the *PIN* of Easley, Hvidkjaer, and O'Hara (2002) and the *AdjPIN* of Duarte and Young (2009). In the portfolio approach, which does not distinguish stocks by institutional ownership level, we find that both *PIN* and *AdjPIN* exhibit significant pricing effects in excess of standard risk factors. In the regression approach, in which we regress returns on the information asymmetry measure over all individual stocks, we find that only *PIN*, and not *AdjPIN*, is priced in the cross section of stock returns. When institutional ownership is considered, however, this controversial result disappears. Using both approaches, we find that information asymmetry as captured by *PIN* or *AdjPIN* affects expected returns only for stocks with low levels of institutional ownership. There is no such evidence for stocks with high institutional ownership levels, though.

A natural question to ask is why the PIN and AdjPIN measures are not priced for stocks with high institutional ownership. Two possibilities emerge. First, empirical literature has well-established that institutional investors act as active monitors by directly intervening in a firm's activities (Shleifer and Vishny, 1986; Gillan and Starks, 2000; Ferreira and Mastos, 2008). This monitoring role can potentially mitigate information asymmetry in a stock. More importantly, institutional monitoring can prevent the leakage and spread of harmful private information, so that uninformed investors take institutional ownership in a firm as a credible signal that the information risk arising from information asymmetry is mitigated. Second, as a measure of information asymmetry, neither *PIN* nor *AdjPIN* distinguishes the types of private information in the trading of a stock. In a sense, while both *PIN* and *AdjPIN* measure the quantity of private signals, they do not take into account the effectiveness of such signals that reveal the fundamental value of a firm. Therefore, uninformed investors may take institutional holdings as an alternative signal for assessing information risk. Taken together, these two factors lead us to the observation that the effects of the *PIN* and *AdjPIN* measures on expected stock returns only materialize for such firms that have a small proportion of institutional investors. Further research, however, is warranted in order to determine whether these explanations are reasonably valid. This remains a topic for our future research.

References

Admati, A.R., Pfleiderer, P., 2009. The "Wall Street Walk" and Shareholder Activism: Exit as a Form of Voice. Review of Financial Studies. volume and page?

Aggarwal, R., Erel, I., Ferreira, M., Matos, P., 2011. Does Governance Travel around the World? Evidence from institutional investors. Journal of Financial Economics 100, 154-181.

Akins, B.K., Ng, J., Verdi, R.S., 2011. Investor Competition over Information and the Pricing of Information Asymmetry. The Accounting Review 87, 35-58.

Aktas, N., de Bodt, E., Declerck, F., Van Oppens, H., 2007. The PIN Anomaly around M&A Announcements. Journal of Financial Markets 10, 169-191.

Amihud, Y., 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. Journal of Financial Markets, 5, 31-56.

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

Baik, B., Kang, J.-K., Kim, J.-M., 2010. Local Institutional Investors, Information Asymmetries, and Equity Returns. Journal of Financial Economics 97, 81-106.

Bennett, J.A., Sias, R.W., Starks, L.T., 2003. Greener Pastures and the Impact of Dynamic Institutional Preferences. Review of Financial Studies 16, 1203-1238.

Benos, E., Jochec, M., 2007. Testing the PIN Variable. SSRN eLibrary. Please give the website address

Brockman, P., Yan, X., 2009. Block Ownership and Firm-specific Information. Journal of Banking and Finance 33, 308-316.

Bushee, B.J., Goodman, T.H., 2007. Which Institutional Investors Trade Based on Private Information About Earnings and Returns? Journal of Accounting Research 45, 289-321.

Carhart, M.M., 1997. On Persistence in Mutual Fund Performance. The Journal of Finance 52, 57-82.

Chen, Q., Goldstein, I., Jiang, W., 2007. Price Informativeness and Investment Sensitivity to Stock Price. Review of Financial Studies 20, 619-650.

Chiang, Y.-M., Qian, Y., Sherman, A.E., 2010. Endogenous Entry and Partial Adjustment in IPO Auctions: Are Institutional Investors Better Informed? Review of Financial Studies 23, 1200-1230.

Chidambaran, N.K., John, K., 1998. Relationship Investing: Large Shareholder Monitoring with

Managerial Cooperation. NYU Working Paper No. FIN-98-044.

Chung, K.H., Zhang, H., 2011. Corporate Governance and Institutional Ownership. Journal of Financial & Quantitative Analysis 46, 247-273.

Dennis, P.J., Weston, J.P., 2001. Who's Informed? An Analysis of Stock Ownership and Informed Trading. SSRN eLibrary. Please give the website address

Duarte, J., Han, X., Harford, J., Young, L., 2008. Information Asymmetry, Information Dissemination and the Effect of Regulation FD on the Cost of Capital. Journal of Financial Economics 87, 24-44.

Duarte, J., Young, L., 2009. Why Is PIN Priced? Journal of Financial Economics 91, 119-138.

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

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

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

Easley, D., O'Hara, M., 1987. Price, Trade Size, and Information in Securities Markets. Journal of Financial Economics 19, 69-90.

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

Easley, D., O'Hara, M., Paperman, J., 1998. Financial Analysts and Information-based Trade. Journal of Financial Markets 1, 175-201.

Edmans, A., 2009. Blockholder Trading, Market Efficiency, and Managerial Myopia. Journal of Finance 64, 2481-2513.

Edmans, A., Manso, G., 2010. Governance through Trading and Intervention: A Theory of Multiple Blockholders. Review of Financial Studies. volume and page?

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

Fama, E.F., MacBeth, J.D., 1973. Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy 81, 607-636.

Ferreira, D., Ferreira, M.A., Raposo, C.C., 2011. Board Structure and Price Informativeness. Journal of Financial Economics 99, 523-545.

Ferreira, M.A., Laux, P.A., 2007. Corporate Governance, Idiosyncratic Risk, and Information Flow. Journal of Finance 62, 951-989.

Ferreira, M.A., Matos, P., 2008. The Colors of Investors' Money: The Role of Institutional Investors around the World. Journal of Financial Economics 88, 499-533.

Gillan, S.L., Starks, L.T., 2000. Corporate Governance Proposals and Shareholder Activism: The Role of Institutional Investors. Journal of Financial Economics 57, 275-305.

