Previous version: January 2007
This version: July 2007

# Is Information Risk (PIN) Priced?

Partha Mohanram
Phillip H. Geier Jr. Associate Professor
Columbia Business School
3022 Broadway, New York, NY 10027
Tel: 212-854-2561, Fax: 212-316-9219
E-mail: pm2128@columbia.edu

Shiva Rajgopal
Herbert O. Whitten Endowed Professor
University of Washington Business School
Box 353200
Seattle, WA 98195
Tel: 206-543-7913, Fax: 206-685-9392
E-mail: rajgopal@u.washington.edu

**Abstract:**

Several recent papers assume that private information (PIN), proposed by Easley, Hvidkjaer and O'Hara (2002, 2004), is a determinant of stock returns. In this paper, we formally test whether PIN is indeed priced. We first replicate Easley, Hvidkjaer and O'Hara (2002) and show that while PIN does predict future returns in the sample they analyze, the effect is not robust to alternative specifications and time periods. Further, we find no evidence that PIN factor loadings predict returns. Finally, PIN exhibits no association with ex-ante measures of cost of capital derived from analysts' earnings forecasts. Overall, our findings cast doubt on whether PIN reflects information risk systematically priced by investors.

**Keywords:** Information Risk, PIN, Risk.

**JEL Classification:** G12, G14, M40

---

We thank our respective schools for financial support. We acknowledge helpful written comments from an anonymous referee, Ross Watts (editor), Hemang Desai, Gus DeFranco, Jefferson Duarte, Mozaffar Khan, Stephen Penman, Gil Sadka, Lakshmanan Shivakumar, K.R. Subramanyam and workshop participants at Columbia University, MIT, University of Washington and Rutgers University. All errors are ours.

# Is Information Risk (PIN) Priced?

## 1. Introduction

An influential set of recent papers by Easley, Hvidkjaer and O'Hara (2002, 2004) suggests that information risk based on private information in a stock and proxied by the probability of informed trading measure, PIN, is a determinant of stock returns. The magnitude of returns affected by PIN is pretty large as well. Easley et al. (2002, 2004) find that (i) a 10% increase in PIN is associated with an increase in annual expected returns of 2.5%, on average; and (ii) a zero-investment portfolio that is size neutral, but long in high PIN stocks and short in low PIN stocks, earns a mean monthly return of 0.27% with a t-statistic of 2.86. Easley et al. (2004) interpret these data as evidence that PIN captures information risk that is systematically priced by investors.

Several recent papers in the finance and accounting literatures (i) explicitly link changes in PIN to changes in cost of capital; (ii) assert that PIN is reflected in stock prices or credit ratings of firms; or (iii) cite the Easley et al. (2002) result that higher PIN is associated with higher cost of capital (see the appendix for detailed citations). Assuming that the pricing of PIN is a robust result (which we demonstrate is not the case), whether PIN even ought to be priced is debatable. In particular, Lambert, Leuz, and Verrecchia (2006) argue that information risk is potentially fully diversifiable and hence does not have to be priced. Similarly, Hughes, Liu, and Liu (2005) conclude that information risk is either diversifiable or subsumed by existing risk factors. Further, Spiegel and Wang (2005, footnote 6) suggest that PIN captures a stock's liquidity characteristics and whether liquidity is a systematic risk is unclear.

We investigate whether PIN reflects information risk that is systematically priced by investors. In particular, we conduct three tests. First, we investigate whether the Easley et al. (2002) result covering the 1984-1998 period that PIN predicts future stock returns is robust to multiple periods of varying time-series length. Second, we conduct Fama and MacBeth (1973) type cross-sectional regressions of returns on PIN, PIN factor loadings, and other average return predictors on several portfolios based on size, PIN and PIN factor loadings. If PIN represented priced information risk, we expect the coefficient on PIN factor loadings in such cross-sectional regressions to be statistically significant. Finally, we investigate whether PIN and ex-ante measures of cost of capital derived from analysts' earnings forecasts are positively correlated. If PIN were priced risk, we would expect higher PIN to be associated with higher ex-ante cost of capital.

Overall, our results do not provide support for the hypothesis that PIN is priced information risk. In particular, the Easley et al. (2002) result that PIN is priced is restricted to the 1984-1988 time period. Moreover, although the Easley et al. (2002) result that PIN predicts returns is primarily driven by small firms, a majority of small firms, counter-intuitively, have negative loadings on the PIN factor, suggesting that cost of capital for small firms is actually decreasing in information risk. In the cross-sectional tests, we find that the PIN factor loadings do not exhibit any statistical association with returns. Moreover, across several specifications, we could not find even one case where PIN exhibits a statistically significant positive association with ex-ante cost of capital measures.

A combined reading of the results is disconcerting to the notion that PIN is priced information risk. Our evidence suggests that there is no robust return premium associated with the PIN factor, and the difference in returns attributed to the PIN factor cannot be confidently viewed as compensation for information risk.

The remainder of the paper is organized as follows. Section 2 replicates the return premium to PIN demonstrated by Easley et al. (2002) and assesses the robustness of that result. Section 3 constructs the PIN factor in line with Easley et al. (2004) and investigates whether PIN factor loadings predict returns. Section 4 investigates the association between PIN and ex-ante cost of capital measures. Section 5 concludes.

## 2. Return premium to the PIN characteristic

### *2.1 Theoretical arguments related to the pricing of information risk*

Traditional asset-pricing theory (e.g., Fama, 1991) assumes that information risk is completely diversifiable and should hence have no effect on expected returns. A recent stream of papers by Easley, Hvidkjaer and O'Hara (2002) and Easley and O'Hara (2004) has developed the theoretical intuition for why information risk ought to be priced. These Easley et al. papers propose a microstructure model to formalize the learning problem confronting a market maker in a world with informed and uninformed traders. When information about the payoff on risky assets is private rather than public, and uninformed investors cannot perfectly infer such private information from prices, they require a greater expected excess return. To completely avoid this risk, uninformed traders would have to hold only the risk-free asset, which would be suboptimal as compared to holding some of the risky, private information assets. Because uninformed investors are rational, they hold an optimally diversified portfolio, but no matter how they diversify, uninformed traders are taken advantage of by informed traders who have learned which assets to hold.

Recent theory work in accounting has questioned the basic result of the Easley et al papers. Lambert, Leuz, and Verrecchia (2006) argue that the Easley and O'Hara (2004) result that information risk ought to be priced is not unambiguous. In particular, they find that when the number of traders becomes sufficiently large, information risk does not have to be priced, i.e. it is fully diversifiable. Lambert, Leuz and Verrecchia (2006) further argue that even if information risk is non-diversifiable, such risk ought to be picked up by a well specified forward-looking beta in traditional asset pricing tests and information risk by itself does not need to be an independent risk factor. Similarly, Hughes, Liu, and Liu (2005) study the role of information risk in a multi-factor asset pricing mode and conclude that information risk is either diversifiable or subsumed by existing risk factors. Given (1) the critical role that the accounting process plays in the generation of information; (2) the reliance in emerging accounting literature on the interpretation that PIN is priced information risk; and (3) the controversy surrounding whether information risk commands a premium, we believe that it is important to empirically investigate whether information risk, as embodied by PIN, is indeed priced.

### *2.2 Estimation of PIN*

The PIN estimation methodology is detailed in Easley et al. (2002, 2004). To summarize this methodology, given a history of trades, the market maker can estimate the probability that the next trade is from an informed trader. Easley et al. (2002) show that this probability of information-based trade is given by:

$$PIN = \frac{\alpha\mu}{\alpha\gamma + \varepsilon_s + \varepsilon_B} \quad (1)$$

where α is the probability that there is new information at the beginning of the trading day, μ is the arrival rate of orders from informed traders, $\varepsilon_S$ is the arrival rate of orders from uninformed sellers and $\varepsilon_B$ is arrival rate of orders from uninformed buyers. The numerator in (1) represents the arrival rate of information based orders and the denominator in (1) is the arrival rate for all orders. Thus, PIN in expression (1) is the fraction of orders that arise from informed traders relative to the overall order flow. Easley et al. (2002, 2004) estimate PIN for specific stocks using maximum likelihood estimation with trade and quote data for stocks listed on the New York Stock Exchange (NYSE) and American Stock Exchange (AMEX) stocks.[1]

We rely on the dataset of PIN estimates graciously provided by Professor Soren Hvidkjaer on his personal website (http://www.smith.umd.edu/faculty/hvidkjaer/). The dataset covers the sample of all ordinary common stocks listed on NYSE and AMEX for the years 1983 − 2001. The dataset excludes REITs, stocks of companies incorporated outside of the U.S, and closed-end funds. Also excluded are stocks in any year in which the stock did not have at least 60 days with quotes or trades, as PIN cannot be reliably estimated for such stocks. Further, since PIN and size portfolios are based on year-end firm size, also excluded are stocks for which this information is not available. The final sample has between 1,863 and 2,414 stocks in the years 1983 − 2001. For further details on the construction and content of the dataset, see Easley et al. (2004).

[1] Asquith, Oman and Safaya (2006) point out that the Lee-Ready algorithm, used by Easley et al. (2002, 2004), to classify trades as buys or sells for computing PIN, is prone to measurement error. While we acknowledge this limitation, we note that Easley et al. (2002, 2004) show that PIN is priced despite the measurement error in classifying trades as buys or sells. Our departure point in the paper is the Easley et al. (2002, 2004) result that PIN is priced. Having said that, we acknowledge that we might have had difficulty showing that PIN is priced on account of such measurement error in PIN.

Descriptive data reported here confirms that our sample matches theirs (Table 1). In particular, the average of the yearly cross-sectional median PINs is 0.196. The means of the yearly 25th and 75th percentiles of PIN are 0.154 and 0.250 respectively. Similar to Easley et al. (2004), there appears to be a strong correlation between PIN and size (average $\rho$ = -0.660).[2] In a recent paper Aslan, Easley, Hvidkjaer and O'Hara (2006) explore the firm characteristics associated with PIN and find that PIN is (i) negatively correlated with analyst following, institutional ownership, share turnover and Tobin's q; and (ii) positively correlated with smaller firms, ROA, and stock return volatility.

### *2.3 Replicating the Easley et al. 2002 result that PIN is priced*

To allow comparability with previous work, we seek to replicate the result (reported in table VI of the Easley et al. (2002) *Journal of Finance* paper) that the PIN characteristic is priced. In particular, Easley et al. (2002) use PIN data over the years 1983-1997 and regress one-year ahead monthly returns in excess of the risk free rate over the years 1984-1998 on beta, PIN, BM (book-to-market) and size characteristics measured at the end of year *t-1*.

Following Easley et al. (2002), we calculate pre-ranking portfolio betas estimated for individual stocks using monthly returns from at least two years, when possible, five years before the test year. Thus, for each stock, we use at least 24 monthly return observations in the estimation. We regress these stock returns on the contemporaneous and lagged value-weighted CRSP NYSE/AMEX index. Pre-ranking portfolio betas are then computed as the sum of the two coefficients. Next, 40 portfolios are sorted every

[2] The number of observations per year appears to be almost identical but not exactly the same as reported by Easley et al. (2004). Given that we use the data provided by them, our explanation for this difference is either that the data were updated, or that a few observations were deleted in their analysis because of the lack of availability of some other data items.

January on the basis of the estimated betas, and monthly portfolio returns are calculated as equal-weighted averages of the individual stock returns. Post-ranking portfolio betas are estimated from the full sample period, such that one beta estimate is obtained for each of the 40 portfolios. Portfolio returns are regressed on contemporaneous and lagged values of CRSP index returns. The portfolio beta, $\beta_p$, is then the sum of the two coefficients. We use individual stocks in the cross-sectional regressions, so individual stock betas are taken as the portfolio to which they belong.

Book value of equity is obtained from annual COMPUSTAT files (data #60). Following Easley et al. (2002), we exclude negative BM values, and set BM outside the 0.005 and 0.995 fractiles equal to these fractiles, respectively. We take logs, such that the explanatory variable, $BM_{it-1}$ is LBM for firm *i*. SIZE is the log of the market value of equity at the end of year *t-1*. For each month in the sample period 1984-1998 related to stock returns, we run the following cross-sectional regression:

$$R_{it} = \gamma_{ot} + \gamma_{1t}\,\beta_p + \gamma_{2t}\,PIN_{it-1} + \gamma_{3t}\,SIZE_{it-1} + \gamma_{4t}\,LBM_{it-1} + error_{it} \qquad (2)$$

where $R_{it}$ is the excess return of stock *i* in month *m* of year *t*, and $\gamma_{jt}$ represents the estimated coefficients. The coefficients from the cross-sectional regressions are averaged over time, using the standard Fama and MacBeth (1973) methodology. To address the inefficiency of this procedure related to time-varying volatility, we also use the correction suggested by Litzenberger and Ramaswamy (1979). This correction weights the coefficients by their precisions when summing across the cross-sectional regressions.

The results of estimating (2) over 1984-1998 are reported in Panel A of Table 2 and mirror closely those reported by Easley et al. (2002). In particular, we find a positive and statistically significant coefficient on PIN (t-statistic = 2.75 under Fama-MacBeth and 3.46 under Litzenberger-Ramaswamy (henceforth LR) correction). Both the magnitude of the mean coefficient on PIN as well as the level of significance is similar to Easley et al. (2002). Thus, we are able to replicate the basic Easley et al. (2002) result that PIN appears to be priced for the sample period 1984-1998.

Interestingly, panel B shows that PIN, by itself, is not related to returns (t-statistic = 0.59 under Fama-MacBeth and 0.34 under LR correction). Thus, PIN appears to load only when accompanied by other variables, especially SIZE, as in panel A.

*2.4 Different time windows*

In panel C, we examine whether the pricing of PIN is robust across several sub-periods for two reasons. First, we want to assess whether the pricing over 1984-1998 extends to the four additional years, 1998 to 2002, for which data is now available but was absent when Easley et al. (2002) wrote their paper. Second, we want to examine whether the basic result that PIN is priced is true in sub-periods within the time period being analyzed. In particular, we replicate the Easley et al. (2002) result for five-year windows (1984-1988; 1989-1993; 1994-1998; 1999-2002). The 1999-2002 window only has four years of data because we do not have access to PIN data for 2003. We argue that these sub-samples should have sufficient statistical power to detect pricing of PIN because we rely on Fama-Macbeth tests that between 48 and 60 months of data.

It is interesting to note that PIN is positively related to returns only in the 1984-1988 period where the L-R t-statistic on PIN is 2.96. The maximum L-R t-statistic attained in any other sub-period is only 1.57. Thus, the basic Easley et al. result that PIN is priced appears to be restricted to the 1984-1988 period. The pricing of PIN does not appear to be robust to an extended time period and specification changes.

*2.5 PIN-size portfolios-dependent sorts*

Given the earlier finding that PIN and size are negatively correlated, we attempt to isolate the effects of PIN by first sorting stocks on the basis of size, and then sorting on PIN within size groups.[3] In particular, at the beginning of the year *t*, we sort stocks into three equal groups based on market capitalization at the end of the prior year (*t-1*). Next, within each size group, we sort into three equal-sized groups based on PINs from the prior year (labeled as S for Small, M for Medium and B for Big). The sequential sorting procedure yields nine portfolios (S/Low PIN, S/Mid PIN, S/High PIN, M/Low PIN, M/Mid PIN, M/High PIN, B/Low PIN, B/Mid PIN, B/High PIN). This sequential sorting process ensures that each of the nine sub-portfolios has roughly an equal number of firm-year observations (approximately 4,375 firm years). We rely on these sequentially sorted portfolios in the remainder of the paper.