Glosten, L.R., Milgrom, P.R., 1985. Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. Journal of Financial Economics 14, 71-100.

Gompers, P.A., Metrick, A., 2001. Institutional Investors and Equity Prices. Quarterly Journal of Economics 116, 229-259.

Hail, L., Leuz, C., 2006. International Differences in the Cost of Equity Capital: Do Legal Institutions and Securities Regulation Matter? Journal of Accounting Research 44, 485-531.

Hartzell, J.C., Starks, L.T., 2003. Institutional Investors and Executive Compensation. Journal of Finance 58, 2351-2374.

Hughes, J.S., Liu, J., Liu, J., 2007. Information Asymmetry, Diversification, and Cost of Capital. Accounting Review 82, 705-729.

Ke, B.I.N., Petroni, K., 2004. How Informed Are Actively Trading Institutional Investors? Evidence from Their Trading Behavior before a Break in a String of Consecutive Earnings Increases. Journal of Accounting Research 42, 895-927.

Lee, C.M.C., Ready, M.J., 1991. Inferring Trade Direction from Intraday Data. Journal of Finance 46, 733-746.

Li, H., Wang, J., Wu, C., He, Y.A.N., 2009. Are Liquidity and Information Risks Priced in the Treasury Bond Market? Journal of Finance 64, 467-503.

Mohanram, P., Rajgopal, S., 2009. Is PIN Priced Risk? Journal of Accounting and Economics 47, 226-243.

Pastor, L., Stambaugh, R.F., 2003. Liquidity Risk and Expected Stock Returns. Journal of Political Economy 111, 642-685.

Petersen, M.A., 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Review of Financial Studies 22, 435-480.

Piotroski, J.D., Roulstone, B.T., 2004. 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.

Richard W. Sias, Laura T. Starks, Sheridan Titman, 2006. Changes in Institutional Ownership and Stock Returns: Assessment and Methodology. Journal of Business 79, 2869-2910.

Shleifer, A., Vishny, R.W., 1986. Large Shareholders and Corporate Control. Journal of Political Economy 94, 461-488.

Sias, R.W., Whidbee, D.A., 2010. Insider Trades and Demand by Institutional and Individual Investors. Review of Financial Studies 23, 1544-1595.

Yan, X., Zhang, Z., 2009. Institutional Investors and Equity Returns: Are Short-term Institutions Better Informed? Review of Financial Studies 22, 893-924.

Appendix: The Likelihood of the Extended PIN Model

Duarte and Young (2009) extend the PIN model of Easley et al. (1996) to allow for the pervasive positive contemporaneous correlations between buyer- and seller-initiated order flows as actually observed in the data. They decompose *PIN* into two components: asymmetric information, which they term "adjusted" *PIN* (*AdjPIN*), and illiquidity. Thus, *AdjPIN* is *PIN* purged of all illiquidity effects unrelated to information asymmetry. Duarte and Young (2009) argue that *AdjPIN* is more accurate as a proxy for information asymmetry. The likelihood function of the *AdjPIN* model is specified as follows:

$$
\begin{aligned}
L(\vartheta|B,S) = {} & (1-a)(1-\theta)e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} \\
& + (1-a)\theta e^{-(\varepsilon_b+\Delta_b)}\frac{(\varepsilon_b+\Delta_b)^B}{B!}e^{-(\varepsilon_s+\Delta_s)}\frac{(\varepsilon_s+\Delta_s)^S}{S!} \\
& + a(1-\theta')(1-d)e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-(u_s+\varepsilon_s)}\frac{(u_s+\varepsilon_s)^S}{S!} \\
& + a\theta'(1-d)e^{-(\varepsilon_b+\Delta_b)}\frac{(\varepsilon_b+\Delta_b)^B}{B!}e^{-(u_s+\varepsilon_s+\Delta_s)}\frac{(u_s+\varepsilon_s+\Delta_s)^S}{S!} \\
& + a(1-\theta')de^{-(u_b+\varepsilon_b)}\frac{(u_b+\varepsilon_b)^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} \\
& + a\theta de^{-(u_b+\varepsilon_b+\Delta_b)}\frac{(u_b+\varepsilon_b+\Delta_b)^B}{B!}e^{-(\varepsilon_s+\Delta_s)}\frac{(\varepsilon_s+\Delta_s)^S}{S!},
\end{aligned}
$$

where B (S) is the number of buys (sells) for a given day, θ (θ') is the probability of a symmetric order-flow shock conditional on the absence (arrival) of private information, Δ_b (Δ_s) is the arrival rate of buys (sells) in the event of symmetric order-flow shock, ε_b (ε_s) is the arrival rate of buy (sell) orders submitted by uninformed traders, a is the probability of an information event occurring during a trading day, d is the probability of good news, $(1-d)$ is the probability of bad news, u_b (u_s) is the arrival rate of buy (sell) orders submitted by informed traders if the information event occurs, and $\vartheta = (a, u_b, u_s, \varepsilon_b, \varepsilon_s, d, \theta, \theta', \Delta_b, \Delta_s)$ is a vector of parameters.

Table 1: Summary Statistics

This table reports the summary statistics of all variables for the sample period from 1993 to 2004. Panel A presents the time-series mean, median, standard deviation, and 25th and 75th quartiles. Panel B reports the correlation matrix. All variables are defined in Subsection 2.2.

Panel A: Descriptive Statistics						
Variables	Mean	SD	P25	P50	P75	Obs.
Informed Trading Measures						
PIN	0.217	0.118	0.150	0.185	0.414	16,866
AdjPIN	0.178	0.082	0.117	0.159	0.306	16,866
Institutional Ownership						
IO	0.481	0.264	0.258	0.514	0.695	16,866
TOP5	0.227	0.118	0.147	0.226	0.299	16,866
IOHHI	0.154	0.186	0.045	0.075	0.182	16,866
ΔIO	0.066	0.076	0.017	0.043	0.088	16,866
Firm Characteristics						
MCAP	266.541	0.019	55.841	271.061	1,175.001	16,866
MTB	1.354	0.819	1.138	1.449	2.107	16,866
SPREAD	0.348	0.238	0.187	0.326	0.446	16,866
TURN	0.079	0.124	0.022	0.046	0.088	16,866
AGE	12.383	1.130	6.209	12.420	12.526	16,866
ROA	0.030	0.478	0.021	0.102	0.168	16,866
VOL	0.078	0.146	0.012	0.027	0.064	16,866
LEV	0.263	0.217	0.086	0.239	0.383	16,866
DD	0.484	0.500	0.000	0.000	1.000	16,866
DIVER	0.947	0.223	1.000	1.000	1.000	16,866
SP500	0.213	0.409	0.000	0.000	0.000	16,866
ALYST	4.341	3.750	1.000	5.006	20.000	16,866
INSIDE	0.011	0.039	0.000	0.000	0.004	16,866

Panel B: Correlation Matrix

	PIN	*AdjPIN*	*IO*	*TOP5*	*IOHHI*	*ΔIO*	*SIZE*	*MTB*	*TURN*	*AGE*	*ROE*	*VOL*	*LEV*	*DD*	*DIVER*	*SP500*	*ALYST*	*INSIDE*
PIN	**1.000**																	
AdjPIN	0.732	**1.000**																
IO	-0.516	-0.483	**1.000**															
TOP5	-0.197	-0.157	0.579	**1.000**														
IOHHI	0.581	0.506	0.434	0.314	**1.000**													
ΔIO	-0.060	-0.067	0.114	0.200	-0.127	**1.000**												
SIZE	-0.283	-0.344	0.192	-0.079	-0.232	-0.077	**1.000**											
MTB	-0.326	-0.316	0.145	-0.082	-0.217	-0.017	0.295	**1.000**										
TURN	-0.381	-0.347	0.256	0.113	-0.159	0.167	0.014	0.079	**1.000**									
AGE	-0.117	-0.136	0.161	0.029	-0.164	-0.139	0.253	-0.042	-0.077	**1.000**								
ROE	-0.120	-0.108	0.231	0.136	-0.255	-0.010	0.127	0.025	-0.119	0.106	**1.000**							
VOL	0.003	-0.010	-0.187	-0.165	0.184	-0.002	0.008	0.348	0.162	-0.077	-0.356	**1.000**						
LEV	-0.001	-0.009	0.026	0.071	0.007	0.083	-0.021	0.074	0.077	-0.054	-0.076	0.235	**1.000**					
DD	-0.220	-0.211	0.264	0.094	-0.296	-0.104	0.264	0.052	-0.173	0.406	0.200	-0.173	-0.084	**1.000**				
DIVER	-0.032	-0.015	0.062	0.045	-0.093	-0.010	0.052	0.000	-0.069	0.124	0.081	-0.083	0.007	0.135	**1.000**			
SP500	-0.370	-0.395	0.335	0.017	-0.303	-0.080	0.570	0.219	0.073	0.375	0.130	-0.037	0.001	0.380	0.091	**1.000**		
ALYST	-0.383	-0.374	0.356	0.094	-0.356	-0.034	0.410	0.140	0.069	0.313	0.145	-0.088	0.019	0.339	0.087	0.526	**1.000**	
INSIDE	0.025	0.034	-0.025	-0.014	0.028	0.064	-0.060	0.039	0.060	-0.148	0.004	0.024	0.012	-0.119	-0.042	-0.101	-0.079	**1.000**

Table 2: Summary Statistics of Portfolios Double-Sorted by *IO* and *PIN* (*AdjPIN*)

This table presents key characteristics of portfolios double-sorted by *IO* and *PIN* (*AdjPIN*) for the years from 1993 through 2004. At the end of each year, stocks are sorted into quintiles based on *IO*. Within each *IO* quintile, stocks are further sorted into terciles based on *PIN* (Panel A) or *AdjPIN* (Panel B) estimated over the year. Each panel reports the mean of *PIN*s (*AdjPIN*s), market capitalizations (in USD millions), *IOs* and monthly returns.

Panel A: Portfolios Double-Sorted by *IO* and *PIN*

PIN / *IO*	Small (A)	Medium (B)	Large (C)	Difference (C) – (A)
		PIN		
Lowest (1)	0.171	0.309	0.512	0.341
(2)	0.156	0.245	0.421	0.265
(3)	0.134	0.198	0.326	0.192
(4)	0.123	0.175	0.265	0.142
Highest (5)	0.115	0.163	0.241	0.126
		Monthly Returns (%)		
Lowest (1)	0.912	1.651	1.735	0.823
(2)	0.884	0.994	1.481	0.597
(3)	0.991	1.116	1.145	0.154
(4)	1.156	1.275	1.278	0.122
Highest (5)	1.227	1.054	1.244	0.017
		MCAP (USD mil.)		
Lowest (1)	614.223	173.124	72.654	-541.569
(2)	1,696.343	492.467	238.044	-1,458.299
(3)	7,466.561	1,657.091	828.398	-6,638.163
(4)	11,861.299	3,322.938	1,656.327	-10,204.972
Highest (5)	12,971.113	4,448.534	2,302.354	-10,668.759
		IO		
Lowest (1)	0.036	0.031	0.025	-0.011
(2)	0.208	0.198	0.198	-0.010
(3)	0.381	0.376	0.367	-0.014
(4)	0.547	0.545	0.521	-0.026
Highest (5)	0.716	0.713	0.683	-0.033

Panel B: Portfolios Double-Sorted by *IO* and *AdjPIN*

IO \ *AdjPIN*	Small (A)	Medium (B)	Large (C)	Difference (C) – (A)
		AdjPIN		
Lowest (1)	0.131	0.239	0.378	0.247
(2)	0.123	0.208	0.334	0.211
(3)	0.115	0.176	0.271	0.156
(4)	0.102	0.154	0.226	0.124
Highest (5)	0.090	0.141	0.215	0.125
		Monthly Returns (%)		
Lowest (1)	1.016	1.460	1.715	0.699
(2)	0.867	1.219	1.456	0.589
(3)	0.984	1.125	1.145	0.161
(4)	1.211	1.181	1.305	0.094
Highest (5)	1.191	1.063	1.207	0.016
		MCAP (USD mil.)		
Lowest (1)	574.338	172.312	91.945	-482.393
(2)	1,624.389	592.648	216.003	-1,408.386
(3)	7,781.513	1,533.731	651.628	-7,129.885
(4)	12,141.779	3,417.978	1,358.434	-10,783.345
Highest (5)	13,222.687	4,334.675	2,511.212	-10,711.475
		IO		
Lowest (1)	0.031	0.021	0.021	-0.010
(2)	0.280	0.192	0.186	-0.094
(3)	0.389	0.376	0.375	-0.014
(4)	0.541	0.545	0.527	-0.014
Highest (5)	0.706	0.697	0.676	-0.030