Table 3 reports descriptive data on PIN, Size and value-weighted monthly returns in excess of one-month T-bill rates (*Exret*) for each of these nine portfolios are computed from January to December of year *t*. The data in Table 3 reveal several interesting patterns. First, sorting PIN into three portfolios, keeping size constant, does appear to capture reasonable independent variation in PIN independent of size. In particular, PIN

[3] The traditional asset-pricing literature has relied on independent sorts of variables whose ability to predict returns is being tested (Fama and French 1993, 1996, Daniel and Titman 1997, Davis, Fama and French 2000, and Hirschleifer, Hou and Teoh 2006). However, we are unable to follow the tradition in this setting because of the high correlation between PIN and SIZE. That is, when independent sorts into nine size-PIN groups are attempted, we find that the spread in returns between Low PIN and High PIN groups is statistically insignificant in every size partition. To avoid unfairly penalizing the ability of PIN to predict returns, due to its high correlation with size, we depart from tradition and perform sequential sorts on PIN within size groups. This approach is also consistent with how Easley et al. (2004) construct the PIN factor. We hasten to add that we have replicated all tables in the paper using independent sorts of PIN and size. The fundamental inferences (related to whether or not PIN is robustly able to predict returns or behave like a risk factor) remain similar to the ones reported in the paper based on sequential sorts.

spreads in each size group, reported in the last two columns of panel A, are strongly significant at conventional levels. It is interesting to note that the spread in PINs of 0.175 is highest for the smallest size group and lowest for the largest size group (spread = 0.086), suggesting that PIN is likely to have greater potential to explain returns for small stocks relative to large stocks.

Second, panel B shows that for a given size category, as PIN increases, the average size declines, given that size and PIN are correlated. That is, the largest firms fall in the B/Low PIN group (average market capitalization = $10.374 billion), while the smallest firms fall in the S/High PIN group (average market capitalization = $29.37 million). This outcome seems intuitive because the largest firms are most likely well followed by information intermediaries such as analysts and are likely to be associated with lower private information relative to the smallest firms.

Third, as panel C reports, for the smallest size group (S), higher PIN is associated with greater returns, consistent with the hypothesis that more information asymmetry (higher PIN) is associated with greater expected returns. In particular, S/Low PIN portfolio earns returns of 0.514% per month while the S/High PIN portfolio earns 1.094% per month. The resultant spread of 0.58% per month is statistically significant (t-statistic = 3.23). This spread can be interpreted as the mean return on a composite zero-investment portfolio formed by taking long (short) positions of equal size in the high (low) PIN portfolios. Further, there is no spread between high PIN and low PIN stocks for both the medium size firms (0.02% per month, t-statistic = 0.18) and large firms (0.031%, t-statistic = 0.22). Thus, the key point that emerges from Table 3 is that we observe economically significant abnormal return to PIN only among small stocks.

## 3. Cross-sectional tests of returns on PIN factor loadings

Easley et al. (2004, page 3) believe that PIN might be a factor in returns because a zero investment PIN portfolio that they create earns abnormal returns not explained by the usual Fama and French (1993) factors. To evaluate the robustness of this assertion and to provide another test of whether stock prices reflect a systematic risk premium for PIN, we estimate a cross-sectional regression of returns on PIN factor loadings, and examine the statistical significance of the coefficients of these loadings. To do so, in the following sections, we describe how we create a PIN factor and estimate PIN factor loadings for every firm.

### *3.1 Creating the PIN factor*

We form PIN factors based on PIN and size groups formed via dependent sorts in accordance with Easley et al. (2004). In particular, at the end of December of each year *t* from 1983 to 2001, all stocks on NYSE and AMEX with non-missing size and PIN data are assigned to size decile, and within each decile, three equal size groups are formed on the basis of PIN. We then compute value-weighted hedge returns for each size decile of portfolios long on high PIN firms and short on low PIN firms. The PIN factor is defined as the (equally weighted) average of the hedge returns for each of the ten size deciles.[4]

The descriptive statistics reported in panel A of Table 4 related to the PIN factor and the other factors closely resemble those reported in Easley et al. (2004). The correlation table in panel B shows that the PIN factor returns exhibit modest correlation with SMB and the HML factor returns ($\rho$ = -0.09 and 0.027, respectively), but reports a

[4] While we use deciles in the creation of the PIN factor to be consistent with Easley et al. (2004), we use three size groups in the rest of the paper because of the additional partitions on PIN and PIN loading used in the remainder of the paper. Our results are similar if we form the PIN factor on the basis of three size groups instead of deciles.

strong correlation with the momentum factor UMD ($\rho$ = 0.576). At first blush, the low correlation between SMB and PIN factor appears to be inconsistent with the high correlation between PIN and size. Recall, however, that we create the PIN factor within size groups using dependent sorts, partly to counter the high correlation between size and PIN.[5] On a different note, the relatively high correlation between UMD and PIN factor underscores the need to control for momentum when considering the PIN factor.

*3.2 Does the PIN factor load for the entire sample?*

In panel C of Table 4, we investigate whether the PIN factor explains returns for the entire sample. In particular, we compute value-weighted returns for the entire sample of firms from January to December of year *t*. We then estimate the Fama-French three-factor model enhanced by the momentum factor (UMD) and a five-factor model that adds the PIN factor (PINF) by regressing the value-weighted monthly returns in excess of the one-month T-bill rates, $R_{it} - R_{ft}$ on the relevant factors. The sample covers 228 months of data (19 years for which PIN data is available and 12 months of data per year). In other words, for each portfolio i we perform the following time series regressions:

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{mt} - R_{ft}) + s_i\, SMB_t + h_i\, HML_t + m_i\, UMD_t + error_{it} \qquad (3)$$

[5] One of the problems with using dependent sorts as in Easley et al. (2004) to create a PIN factor is that the second sort on PIN is almost a second (inverted) sort on size, given the high negative correlation between size and PIN. One solution to this is to employ a methodology motivated by Penman (1983), a paper which analyzes the information content of management earnings forecasts to that of dividend announcements which are two phenomena that are also highly correlated. In Easley et al. (2004), the hedge return in each decile is defined as (VRET High PIN – VRET Low PIN), where VRET refers to the value weighted average return for each group. In our setting, we implement a second sort both on PIN as well as a repeated second sort on size. Given that PIN is likely to be highly negatively correlated with size, low PIN is likely to pick up the large firms within each decile, while high PIN will pick up the small firms within each decile. To remove the effect of size, we define the hedge return instead as (VRET High PIN – VRET Small) - (VRET High PIN – VRET Large), where large and small refer to the size groupings under the second sort on size within each decile. As before, we average the hedge returns across all decile to create the alternate PIN factor. When we use this alternate definition of the PIN factor, none of our results changes in any substantive manner. Hence it is unlikely that the weak performance of the PIN characteristic or loading (as shown later) in explaining returns is driven by the strong negative correlation between PIN and size.

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{mt} - R_{ft}) + s_i\, SMB_t + h_i\, HML_t + + m_i\, UMD_t + p_i\, PINF_t + \varepsilon_{it} \qquad (4)$$

The first row of panel C presents the results from estimating the standard three-factor Fama-French model for our sample. The adjusted $R^2$ for this regression is a high 94.3% and more notably, the intercept term from such an estimation is statistically insignificant (t-statistic = -0.12). Thus, the three-factor model seems to adequately capture the cross-section of returns in our sample period.[6] Row 2 shows that the momentum factor loads (t-statistic = -1.96), and the adjusted $R^2$ increases a bit to 94.4%. Row 3 reports that the PIN factor also weakly predicts returns for the entire sample when introduced by itself (coefficient = 0.0826 and t-statistic = 1.87). However, when UMD and PIN are introduced together, as in Row 4, both factors attain strong statistical significance, and the adjusted $R^2$ goes up to 94.7% from the earlier 94.4%. In particular, consistent with Easley et al. (2004)'s interpretation of PIN as a priced factor, we find that the coefficient of 0.1999 on the PIN factor is positive, large, and statistically significant at the 1% level (t-statistic = 3.81).[7]

However, evidence of a statistically significant coefficient on PINF is not sufficient to interpret PINF as a priced risk factor. To confirm that PINF is a priced risk factor, we need to show that PIN factor loadings, by themselves, can predict returns. To

---

[6] A reader might wonder why the coefficients on SMB, HML, UMD and PIN factors are not zero by construction when equation (4) is estimated for the entire sample. Note that SMB, HML, UMD are constructed from the universe of firms (NYSE, NASDAQ and AMEX) whereas the PIN sample is available only for NYSE and AMEX firms. Further, the PIN factor is constructed by averaging returns on PIN factor mimicking portfolios within size groups.

[7] Why does the PIN factor attain statistical significance when the intercept term from the three-factor model is zero? We believe there are two explanations for this. First, as seen in Table 4, UMD factor returns are highly correlated with PIN returns ($\rho = 0.576$). Thus, PIN and UMD borrow some of their explanatory power from each other and collectively do not seem to provide significantly new explanatory power. Second, the PIN characteristic predicts stock returns only for a sub-section of firms (small firms) and not the entire portfolio of firms. Hence, the intercept term for the sample as a whole might be insignificant although sub-samples partitioned on size might yield significant intercepts.

interpret a variable as a risk factor, it is standard in the asset-pricing literature to verify that factor loadings with respect to that variable predict returns (e.g., Fama and Macbeth 1973 and Davis, Fama and French 2000). We turn to this test next.

### *3.3 Cross-sectional tests of returns on PIN factor loadings*

#### *3.3.1 PIN factor loadings*

In this section, we investigate whether there is a risk premium for PINF by (i) estimating a cross-sectional regression of returns on factor loadings of PINF; and (ii) assessing the statistical significance of the coefficients of these factor loadings. To do so, we need to estimate factor loadings. At the beginning of each year from 1984–2002, stocks are sorted into three groups based on the market capitalization at the end of the prior year, and within each size groups, three portfolios are formed based on the PINs estimated over the prior year. The following regressions are run at the firm-year level, using five years of lagged monthly returns, ensuring at least 2 years (24 months) data is available:

$$R_i = a_i + b_i(r_m - r_f) + s_i SMB + h_i HML + m_i UMD + p_i PINF + e_i \quad (5)$$

The sample size is reduced to 32,630 as we lose observations for the first two years (1984 and 1985), for which we do not have adequate lagged information. The firm level regression coefficients are winsorized at 1% at 99% using each year's distribution. Table 5, Panel A, reports the average estimates for each of the coefficients and their t-statistics, along with the average adjusted $R^2$.

An examination of the factor loadings reveals several interesting patterns. First, from panel A, we can see that the average factor loading on the PIN factor increases within each size group as PIN increases. For instance, among small firms, the average factor loading increases from -0.832 for low PIN firms to -0.528 for high PIN firms, while for large firms, the average factor loading increases from 0.018 for low PIN firms to 0.207 for high PIN firms. This suggests that factor loadings and characteristics are likely to be correlated necessitating the need to ascertain whether the PIN factor loadings, by themselves, or the PIN characteristic predict returns.

Secondly, the average PIN factor loading for small firms is negative whereas for medium and large firms, the average PIN factor loading is positive, as is obvious from panel A. This pattern, counter-intuitively, suggests that smaller firms have a negative sensitivity to the information risk factor (hedge against information risk!) whereas medium and larger firms are sensitive to information risk. In other words, these results imply that the information risk component in the cost of capital of larger firms is higher than that for smaller firms. One would have expected the opposite as a greater number of informed intermediaries follow large firms, and hence, the related premium for private information ought to be lower.

Finally, the frequency of negative PIN factor loadings, reported in panel B, for each of the nine portfolios, appears unusually high for small firms, ranging from 60.6% in the S/Low PIN group to 54.5% in the S/High PIN group. In contrast, negative PIN factor loadings in Big firms range from 48.7% in B/Low PIN to 41.7% in B/High PIN sample. Even after allowing for the noise in estimating firm-level loadings, such a high frequency of negative loadings for small firms is disconcerting for the interpretation of PIN as a priced risk factor. Recall that the PIN characteristic appears to predict returns robustly only for small firms. Given that result, we would have expected to see a greater frequency of positive loadings in small firms.

*3.3.2 Cross-sectional portfolio tests*

Next, we conduct monthly Fama and MacBeth (1973) cross-sectional regressions of returns on factor loadings and firm characteristics to ascertain whether PIN factor loadings predict returns. It is well known that factor loadings for individual stocks are noisy (Fama and French 1992) and using such noisy factor loadings in the Fama-Macbeth regression will unfairly bias the tests against finding that PIN is a priced risk factor. To increase power, we will run these tests at the portfolio level instead of at the firm level. Khan (2007) discusses the advantages of using portfolios instead of individual firms as the unit of analysis for such cross-sectional tests. These tests are run on several portfolios based on size (market capitalization at the end of the year), PIN and the PIN factor loadings estimated earlier at the firm-level. Specifically, we run the tests on (i) 50 SIZE based portfolios; (ii) 50 PIN characteristic based portfolios; (iii) 100 portfolios based on a combination of SIZE and PIN (10 SIZE x 10 PIN portfolios); (iv) 125 portfolios based on SIZE, PIN and the PIN factor loadings (5 SIZE x 5 PIN x 5 PIN loading portfolios).

For each portfolio, we compute average returns, characteristics and loadings from the firm-level information. The dependent variable is the equally weighted return of the firms in the given portfolio. Portfolio loadings with respect to the market factor $R_m – R_f$, SMB, HML, UMD and PIN are also estimated as the equally weighted average of the firm level loadings of all firms in the portfolio and are labeled LRMRF, LSMB, LHML, LUMD and LPIN respectively.[8] The constraint of needing 24 months of return information and factors including the PIN factor causes us to drop observations for 1983

[8] As an alternative, we estimate loadings at a portfolio level as in Hou and Moskowitz (2005). The correlation between portfolio loadings estimated directly and the portfolio loadings estimated as the average of firm-level loadings (the approach we use) varies from 0.85 to 0.95 across the factors. Our results are virtually identical if we use the portfolio loadings approach.

and 1984. We hence have seventeen years from 1985 to 2001 for which we can generate portfolio loadings. Our cross-sectional regressions for the year-ahead returns thus span the 204 months from January of 1986 to December of 2002.

We first examine whether PIN loadings predict returns after controlling for loadings on all other risk factors. Panel A of Table 6 reports time series averages of the monthly cross-sectional regression coefficients from January of 1986 to December of 2004 and their time series L-R t-statistics using a Fama-Macbeth (1973) procedure. The regression shows that PIN factor loadings are insignificantly related to average returns for every portfolio combination considered. In particular, the most successful L-R t-statistic ever attained on the PIN factor loadings is a counter-intuitive –1.14 for the 50 PIN based portfolios. In sum, PIN factor loadings are unrelated to returns casting doubt on whether PIN is a priced information risk factor.

Turning to the other factor loadings, the coefficient on (i) HML loadings (LHML) is positive and significant in three of the four tests; (ii) SMB loadings (LSMB) are negative in three out of four tests; and (iii) beta (LRMRF) is negative in three of the four portfolios. The negative loading on beta in cross-sectional tests, while inconsistent with rational pricing, is consistent with results in prior literature; for instance Easley et al. (2002) report an L-R t-statistic of -6.22 on beta in their cross-sectional regression (Table VII, page 2214). The negative loading on SMB may be related to the fact that the composition of firms for which PIN data is available tend to be large firms which have a negative loading on SMB by construction.

As mentioned before, one potential econometric explanation for the results in panel A is that because PIN factor loadings and the PIN characteristic are correlated, the omission of PIN characteristic from the above regressions may have biased the coefficient on PIN factor loadings to zero. To investigate this conjecture, panel B supplements the factor loadings with size, B/M and PIN characteristics. In particular, we introduce portfolio averages of PIN characteristics, which are measured as at the fiscal year end of the previous year in the regression along with portfolio averages of SIZE (log of a firm's market capitalization at the end of previous year) and LBM (log of the book-to-market ratio at the fiscal year-end of the previous year) in the regression. To control for the effect of other known predictors of returns, we also introduce R1 (the previous month's return to control for the short-term reversal effect of Jegadeesh 1990), R2_12 (the return from month –12 to month –2 to account for the medium-term momentum effect of Jegadeesh and Titman 1993), and R13_36 (the return from month –36 to month –13 to control for the long-term winner/loser effect of DeBondt and Thaler 1985).

Results reported in panel B show that the coefficient on PIN factor loadings continues to be insignificant. Interestingly, the PIN characteristic does not attain significance either. Thus, the cross-sectional regression test appears to strongly reject the interpretation that PIN captures priced information risk.