Table 3: Institutional Ownership and Informed Trading

This table reports the coefficient estimates from the following regression:

$$INFO_{i,t} = \beta_0 + \beta_1 IO_{i,t-1} + \gamma'^{CONTROL_{i,t-1}} + e_{i,t},$$

where the dependent variable is either *PIN* or *AdjPIN*, both measured in year *t*. *IO* is institutional ownership, measured in year $t-1$. *CONTROL* is a vector of firm characteristics as discussed in Subsection 2.2. Panels A and B report the estimates of pooled OLS regressions where *PIN* and *AdjPIN* are used, respectively, as the dependent variables. Robust *t*-statistics in parentheses are based on standard errors clustered by firm and year. Sings *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Probability of Informed Trading (*PIN*)

	(1)	(2)	(3)	(4)	(5)	(6)
IO	-0.024***	-0.041***	-0.014***	-0.026***	-0.042***	-0.017***
	(-4.07)	(-5.61)	(-2.61)	(-4.66)	(-5.76)	(-3.23)
TOP5		0.039***			0.037***	
		(3.01)			(2.76)	
IOHHI			0.108***			0.110***
			(7.03)			(7.29)
SPREAD				0.027***	0.027***	0.029***
				(4.70)	(4.66)	(5.08)
ΔIO	0.010	0.004	0.024	0.011	0.005	0.025*
	(0.61)	(0.23)	(1.49)	(0.76)	(0.35)	(1.80)
SIZE	-0.032***	-0.031***	-0.028***	-0.032***	-0.032***	-0.028***
	(-26.37)	(-26.66)	(-21.34)	(-25.56)	(-26.09)	(-21.17)
LAGE	-0.001	-0.001	-0.001	-0.001	-0.001	-0.000
	(-1.47)	(-1.33)	(-0.65)	(-1.06)	(-0.94)	(-0.22)
MTB	-0.003*	-0.003	-0.004**	-0.005**	-0.004**	-0.006***
	(-1.67)	(-1.48)	(-2.17)	(-2.14)	(-1.97)	(-2.65)
TURN	-0.032***	-0.031***	-0.031***	-0.030***	-0.030***	-0.030***
	(-27.16)	(-27.68)	(-27.68)	(-25.41)	(-25.54)	(-25.41)
SP500	0.030***	0.031***	0.022***	0.031***	0.032***	0.022***
	(8.89)	(9.07)	(7.46)	(8.67)	(8.84)	(7.31)
LEV	0.002***	0.001***	0.002***	0.002***	0.002***	0.002***
	(3.02)	(2.76)	(3.39)	(3.23)	(3.00)	(3.62)
ROE	0.007***	0.007***	0.008***	0.006***	0.006***	0.007***
	(3.76)	(3.73)	(4.26)	(3.47)	(3.45)	(3.98)
VOL	-0.026***	-0.027***	-0.031***	-0.023***	-0.023***	-0.027***
	(-3.25)	(-3.26)	(-3.69)	(-2.65)	(-2.66)	(-3.07)
DD	-0.009***	-0.008***	-0.007***	-0.009***	-0.009***	-0.008***
	(-2.90)	(-2.87)	(-2.62)	(-3.15)	(-3.12)	(-2.88)
DIVER	0.002	0.001	0.004	-0.000	-0.001	0.002
	(0.43)	(0.35)	(1.02)	(-0.11)	(-0.19)	(0.56)
ALYST	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	(-3.09)	(-2.99)	(-3.29)	(-3.53)	(-3.42)	(-3.82)
INSIDE	0.028*	0.030*	0.027*	0.028*	0.029*	0.026*
	(1.75)	(1.73)	(1.68)	(1.75)	(1.76)	(1.71)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R-square	0.561	0.562	0.573	0.564	0.565	0.577
Obs.	16,866	16,866	16,866	16,866	16,866	16,866

Panel B: Adjusted Probability of Informed Trading (*AdjPIN*)

	(1)	(2)	(3)	(4)	(5)	(6)
IO	-0.013***	-0.039***	-0.011***	-0.014***	-0.040***	-0.012***
	(-3.28)	(-8.46)	(-2.74)	(-3.57)	(-8.80)	(-3.03)
TOP5		0.061***			0.060***	
		(6.76)			(6.50)	
IOHHI			0.022**			0.023**
			(1.97)			(2.10)
SPREAD				0.013**	0.013*	0.014**
				(2.03)	(1.91)	(2.02)
ΔIO	-0.015	-0.025*	-0.012	-0.015	-0.024**	-0.012
	(-1.34)	(-1.94)	(-1.08)	(-1.43)	(-2.04)	(-1.15)
SIZE	-0.023***	-0.022***	-0.022***	-0.023***	-0.022***	-0.022***
	(-24.71)	(-23.92)	(-26.11)	(-24.47)	(-23.82)	(-25.88)
LAGE	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(-4.23)	(-3.81)	(-3.98)	(-3.76)	(-3.43)	(-3.54)
MTB	-0.002	-0.001	-0.002	-0.002	-0.001	-0.002
	(-1.04)	(-0.63)	(-1.16)	(-1.38)	(-0.98)	(-1.52)
TURN	-0.019***	-0.019***	-0.019***	-0.018***	-0.018***	-0.018***
	(-20.10)	(-20.53)	(-19.64)	(-17.85)	(-18.03)	(-17.48)
SP500	0.010***	0.012***	0.009***	0.011***	0.012***	0.009***
	(5.69)	(6.09)	(5.33)	(5.98)	(6.39)	(5.63)
LEV	0.001**	0.001*	0.001**	0.001**	0.001*	0.001**
	(2.32)	(1.73)	(2.41)	(2.47)	(1.89)	(2.55)
ROE	0.005***	0.005***	0.006***	0.005***	0.005***	0.005***
	(4.90)	(4.80)	(5.04)	(4.46)	(4.39)	(4.59)
VOL	-0.027***	-0.027***	-0.028***	-0.025***	-0.025***	-0.026***
	(-4.53)	(-4.66)	(-4.50)	(-4.12)	(-4.21)	(-4.12)
DD	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(-1.35)	(-1.31)	(-1.21)	(-1.57)	(-1.54)	(-1.43)
DIVER	0.008**	0.008**	0.009***	0.007**	0.007**	0.008**
	(2.46)	(2.33)	(2.58)	(2.33)	(2.19)	(2.45)
ALYST	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(-3.04)	(-2.73)	(-3.08)	(-3.63)	(-3.28)	(-3.70)
INSIDE	0.032***	0.035***	0.032***	0.032***	0.035***	0.031***
	(2.78)	(3.13)	(2.80)	(2.81)	(3.15)	(2.83)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-square	0.508	0.511	0.509	0.509	0.512	0.510
Obs.	16866	16866	16866	16866	16866	16866