## 4. Ex-Ante Cost of Capital

Another potential way to test whether PIN is priced is to examine whether higher PIN is associated with a higher ex-ante cost of capital, derived from analysts' earnings forecasts. Testing whether PIN is associated with ex-ante measures of risk premium seems reasonable in our context for at least three reasons. First, one can potentially argue that monthly cross-sectional portfolio regressions of the type just discussed in Table 6 are potentially harsh asset pricing tests given that none of the coefficients on factor loadings consistently attain statistical significance. Second, prior research has shown that the correlation between expected returns and realized returns is weak (Elton 1999). This has led to attempts to infer the risk premium ex-ante (Claus and Thomas 2001), Gebhardt, Lee, and Swaminathan 2001). Third, prior research (Gode and Mohanram 2003) has shown that ex-ante or implied cost of capital estimates are associated in the expected direction with risk measures such as systematic risk (β +), idiosyncratic risk (+), earnings volatility (+) and more importantly with measures of information environment quality such as size (-), analyst following (-) and forecast dispersion (+).

In the ex-ante approach, the researcher infers the risk premium from the current price and future expected dividends from earnings estimates provided by sell-side analysts. One can view the ex-ante cost of capital as a summary statistic for the total priced risk. If any aspect of risk, such as PIN, is indeed priced, it ought to be positively correlated with the ex-ante measure of risk.

We use the Ohlson & Juettner-Nauroth (2005) model, commonly referred to as the OJ model, to infer the implied cost of capital. We use this model as it parsimoniously provides us with closed form estimates of the implied cost of capital without any assumptions regarding long run industry profitability. The OJ model calculates the cost of capital ($r_e$) as follows

$$\boxed{\begin{array}{l} r_e = A + \sqrt{A^2 + \dfrac{eps_1}{P_0}\left(g_2 - (\gamma - 1)\right)} \\ \text{where } A \equiv \dfrac{1}{2}\left((\gamma - 1) + \dfrac{dps_1}{P_0}\right) \text{ and } g_2 = \dfrac{(eps_2 - eps_1)}{eps_1} \end{array}}$$

where $eps_1$ and $eps_2$ are consensus estimates of one-year-ahead and two-year-ahead annual EPS, $g_2$ consequently is the short-term growth rate in earnings, $dps_1$ is the estimated dividend in the next period assuming historical payout and $\gamma$ is the estimate of the long run economy-wide growth rate. Consistent with the Gode-Mohanram (2003) implementation of the OJ model, we use the average of short-term growth rate ($eps_2/eps_1$ -1) and analyst five-year earnings growth forecasts as our measure of $g_2$. Further, we set set $\gamma-1 = r_f$ -3%, where $r_f$ is the yield on 10-year notes. To ensure comparability across time, we conduct all tests in terms of the risk premium by subtracting out the risk-free rate. We label our measure of risk premium as $RP_{OJ}$.[9]

Of the 39,376 observations for which we have PIN information from 1984 to 2002, we were able to get adequate information on analyst forecasts for 18,086 observations, or around 47% of the sample. Panel A of Table 7 provides descriptive statistics for the sub-sample with valid information, partitioned by size and PIN. Clearly, we lose most of our observations for small firms who are far less likely to have analyst following. For small firms, high PIN is associated with a lower ex-ante cost of capital than low PIN. Small firms with low PIN had an average $RP_{OJ}$ of 10.86% as opposed to an average $RP_{OJ}$ of 9.97% for firms with low PIN. These differences are statistically significant at the 1% level despite the small sample size. Similarly, for medium sized firms, higher PIN is associated with lower ex-ante cost of capital. Large firms are the

[9] As an alternative to the OJ model, we also calculate implied cost of capital based on the simplification of the OJ model that uses the PE to growth ratio or the PEG ratio as the basis of calculation. If one sets g to 1 and assumes a zero dividend payout, the formula simplifies to re = $\sqrt{(g_2/(P_o/eps_1))}$, or in other words, the inverse of the square root of the PEG ratio. Easton and Monahan (2005) amongst others show that this measure is less likely to suffer from issues of measurement error associated with parameter choices in the OJ model. Our results using this approach are virtually identical and are hence not presented.

only subgroup where PIN shows an increasing relationship with ex-ante risk, with low PIN firms having a mean $RP_{OJ}$ of 6.34% as opposed to 6.61% for high PIN firms.

Panel B of Table 7 presents a correlation matrix between $RP_{OJ}$, PIN and other measures of risk studied in the prior literature on ex-ante cost of capital, namely: systematic risk, measured using the same portfolio beta approach used in our earlier tests (BETA); dispersion in analysts forecasts (DISP), measured as the log of the standard deviation of one-year-ahead EPS estimates scaled by price; analyst following, proxied for by the log of the number of analysts issuing one-year-ahead EPS forecasts (LNUM); size, proxied for by the log of market capitalization at year end (SIZE); leverage, measured as the log of 1 + the ratio of long term debt to market capitalization; growth, proxied for by analysts' consensus long term growth forecasts (LTG); and finally, the log of the book-to-market ratio (LBM). Consistent with prior research, $RP_{OJ}$ is positively correlated with factors related to systematic risk (BETA), unsystematic risk (DISP), leverage (LDM), growth (LTG) and book-to-market (LBM), while it is negatively correlated to size (SIZE) and analyst following (LNUM). $RP_{OJ}$ is positively correlated with PIN; however, do note that PIN is highly correlated with size and analyst following. To isolate the relationship between ex-ante cost of capital and PIN, we will run multivariate regressions with $RP_{OJ}$ as the dependent variable with all the control variables listed above.

Panel C reports the results of a pooled regression of ex-ante cost of capital measures on PIN and the control variables. Surprisingly, their appears to be an inverse relationship between PIN and ex-ante risk in that the coefficient on PIN is –0.043 (t-statistic of –5.37). That is, firms with a higher level of PIN have lower ex-ante cost of capital. All the other control variables show a strong relationship with ex-ante cost of capital consistent with prior research.

To ensure that the result is not driven by the strong inverse relationship between PIN and SIZE, we rerun the regression excluding SIZE and run the regressions separately for each size group. When SIZE is removed from the model, PIN does not appear to be associated with implied cost of capital (t-statistic of –1.32 on PIN). When the regressions are repeated over the three SIZE based portfolios, PIN is consistently negatively associated with the implied cost of capital as four of the six coefficients on PIN acquire t-statistics in excess of an absolute value of two. Panel D reports the summary of annual regressions instead of the pooled cross-sectional regression reported in panel C. In no case, does PIN acquire a statistically significant positive coefficient at conventional levels. Clearly, these results are disconcerting to the notion that PIN is priced information risk.

The inconsistent relationship between PIN and ex-ante risk corroborates our earlier results that PIN cannot be considered a reliable source of priced information risk. They also highlight one significant difference between PIN based measures of information risk and the accounting quality (AQ) based measures proposed by Francis et al (2004). A current paper by Core, Guay and Verdi (2007) questions whether accruals quality, an accounting proxy for information risk, is priced. However, Francis et al (2004) demonstrate that AQ based measures are reliably associated with ex-ante cost of capital estimates, unlike our results for PIN.

## 5. Conclusions

Easley et al. (2002, 2004) present a theoretical and an empirical case for why PIN, a measure of private information derived from a market microstructure model, is a priced. The Easley et al. (2002, 2004) papers have been very influential in that empirical researchers in finance and especially accounting have begun to rely extensively on the premise that PIN is a priced measure of information risk.

However, a closer scrutiny reveals that such enthusiasm for the interpretation of PIN as priced information risk might be somewhat premature. First, the Easley et al. (2002) result that PIN is priced appears to be limited to the 1984-1988 time period. Second, the average PIN factor loading for small firms is negative, whereas that for large firms is positive. This finding implies that the information cost component of cost of capital for large firms is greater than for small firms, although the PIN characteristic seems to predict returns only for small, not large stocks. Third, PIN factor loadings do not appear to be related to returns. Finally, there appears to be no robust relationship between PIN and ex-ante measures of cost of capital.

A combined reading of the findings presented here suggests that there is not much evidence to support the interpretation that information risk, proxied by PIN, is a source of priced information risk. Future empirical research might want to be cautious about the premise that information risk represented by PIN is priced. We acknowledge that tests of asset–pricing factors ideally require a long time-series of data and our endeavor is hampered by the availability of PIN data from only 1983 onwards. However, this is yet another reason why empirical research might want to be cautious about interpreting PIN as compensation for information risk.

## References


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

Aslan, H., D. Easley, S. Hvidkjaer and M. O'Hara. 2006. The Determinants of Informed Trading: Implications for Asset Pricing. Working paper, University of Houston, Cornell University and University of Maryland.

Asquith, P., R. Oman and C. Safaya. 2007. Short Sales and Trade Classification Algorithms. Working paper, Massachusetts Institute of Technology.

Botosan, C.A., M.A. Plumlee, and Y. Xie. 2004. The Role of Information Precision in Determining the Cost of Equity Capital. *Review of Accounting Studies* 9(2-3): 233-259.

Brown, S., S.A. Hillegeist, and K. Lo. 2004. Conference Calls and Information Asymmetry. *Journal of Accounting and Economics* 37(3): 343-366.

Brown, S., and S.A. Hillegeist. 2007. How Disclosure Quality Affects the Level of Information Asymmetry. *Review of Accounting Studies* 12(2-3): 443-477.

Claus, J. and J. Thomas. 2001. Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets. *Journal of Finance* 56(5): 1629-1666.

Core, J., W. Guay and R. Verdi. 2007. Is Accruals Quality a Priced Risk Factor? Working paper, University of Pennsylvania.

Chen S, T. Shevlin and Y. Tong. 2007. Does the Pricing of Financial Reporting Quality Change Around Dividend Changes? *Journal of Accounting Research* 45(1): 1-40.

Chen, Q., I. Goldstein and W. Jiang. 2007.Price Informativeness and Investment Sensitivity to Stock Price. *Review of Financial Studies* 20(3): 619-650.

Daniel, K. and S. Titman. 1997. Evidence on the Characteristics of Cross-sectional Variation in Common Stock Returns. *Journal of Finance* 52: 1-33.

Davis, J., E. F. Fama, and K. R. French. 2000. Characteristics, Covariances, and Average Returns: 1929 to 1997. *Journal of Finance* 55: 389–406.

De Bondt, W. and R. Thaler. 1985. Does the Stock Market Overreact? *Journal of Finance* 40: 793-815.

Duarte, J., X. Han, J. Harford and L. Young. 2007. Information Asymmetry, Information Dissemination and the Effect of Regulation FD on the Cost of Capital. *Journal of Financial Economics* (forthcoming).

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

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

Easley, D., S. Hvidkjaer, and M. O'Hara. 2004. Factoring Information into Returns. Working paper, Cornell University.

Easton, P. and S. Monahan. 2005. An Evaluation of Accounting-Based Measures of Expected Returns. *The Accounting Review* 80: 501-538.

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

Ellul, A., and M. Pagano. 2006. IPO Underpricing and After-Market Liquidity. *Review of Financial Studies* 19(2): 381-421.

Elton, E. 1999. Expected Return, Realized Return and Asset Pricing Tests. *Journal of Finance* 54(4): 1199-1220.

Fama, E. 1991. Efficient Markets: II. *Journal of Finance* 46: 1575-1617.

Fama, E.F. and K. R. French. 1992. The Cross-section of Expected Stock Returns. *Journal of Finance* 47: 427–465.

Fama, E.F. and K. R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33: 3-56.

Fama, E.F. and K. R. French. 1996. Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance* 51: 55-84.

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

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

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The Market Pricing of Accruals Quality. *Journal of Accounting and Economics* 39: 295-327.

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

Gode, D. and P. Mohanram. 2003. Inferring the Cost of Capital using the Ohlson-Juettner Model. *Review of Accounting Studies* 8: 399-431

Hilary, G. 2006. Organized Labor and Information Asymmetry in Financial Markets. *Review of Accounting Studies* 11(4): 525-548.

Hirshleifer, D., K. Hou and S.H. Teoh. 2006. The Accrual Anomaly: Risk or Mispricing Working paper, Ohio State University.

Hou, K. and T.J. Moskowitz. 2005. Market Frictions, Price Delay, and the Cross-Section of Expected Returns. *Review of Financial Studies* 18(3): 981-1020.

Huddart, S. and B. Ke. 2007. Information Asymmetry and Cross-sectional Determinants of Insider Trading. *Contemporary Accounting Research* 24(1): 195-232.

Hughes, J., J. Liu, and J. Liu. 2005. Information, Diversification, and Cost of Capital. Working Paper, University of California at Los Angeles.

Jegadeesh, N. 1990. Evidence of Predictable Behavior of Security Returns. *Journal of Finance* 45: 881-898.

Jegadeesh, N. and S. Titman. 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance* 48: 65-91.

Khan, M. 2007. Are Accruals Mispriced? Evidence from Tests of an Intertemporal Capital Asset Pricing Model. *Journal of Accounting and Economics* (forthcoming).

Lambert, R., C. Leuz, and R. Verrecchia. 2006. Accounting Information, Disclosure, and the Cost of Capital. Working paper, University of Pennsylvania.

Litzenberger, R., and K. Ramaswamy. 1979. The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence. *Journal of Financial Economics* 7: 163–196.

Odders-White, E. and M. J. Ready. 2006. Credit Ratings and Stock Liquidity. *Review of Financial Studies* 19: 119 - 157.

Ohlson, J. and B. Juettner-Nauroth. 2005. Expected EPS and EPS Growth as Determinants of Value. *Review of Accounting Studies* 10: 349-365.

Pan, J. and A. M. Poteshman. 2006. The Information in Option Volume for Future Stock Prices. *Review of Financial Studies* 19: 871 - 908.

Penman, S. 1983. The Predictive Content of Earnings Forecasts and Dividends. *Journal of Finance* 38(4): 1181-1199.

Spiegel, M. and X.Wang. 2005. Cross-sectional Variation in Stock Returns. Liquidity and Idiosyncratic Risk. Working paper no. 05-13, Yale University.

**Appendix listing published and forthcoming papers that posit that PIN is priced**

To motivate the claim that several extant papers rely, either explicitly or implicitly, on Easley et al. (2002)'s result that PIN is priced, we identify three sets of papers (published or forthcoming in top tier journals in accounting and finance). The first set of papers explicitly links changes in PIN to changes in cost of capital. The second set of papers asserts that PIN is reflected in stock prices or credit ratings of firms. The third set cites the Easley et al (2002) result that higher PIN is associated with higher cost of capital:

<u>Assertions that changes in PIN directly map into changes in cost of capital</u>

- Duarte, Han, Harford and Young (2007 JFE forthcoming) in their equation (2) on page 21 posit that a firm's cost of capital is a function of the firm's beta, its size, its market-to-book and firm-level PIN characteristic. In equation (4) of their paper on page 25, the authors posit that change in cost of capital after Regulation FD is the loading on the PIN characteristic times change in firm-level PIN pre and post FD.

- Brown, Hillegeist and Lo (2004 JAE, page 18), state: "The *Calls* coefficient in the pooled regression indicates that PINs in one quarter are 0.59 percentage points lower for each conference call held during the prior quarter. This represents a 3.2% decrease relative to the mean PIN of 18.24, which in our view represents a moderate and economically plausible effect on the level of information asymmetry. Combined with the findings in Easley et al. (2002) on the association between PINs and the cost of equity capital, our results suggest that holding a conference call each quarter is associated with a 15 basis point <u>reduction in the annual cost of equity capital</u>." (emphasis added)

- Brown and Hillegeist (RAST 2007, page 460) state: "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 <u>reduction in the cost of capital</u>." (emphasis added)

- Chen, Shevlin and Tong (2007 JAR, page 8) hypothesize that: "Ceteris paribus, firms that initiate or increase existing dividends (decrease dividends) exhibit a decrease (an increase) in the pricing of information risk." One of their proxies for information risk is PIN.

Assertions that PIN is reflected in stock prices or credit ratings

- Chen, Goldstein and Jiang (RFS 2006, page 632) state: "PIN capture private firm-specific information impounded in stock price."