Table 4: Institutional Ownership and Informed Trading – Robustness Tests

This table reports the alternative regressions of *PIN* (*AdjPIN*) on institutional ownership (*IO*) and other control variables. Columns 1–3 report the Fama-MacBeth, firm-fixed effects, and pooled OLS regressions, respectively. Columns 4–5 report the results from a system of simultaneous equations in which the dependent variables are *PIN* (*AdjPIN*) (column 4) and *IO* (column 5). Robust *t*-statistics are in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Probability of Informed Trading (*PIN*)					
	Fama-Macbeth	Fixed-Effects	Pool OLS	Simultaneous Equation	
	(1)	(2)	(3)	(4)	(5)
IO	-0.025***	-0.019***	-0.015***	-0.048***	
	(-6.46)	(-3.73)	(-4.06)	(-10.43)	
TOP5	0.117***	0.063***	0.075***	0.031***	1.224***
	(6.73)	(8.92)	(6.81)	(3.92)	(14.12)
SPREAD	0.054***	0.014***	0.021***	0.031***	-0.054***
	(4.42)	(5.15)	(3.73)	(12.15)	(-12.56)
ΔIO	0.033**	0.021***	0.011	0.003	0.136***
	(2.51)	(2.84)	(0.89)	(0.36)	(10.91)
SIZE	-0.030***	-0.031***	-0.018***	-0.033***	0.034***
	(-20.09)	(-22.33)	(-18.02)	(-59.15)	(34.18)
LAGE	0.001*	-0.008***	0.001	0.000	0.001
	(2.00)	(-5.03)	(1.24)	(0.26)	(1.58)
MTB	-0.004***	0.004***	-0.007***	-0.005***	-0.002
	(-3.61)	(2.85)	(-4.04)	(-5.53)	(-1.64)
LEV	0.002***	-0.001	0.002***	0.002***	-0.001
	(6.24)	(-1.30)	(3.84)	(5.48)	(-1.05)
ROE	0.006***	-0.001	0.003	0.005***	0.001
	(3.54)	(-0.45)	(1.59)	(3.51)	(0.31)
VOL	-0.034***	-0.021***	-0.015**	-0.022***	-0.038***
	(-4.65)	(-3.18)	(-2.47)	(-4.37)	(-4.85)
TURN	-0.026***	-0.017***	-0.016***	-0.026***	0.035***
	(-31.68)	(-21.45)	(-19.10)	(-44.76)	(36.34)
DD	-0.006**	-0.000	-0.003	-0.007***	-0.004*
	(-2.83)	(-0.09)	(-1.64)	(-5.24)	(-1.80)
DIVER	0.014*	-0.011***	-0.000	-0.004	-0.014***
	(2.05)	(-3.44)	(-0.16)	(-1.37)	(-3.32)
SP500	0.024***	0.004	0.014***	0.032***	-0.002
	(11.43)	(0.96)	(6.26)	(15.94)	(-0.62)
ALYST	-0.004***	-0.006***	-0.002**	-0.004***	0.003***
	(-4.02)	(-4.07)	(-2.53)	(-7.21)	(3.76)
INSIDE	0.041**	-0.001	0.006	0.043***	0.021
	(2.27)	(-0.08)	(0.50)	(2.88)	(0.88)
LagPIN			0.394***		-0.155***
			(23.61)		(-13.45)
MOM					0.049***
					(16.47)
LPRC					0.058***
					(38.54)
Industry FE	Yes	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	No	No
R-square	0.566	0.750	0.644	0.553	0.802
Obs.	16,866	16,866	16,866	16,665	16,665

Panel B: Adjusted Probability of Informed Trading (*AdjPIN*)					
	Fama-Macbeth	Fixed-Effects	Pool OLS	Simultaneous Equation	
	(1)	(2)	(3)	(4)	(5)
IO	-0.040***	-0.017***	-0.033***	-0.032***	
	(-11.69)	(-3.19)	(-9.66)	(-9.15)	
TOP5	0.059***	0.009**	0.043***	0.041***	1.230***
	(9.02)	(2.36)	(5.35)	(6.72)	(14.21)
SPREAD	0.021**	0.002	0.010*	0.014***	-0.056***
	(2.45)	(0.91)	(1.87)	(7.29)	(-12.97)
ΔIO	-0.018*	-0.021***	-0.025**	-0.025***	0.134***
	(-2.13)	(-3.43)	(-2.35)	(-4.08)	(10.80)
SIZE	-0.022***	-0.024***	-0.017***	-0.024***	0.035***
	(-24.10)	(-20.52)	(-23.17)	(-56.60)	(35.16)
LAGE	-0.001	-0.011***	-0.001*	-0.001***	0.001
	(-1.76)	(-8.53)	(-1.78)	(-2.71)	(1.13)
MTB	-0.001	0.004***	-0.002*	-0.002***	-0.002*
	(-1.29)	(3.36)	(-1.65)	(-2.64)	(-1.77)
LEV	0.001**	-0.001	0.001***	0.001***	-0.001
	(2.98)	(-1.38)	(2.68)	(3.27)	(-1.25)
ROE	0.005***	0.001	0.003***	0.005***	0.000
	(3.90)	(0.43)	(2.82)	(4.90)	(0.12)
VOL	-0.024***	-0.003	-0.017***	-0.024***	-0.039***
	(-5.02)	(-0.48)	(-3.59)	(-6.37)	(-4.93)
TURN	-0.016***	-0.012***	-0.011***	-0.017***	0.035***
	(-24.75)	(-17.64)	(-16.95)	(-37.06)	(36.94)
DD	-0.002***	0.001	-0.001	-0.002*	-0.003
	(-3.13)	(0.48)	(-0.59)	(-1.69)	(-1.51)
DIVER	0.005	-0.002	0.004	0.006***	-0.013***
	(0.67)	(-0.56)	(1.28)	(3.13)	(-3.03)
SP500	0.011***	0.005	0.009***	0.013***	-0.004
	(6.36)	(1.51)	(5.89)	(8.52)	(-1.29)
ALYST	-0.002***	-0.000	-0.001**	-0.002***	0.004***
	(-5.51)	(-0.33)	(-2.49)	(-4.23)	(4.24)
INSIDE	0.041***	0.034***	0.031***	0.043***	0.015
	(3.99)	(2.76)	(3.36)	(3.80)	(0.65)
LagAdjPIN			0.278***		-0.188***
			(18.14)		(-12.26)
MOM					0.048***
					(16.03)
LPRC					0.058***
					(38.55)
Industry FE	Yes	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	No	No
R square	0.509	0.672	0.551	0.516	0.802
Obs.	16866	16866	16866	16665	16665

Table 5: Information-Hedge Portfolios and Abnormal Stock Returns

This table reports the regression results of Carhart's (1997) 4-factor model,

$$R_{hedge,t} = \alpha_{INFO} + \beta_1 MKTRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + e_{t,}$$

where the dependent variables are the monthly returns of the three tercile portfolios sorted by *PIN* (*AdjPIN*) and of a hedge portfolio with a long position in the large-*PIN* (large-*AdjPIN*) stocks and a short position in the small-*PIN* (small-*AdjPIN*) stocks. The intercept α_{PIN} (α_{AdjPIN}) is the abnormal return, *MKTRF*, *SMB*, and *HML* are the three factors of Fama and French (1993), and *UMD* is the momentum factor of Carhart (1997). Panel A presents the results for the equally weighted portfolios. Panel B reports the results for the value-weighted portfolios. Robust *t*-statistics are in parentheses. Signs *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Equally Weighted Portfolios

	Portfolios Sorted by *PIN* (*AdjPIN*)			
	Small	Medium	Large	Hedge Portfolio
	(A)	(B)	(C)	(C) – (A)
α_{PIN}	0.003***	0.004***	0.009***	0.005***
	(3.62)	(2.74)	(3.69)	(2.93)
MKTRF	1.028***	1.001***	0.763***	-0.265***
	(32.39)	(21.85)	(12.20)	(-4.12)
SMB	0.237***	0.639***	0.722***	0.485***
	(5.83)	(11.09)	(10.77)	(7.43)
HML	0.627***	0.6069***	0.470***	-0.157***
	(14.43)	(11.92)	(5.86)	(-1.67)
UMD	-0.191***	-0.210***	-0.203***	-0.013
	(-9.67)	(-4.34)	(-2.50)	(-0.15)
R^2	0.937	0.921	0.803	0.511
Obs.	144	144	144	144
α_{AdjPIN}	0.004***	0.003**	0.008***	0.004**
	(4.89)	(2.01)	(4.13)	(2.05)
MKTRF	1.019***	1.000***	0.770***	-0.249***
	(32.48)	(21.21)	(13.23)	(-4.57)
SMB	0.258***	0.649***	0.693***	0.435***
	(6.29)	(11.33)	(10.87)	(7.85)
HML	0.565***	0.649***	0.494***	-0.070
	(13.46)	(12.40)	(6.74)	(-0.86)
UMD	-0.185***	-0.221***	-0.199***	-0.014
	(-8.86)	(-3.95)	(-3.20)	(-0.25)
R^2	0.937	0.912	0.817	0.487
Obs.	144	144	144	144

Panel B: Value-Weighted Portfolios				
	Portfolios Sorted by *PIN* (*AdjPIN*)			
	Small	Medium	Large	Hedge Portfolio
	(A)	(B)	(C)	(C) – (A)
α_{PIN}	0.004***	0.004***	0.008***	0.004**
	(3.72)	(2.95)	(3.88)	(2.13)
MKTRF	1.026***	1.010***	0.785***	-0.240***
	(32.91)	(23.23)	(14.27)	(-4.19)
SMB	0.194***	0.601***	0.688***	0.493***
	(5.04)	(10.85)	(11.56)	(8.49)
HML	0.620***	0.626***	0.487***	-0.133
	(14.49)	(12.86)	(6.99)	(-1.57)
UMD	-0.172***	-0.177***	-0.166**	0.006
	(-8.72)	(-4.49)	(-2.45)	(0.08)
R^2	0.932	0.917	0.838	0.555
Obs.	144	144	144	144
α_{AdjPIN}	0.004***	0.003**	0.008***	0.003**
	(4.75)	(2.20)	(4.37)	(1.98)
MKTRF	1.024***	1.008***	0.789***	-0.236***
	(34.26)	(22.66)	(15.18)	(-4.68)
SMB	0.209***	0.608***	0.660***	0.451***
	(5.48)	(11.24)	(11.27)	(8.69)
HML	0.569***	0.665***	0.512***	-0.056
	(13.87)	(13.67)	(8.01)	(-0.76)
UMD	-0.163***	-0.190***	-0.162***	0.001
	(-9.05)	(-4.26)	(-3.07)	(0.02)
R2	0.939	0.914	0.847	0.545
Obs.	144	144	144	144

Table 6: Information Asymmetry, Institutional Ownership, and Expected Stock Returns – The Fama-MacBeth Approach