- In Table 6, Pan and Poteshman (RFS 2006, page 900) find that put-call predicts next-day risk-adjusted stock returns better when interacted with PIN. The quote from their paper is as follows: "adding an interaction term with PIN reveals a very interesting result. By itself, the put-call ratio provides markedly lower predictability than before. At the same time, the interaction term with PIN picks up a large degree of predictability."

- Ellul and Pagano (RFS 2006) find that IPO underpricing and PIN are related. On page 414, they argue that " increase in the PIN (from its average level of 0.286 to 0.42) is associated with an increase of 16 percentage points in under pricing."

- In Table 7 of their paper, Odders-White and Ready (2006 RFS, page 142) document that S&P credit ratings are negatively related to the adverse selection component of the PIN measure.

Papers (other than the above) that cite the Easley et al. 2002 result that PIN affects cost of capital

- Hilary (RAST 2006, page 534): "For example, Easley et al. (2002) report that PIN is positively associated with spreads, the cost of capital."

- Francis, Lafond, Olsson and Schipper (2004 TAR, page 971): "Empirical tests of the predicted positive relation between information risk and cost of capital use different characteristics of information risk. For example, Easley et al. (2002) focus on the information asymmetry between informed and uninformed traders, which they operationalize using probability of informed trading (PIN) scores."

- Francis, Lafond, Olsson and Schipper (JAE 2005, page 301): "Easley et al. (2002) find results that are broadly consistent with the prediction that firms with more private information (as measured by PIN scores, a market microstructure measure of informed trading) and less public information have larger expected returns."

- Huddart and Ke (2007 CAR, pages 218, 219):"Easley, Hvidkjaer, and O'Hara (2002) argue that stocks that have a higher probability of information-based trading provide a higher equilibrium return to compensate for the added risk associated with adverse selection. From intra-day trade data for a portion of New York Stock Exchange (NYSE) firm-years that overlap with our sample, they compute a probability of information-based trading (PIN) and present evidence that PIN is a source of risk that is priced by the market."

- Aboody, Hughes and Liu (JAR 2005, page 653): “To be consistent with the idea of diversification in neoclassical asset pricing theory, following Francis et al. [2005] and Easley, Hvidkjaer, and O’Hara [2002], we estimate cost of capital using a factor model based on APT from ex post regressions.”

- Botosan, Plumlee and Xie (RAST, 2004, page 235): “..EHO document a strong positive association between averaged realized returns and PIN.”

- Ecker et al. (2006 TAR, page 750): “In Easley and O’Hara’s (2004) model, the risk premium associated with information uncertainty is a function of private information (which pertains to information asymmetry) and the precision of public and private information.”

## Table 1: Summary Statistics of PIN by year

This table presents the summary statistics on PIN by year and a summary of yearly distributions. P1, P25, P75 and P99 refer to percentiles of the yearly cross-sectional distribution; Std is the standard deviation. SIZE is measured as year-end market capitalization.

| Year | N | Mean | Std | P1 | P25 | Median | P75 | P99 | ρ(SIZE,PIN) |
|---|---|---|---|---|---|---|---|---|---|
| 1983 | 2091 | 0.224 | 0.075 | 0.100 | 0.174 | 0.212 | 0.262 | 0.456 | -0.624 |
| 1984 | 2043 | 0.210 | 0.070 | 0.095 | 0.162 | 0.199 | 0.246 | 0.461 | -0.482 |
| 1985 | 1992 | 0.217 | 0.067 | 0.100 | 0.172 | 0.209 | 0.251 | 0.448 | -0.513 |
| 1986 | 1916 | 0.218 | 0.067 | 0.095 | 0.173 | 0.209 | 0.254 | 0.426 | -0.575 |
| 1987 | 1975 | 0.224 | 0.073 | 0.097 | 0.172 | 0.215 | 0.262 | 0.444 | -0.687 |
| 1988 | 1948 | 0.220 | 0.071 | 0.095 | 0.172 | 0.210 | 0.260 | 0.437 | -0.544 |
| 1989 | 1901 | 0.218 | 0.074 | 0.095 | 0.170 | 0.207 | 0.251 | 0.457 | -0.532 |
| 1990 | 1854 | 0.229 | 0.080 | 0.096 | 0.174 | 0.216 | 0.267 | 0.483 | -0.617 |
| 1991 | 1938 | 0.231 | 0.084 | 0.093 | 0.171 | 0.218 | 0.273 | 0.506 | -0.727 |
| 1992 | 2015 | 0.224 | 0.081 | 0.088 | 0.166 | 0.212 | 0.265 | 0.467 | -0.749 |
| 1993 | 2140 | 0.208 | 0.073 | 0.089 | 0.159 | 0.197 | 0.244 | 0.460 | -0.622 |
| 1994 | 2199 | 0.206 | 0.076 | 0.092 | 0.155 | 0.195 | 0.239 | 0.458 | -0.639 |
| 1995 | 2207 | 0.203 | 0.077 | 0.080 | 0.151 | 0.190 | 0.237 | 0.478 | -0.591 |
| 1996 | 2243 | 0.201 | 0.077 | 0.083 | 0.145 | 0.188 | 0.238 | 0.444 | -0.708 |
| 1997 | 2309 | 0.191 | 0.080 | 0.069 | 0.131 | 0.177 | 0.232 | 0.435 | -0.722 |
| 1998 | 2337 | 0.183 | 0.088 | 0.064 | 0.122 | 0.162 | 0.223 | 0.480 | -0.761 |
| 1999 | 2208 | 0.185 | 0.091 | 0.059 | 0.118 | 0.163 | 0.231 | 0.463 | -0.797 |
| 2000 | 2083 | 0.193 | 0.098 | 0.066 | 0.118 | 0.168 | 0.243 | 0.506 | -0.804 |
| 2001 | 1977 | 0.207 | 0.105 | 0.074 | 0.125 | 0.180 | 0.269 | 0.524 | -0.843 |
| Summary | 2072.4 | 0.210 | 0.079 | 0.086 | 0.154 | 0.196 | 0.250 | 0.465 | -0.660 |

## Table 2: Asset Pricing Tests for the PIN Characteristics

This table presents results from firm-level cross-sectional regressions estimated every month between January 1984 and December 2002 for various time windows using both standard Fama and Macbeth (1973) methodology as well as Litzenberger and Ramaswamy (L-R) (1979) precision weighted means (weighted least squares). The dependent variable is the percentage monthly return (RET). BETA is a portfolio beta based on 40 portfolios using the procedure described in section 2.3. PIN is measured at prior year end. SIZE is the log of market capitalization at prior year end. LBM is the log of the book-to-market ratio at prior year end. Time-series means of monthly regression coefficients are reported with their time-series t-statistics below in parentheses.

| Time Period | Method | Intercept | Beta | PIN | SIZE | LBM | Avg. Adjusted $R^2$ |
|---|---|---|---|---|---|---|---|
| Panel A: Replication of Easley et al. (2002) | | | | | | | |
| 1984-1998 | Fama-Macbeth | 0.718 | -0.438 | 1.638 | 0.107 | 0.192 | 2.78% |
| | | (1.51) | (-1.07) | (2.75) | (1.70) | (2.14) | |
| 1984-1998 | L-R WLS | 0.512 | -0.751 | 1.931 | 0.148 | 0.207 | 2.78% |
| | | (1.19) | (-1.90) | (3.46) | (2.56) | (2.37) | |
| Panel B: Regressions with just PIN in the 1984-1998 time period | | | | | | | |
| 1984-1998 | Fama-Macbeth | 1.01 | | 0.54 | | | 0.44% |
| | | (2.77) | | (0.59) | | | |
| 1984-1998 | L-R WLS | 0.894 | | 0.283 | | | 0.44% |
| | | (2.54) | | (0.34) | | | |
| Panel C: Replication of Easley et al. (2002) over different time periods | | | | | | | |
| 1984-1988 | Fama-Macbeth | 0.817 | -1.517 | 2.670 | 0.224 | 0.294 | 3.5% |
| | | (1.14) | (-1.88) | (2.76) | (2.52) | (1.86) | |
| 1984-1988 | L-R WLS | 0.780 | -1.791 | 2.800 | 0.239 | 0.303 | 3.5% |
| | | (1.15) | (-2.27) | (2.96) | (2.74) | (1.94) | |
| 1989-1993 | Fama-Macbeth | 0.955 | 0.220 | 0.775 | -0.019 | 0.019 | 2.8% |
| | | (0.91) | (0.28) | (0.69) | (-0.14) | (0.11) | |
| 1989-1993 | L-R WLS | 0.433 | -0.219 | 1.418 | 0.071 | -0.01 | 2.8% |
| | | (0.48) | (-0.29) | (1.41) | (0.60) | (-0.07) | |
| 1994-1998 | Fama-Macbeth | 0.381 | -0.017 | 1.469 | 0.117 | 0.263 | 2.0% |
| | | (0.56) | (-0.04) | (1.47) | (1.20) | (1.84) | |
| 1994-1998 | L-R WLS | 0.315 | -0.126 | 1.510 | 0.122 | 0.307 | 2.0% |
| | | (0.48) | (-0.28) | (1.57) | (1.27) | (2.17) | |
| 1999-2002 | Fama-Macbeth | 2.053 | 0.053 | -1.762 | -0.142 | 0.332 | 3.8% |
| | | (1.43) | (0.04) | (-0.76) | (-0.6) | (1.16) | |
| 1999-2002 | L-R WLS | 1.414 | -0.767 | -0.253 | -0.013 | 0.371 | 3.8% |
| | | (1.09) | (-0.69) | (-0.12) | (-0.06) | (1.42) | |