This table reports the coefficient estimates from the following the Fama-MacBeth (1973) regression models:

$$R_{i,t} = \beta_0 + \beta_1 INFO_{i,t-1} + \beta_2 Beta_{i,t-1} + \beta_3 SIZE_{i,t-1} + \beta_4 BTM_{i,t-1} + e_{i,t}\ ,$$

$$R_{i,t} = \beta_0 + \beta_1 (INFO_{i,t-1} \times DL) + \beta_2 (INFO_{i,t-1} \times DM) + \beta_3 (INFO_{i,t-1} \times DH) + \beta_4 DM + \beta_5 DH + \beta_6 Beta_{i,t-1} + \beta_7 SIZE_{i,t-1} + \beta_8 BTM_{i,t-1} + e_{i,t},$$

where the dependent variable is the monthly return of stock *i* in excess of a risk-free market return, measured in month *t*. *INFO* is measured by either *PIN* or *AdjPIN*, *Beta* is the stock beta estimated based on the data of the past 60 months, *SIZE* is the natural logarithm of market capitalization, and *BMT* is the natural logarithm of book-to-market ratio. The three dummies, *DL*, *DM*, and *DL*, denote, respectively, whether the stock belongs to the low-, medium- or high-*IO* group. Columns 1 and 3 report the results of the first model. Columns 2 and 4 show the results of the second model. Robust *t*-statistics are in parentheses. Signs *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	PIN	*PIN*	*AdjPIN*	*AdjPIN*
	(1)	(2)	(3)	(4)
INFO	0.009**		0.007	
	(1.98)		(1.44)	
INFO × *DL*		0.020***		0.024***
		(4.03)		(4.71)
INFO × *DM*		0.008		0.010
		(1.02)		(1.11)
INFO × *DH*		0.009		0.007
		(1.53)		(0.52)
DM		0.013***		0.011***
		(6.14)		(4.17)
DH		0.013***		0.010***
		(5.79)		(3.50)
SPREAD	0.007***	0.007***	0.008***	0.007***
	(6.77)	(6.31)	(7.08)	(6.78)
SIZE	0.002***	0.002***	0.001***	0.001
	(4.74)	(3.14)	(2.15)	(1.61)
BETA	-0.000	-0.001	-0.000	-0.001
	(-0.10)	(-0.51)	(-0.09)	(-0.68)
BTM	0.006***	0.007***	0.006***	0.007***
	(6.04)	(7.37)	(6.00)	(7.35)
Intercept	0.002	0.001	0.004	0.005
	(0.24)	(0.18)	(0.55)	(0.81)
Industry FE	Yes	Yes	Yes	Yes
No. of. Months	144	144	144	144

Table 7: Information Asymmetry, Institutional Ownership, and Abnormal Stock Returns – The 4-Factor Model

This table reports the coefficient estimates from the following 4-factor regression:

$$R_{hedge,t} = \alpha_{INFO} + \beta_1 MKTRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + e_t,$$

where the dependent variable is the monthly return of the hedge portfolio formed by taking a long position in the large-*PIN* (large-*AdjPIN*) stocks and a short position in the small-*PIN* (small-*AdjPIN*) stocks. The intercept α_{PIN} (α_{AdjPIN}) is the abnormal return of the portfolio, *MKTRF*, *SMB*, and *HML* are the three factors of Fama and French (1993), and *UMD* is the momentum factor of Carhart (1997). Panel A reports the results for the equally weighted portfolios. Panel B displays the results for the value-weighted portfolios. Robust *t*-statistics are in parentheses. Signs *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Equally Weighted Portfolios

	Hedge Portfolios Sorted by *IO*				
	Lowest (1)	(2)	(3)	(4)	Highest (5)
α_{PIN}	0.009***	0.006***	0.002	0.001	0.001
	(5.11)	(2.64)	(1.43)	(1.23)	(1.47)
MKTRF	-0.220***	-0.191**	-0.101**	-0.094**	-0.062*
	(-3.94)	(-2.48)	(-2.21)	(-2.39)	(-1.91)
SMB	-0.050	0.126*	0.485***	0.447***	0.366***
	(-0.73)	(1.75)	(8.14)	(7.42)	(7.61)
HML	0.027	0.101	0.058	0.157***	0.036
	(0.38)	(0.98)	(0.75)	(2.93)	(0.80)
UMD	-0.023	0.093*	0.037	0.044	0.085***
	(-0.67)	(1.68)	(0.93)	(1.12)	(2.79)
R^2	0.214	0.208	0.551	0.534	0.558
Obs.	144	144	144	144	144
α_{AdjPIN}	0.007***	0.004**	0.002*	0.001	0.000
	(3.21)	(2.03)	(1.64)	(0.19)	(0.35)
MKTRF	-0.159***	-0.108*	-0.084	-0.103**	-0.121***
	(-2.97)	(-1.67)	(-1.77)	(-2.38)	(-3.55)
SMB	-0.049	0.111**	0.457***	0.504***	0.394***
	(-0.75)	(2.22)	(9.08)	(8.89)	(7.83)
HML	-0.002	-0.002	0.082	0.194***	0.098**
	(-0.03)	(-0.03)	(1.20)	(3.87)	(2.15)
UMD	-0.002	0.029	0.045	0.030	0.060**
	(-0.05)	(0.90)	(1.42)	(0.69)	(2.02)
R^2	0.124	0.091	0.550	0.588	0.589
Obs.	144	144	144	144	144

Panel B: Value-Weighted Portfolios					
	Hedge Portfolios Sorted by *IO*				
	Lowest (1)	(2)	(3)	(4)	Highest (5)
α_{PIN}	0.008***	0.005***	0.002	0.001	0.001
	(4.83)	(2.71)	(1.23)	(1.02)	(0.29)
MKTRF	-0.230***	-0.180**	-0.085*	-0.086**	-0.054*
	(-4.41)	(-2.44)	(-1.90)	(-2.28)	(-1.65)
SMB	-0.052	0.154**	0.501***	0.452***	0.364***
	(-0.78)	(2.19)	(8.49)	(7.60)	(7.74)
HML	0.006	0.068	0.061	0.161***	0.028*
	(0.09)	(0.66)	(0.81)	(3.08)	(1.64)
UMD	-0.023	0.081*	0.041	0.043	0.089***
	(-0.71)	(1.70)	(1.09)	(1.18)	(2.93)
R^2	0.217	0.194	0.578	0.539	0.574
Obs.	144	144	144	144	144
α_{AdjPIN}	0.006***	0.004*	0.002	0.001	0.000
	(3.15)	(1.93)	(0.34)	(0.08)	(0.49)
MKTRF	-0.175***	-0.104	-0.079*	-0.096**	-0.112***
	(-3.26)	(-1.60)	(-1.71)	(-2.30)	(-3.37)
SMB	-0.045	0.136***	0.477***	0.509***	0.385***
	(-0.68)	(2.61)	(9.59)	(9.03)	(7.69)
HML	-0.020	0.001	0.084	0.197***	0.096**
	(-0.27)	(0.01)	(1.30)	(4.01)	(2.17)
UMD	-0.007	0.037	0.047	0.032	0.062**
	(-0.17)	(1.14)	(1.55)	(0.77)	(2.12)
R^2	0.134	0.108	0.591	0.604	0.594
Obs.	144	144	144	144	144

Table 8: Information Asymmetry, Institutional Ownership, and Abnormal Stock Returns – The 5-Factor Model

This table reports the coefficient estimates from the following 5-factor model:

$$R_{hedge,t} = \alpha_{INFO} + \beta_1 MKTRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \beta_5 PSLIQ_t + e_t,$$

where the dependent variables are the monthly returns of the hedge portfolios formed by taking a long position in the large-*PIN* (large-*AdjPIN*) stocks and a short position in the small-*PIN* (small-*AdjPIN*) stocks. The intercept α_{PIN} (α_{AdjPIN}) is the abnormal return of the hedge portfolios, *MKTRF*, *SMB*, and *HML* are the three factors of Fama and French (1993), *UMD* is the momentum factor of Cahart (1997), and *PSLIQ* is the liquidity factor of Pastor and Stambaugh (2003). Panel A reports the results for the equally weighted portfolios. Panel B provides the results for the value-weighted portfolios. Robust *t*-statistics are in parentheses. Signs *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Equally Weighted Portfolios					
	Hedge Portfolios Sorted by *IO*				
	Lowest (1)	(2)	(3)	(4)	Highest (5)
α_{PIN}	0.008***	0.006***	0.001	0.001	0.000
	(4.80)	(2.96)	(0.24)	(0.02)	(0.37)
MKTRF	-0.247***	-0.165**	-0.086*	-0.076*	-0.057*
	(-4.38)	(-2.04)	(-1.81)	(-1.81)	(-1.70)
SMB	-0.064	0.139*	0.492***	0.456***	0.369***
	(-0.97)	(1.92)	(8.21)	(7.61)	(7.68)
HML	-0.005	0.130	0.075	0.177***	0.042
	(-0.06)	(1.21)	(0.94)	(3.18)	(0.91)
UMD	-0.018	0.089*	0.034	0.041	0.084***
	(-0.52)	(1.68)	(0.84)	(1.01)	(2.73)
PSLIQ	0.104**	-0.098	-0.057	-0.068*	-0.021
	(2.17)	(-1.63)	(-1.19)	(-1.84)	(-0.59)
R^2	0.241	0.217	0.563	0.541	0.564
Obs.	144	144	144	144	144
α_{AdjPIN}	0.006***	0.004**	0.002	0.001	0.001
	(3.18)	(2.10)	(0.57)	(0.09)	(0.60)
MKTRF	-0.168***	-0.100	-0.079	-0.081	-0.105**
	(-2.93)	(-1.48)	(-1.59)	(-1.77)	(-3.10)
SMB	-0.053	0.114**	0.460***	0.514***	0.402***
	(-0.81)	(2.30)	(9.01)	(9.32)	(8.16)
HML	-0.012	0.006	0.088	0.218***	0.116**
	(-0.16)	(0.08)	(1.23)	(4.12)	(2.42)
UMD	-0.001	0.027	0.044	0.027	0.057*
	(-0.02)	(0.86)	(1.38)	(0.58)	(1.92)
PSLIQ	0.034	-0.027	-0.019	-0.080**	-0.059**
	(0.69)	(-0.54)	(-0.47)	(-2.44)	(-2.06)
R^2	0.124	0.089	0.551	0.604	0.598
Obs.	144	144	144	144	144

Panel B: Value-Weighted Portfolios					
	Hedge Portfolios Sorted by *IO*				
	Lowest (1)	(2)	(3)	(4)	Highest (5)
α_{PIN}	0.007***	0.005***	0.001	0.001	0.000
	(4.52)	(2.97)	(1.03)	(0.29)	(0.21)
MKTRF	-0.256***	-0.158**	-0.070	-0.067*	-0.049
	(-4.79)	(-2.04)	(-1.50)	(-1.67)	(-1.48)
SMB	-0.065	0.165**	0.509***	0.461***	0.366***
	(-1.00)	(2.33)	(8.57)	(7.80)	(7.80)
HML	-0.023	0.093	0.078	0.183***	0.033
	(-0.34)	(0.87)	(1.00)	(3.37)	(0.73)
UMD	-0.019	0.078*	0.039	0.040	0.088***
	(-0.57)	(1.68)	(1.00)	(1.05)	(2.87)
PSLIQ	0.096**	-0.084	-0.057	-0.072*	-0.017
	(2.05)	(-1.45)	(-1.21)	(-1.93)	(-0.48)
R^2	0.241	0.204	0.587	0.551	0.574
Obs.	144	144	144	144	144
α_{AdjPIN}	0.006**	0.004**	0.002	0.001	0.001
	(3.12)	(2.01)	(1.26)	(0.39)	(0.76)
MKTRF	-0.184**	-0.095	-0.073	-0.073*	-0.095***
	(-3.21)	(-1.39)	(-1.51)	(-1.65)	(-2.89)
SMB	-0.050	0.140***	0.480***	0.521***	0.394***
	(-0.74)	(2.70)	(9.55)	(9.52)	(8.09)
HML	-0.029	0.011	0.090	0.224***	0.115**
	(-0.38)	(0.14)	(1.34)	(4.34)	(2.50)
UMD	-0.006	0.036	0.046	0.028	0.059**
	(-0.13)	(1.10)	(1.49)	(0.64)	(2.01)
PSLIQ	0.033	-0.033	-0.022	-0.087***	-0.063**
	(0.66)	(-0.66)	(-0.56)	(-2.67)	(-2.23)
R^2	0.133	0.114	0.587	0.617	0.603
Obs.	144	144	144	144	144

Figure 1: Institutional Ownership and Informed Trading

This figure plots the mean *IO* by *PIN* (left panel) and *AdjPIN* (right panel) quintile using yearly sample data for the period from 1993 to 2004.