## Table 3: Characteristics and Returns of PIN portfolios based on Sequential Sorts on Size and PIN

At the beginning of each year from 1984–2002, stocks are sorted into three groups based on market capitalization at the end of the prior year, and within each size groups, three portfolios are formed based on the PINs estimated over the prior year. There were 39,376 observations in total, or approximately 4,375 per group based on size-PIN grouping. Panel A contains the time-series average of the yearly value-weighted mean PIN for each portfolio. High-Low is the average return difference between the examined variable in the respective panel between high and low PIN portfolios, and t(High-Low) is the t-statistic of High-Low. Panel B contain the time series average of the average firm size in each portfolio. Panel C contains the time series average of the monthly returns of each portfolio. Returns are weighted by the prior year-end market value.

Panel A: PIN

| | PIN Group | | | | |
|---|---|---|---|---|---|
| Size Group | Low | Medium | High | High - Low | $t_{(High-Low)}$ |
| Small | 0.186 | 0.255 | 0.360 | 0.175 | 142.59 |
| Medium | 0.153 | 0.201 | 0.268 | 0.115 | 128.56 |
| Big | 0.113 | 0.149 | 0.199 | 0.086 | 114.31 |

Panel B: Size

| | PIN Group | | | | |
|---|---|---|---|---|---|
| Size Group | Low | Medium | High | High - Low | $t_{(High-Low)}$ |
| Small | 48.41 | 40.09 | 29.37 | -19.04 | -24.81 |
| Medium | 402.17 | 343.81 | 288.06 | -114.11 | -21.55 |
| Big | 10374 | 3747.76 | 2417.37 | -7956.63 | -27.12 |

Panel C: Returns

| | PIN Group | | | | |
|---|---|---|---|---|---|
| Size Group | Low | Medium | High | High - Low | $t_{(High-Low)}$ |
| Small | 0.514 | 0.608 | 1.094 | 0.580 | 3.23 |
| Medium | 0.964 | 0.990 | 0.984 | 0.020 | 0.18 |
| Big | 1.119 | 1.135 | 1.150 | 0.031 | 0.22 |

## Table 4: Descriptive Statistics and Correlations Among Factors

Panel A contains summary statistics on the Fama-French factor portfolio monthly returns in 1984–2002: market excess return ($R_m - R_f$), small stock returns minus large stock returns (SMB), high book-to-market stock returns minus low book-to-market stock returns (HML), and past 1-year winner stock returns minus past loser stock returns (UMD); and on portfolio returns based on pin-sorted portfolios (PINF). The construction of the PINF portfolio is explained in the text. Panel B contains the time-series correlations between the factor portfolios over the sample period.

Panel A: Summary Statistics

| Factor | Mean | Std. Deviation | t-stat |
|---|---|---|---|
| $R_m - R_f$ | 0.549 | 4.612 | 1.80 |
| SMB | -0.072 | 3.521 | -0.31 |
| HML | 0.356 | 3.402 | 1.58 |
| UMD | 0.994 | 4.514 | 3.33 |
| PINF | 0.239 | 1.606 | 2.25 |

Panel B: Correlations

| | SMB | HML | UMD | PINF |
|---|---|---|---|---|
| $R_m - R_f$ | 0.164 | -0.524 | -0.087 | -0.246 |
| SMB | | -0.452 | 0.108 | -0.090 |
| HML | | | -0.078 | 0.027 |
| UMD | | | | 0.576 |

Panel C: Fama-French Regressions on the Entire Sample

At the beginning of each year from 1984–2002, stocks are sorted into three groups based on the market capitalization at the end of the prior year, and within each size groups, three portfolios are formed based on the PINs estimated over the prior year. There were 39,376 observations in total, or approximately 4,375 per group based on size-PIN grouping. Monthly excess returns for these portfolios are regressed against $r_m - r_f$, SMB, HML , UMD and PINF using the following specification: $R_i = \alpha_i + \beta_i(r_m - r_f) + s_iSMB + h_iHML + m_iUMD + p_iPINF + \varepsilon_i$ The table reports the point estimates for each of the coefficients and their t-statistics in parantheses, along with the adjusted $R^2$. The sample period is 1984–2002.

| $\alpha_i$ | $\beta_i$ | $s_i$ | $h_i$ | $m_i$ | $p_i$ | Adj. $R^2$ |
|---|---|---|---|---|---|---|
| -0.008 | 1.003 | -0.128 | 0.305 | | | 94.3% |
| (-0.12) | (57.71) | (-5.91) | (11.73) | | | |
| 0.026 | 0.998 | -0.125 | 0.300 | -0.029 | | 94.4% |
| (0.36) | (57.18) | (-5.80) | (11.54) | (-1.96) | | |
| -0.036 | 1.013 | -0.123 | 0.313 | | 0.082 | 94.4% |
| (-0.51) | (56.18) | (-5.68) | (11.95) | | (1.87) | |
| 0.005 | 1.015 | -0.110 | 0.312 | -0.069 | 0.199 | 94.7% |
| (0.08) | (57.99) | (-5.14) | (12.27) | (-3.85) | (3.81) | |

## Table 5: Firm Level Regressions to Estimate PIN Loadings

At the beginning of each year from 1984–2002, stocks are sorted into three groups based on the market capitalization at the end of the prior year, and within each size groups, three portfolios are formed based on the PINs estimated over the prior year. The following regressions are run at the firm-year level, using five years of lagged monthly returns, ensuring at least 2 years (24 months) data is available $R_i = \alpha_i + \beta_i(r_m - r_f) + s_i SMB + h_i HML + m_i UMD + p_i PINF + \varepsilon_i$ The sample size is reduced to 32,630 as we lose observations for the first two years (1984 and 1985). The firm level regression coefficients are winsorized at 1% at 99% using each year's distribution. The table reports the average estimates for each of the coefficients and their t-statistics in parantheses, along with the average adjusted $R^2$.

Panel A: Summary of Firm-Level Factor Regressions

| Size Group | PIN Group | $\alpha_i$ | $\beta_i$ | $s_i$ | $h_i$ | $m_i$ | $p_i$ | Adj. $R^2$ |
|---|---|---|---|---|---|---|---|---|
| Small | LOW | -0.383 | 0.987 | 1.188 | 0.217 | -0.149 | -0.832 | 16.2% |
| | | (-9.93) | (81.82) | (60.87) | (10.68) | (-9.36) | (-18.99) | |
| Small | MEDIUM | -0.201 | 0.940 | 1.148 | 0.286 | -0.110 | -0.742 | 15.0% |
| | | (-5.2) | (78.85) | (56.72) | (14.13) | (-6.89) | (-17.1) | |
| Small | HIGH | -0.005 | 0.783 | 0.993 | 0.334 | -0.064 | -0.528 | 12.6% |
| | | (-0.14) | (67.63) | (50.78) | (17.79) | (-4.15) | (-12.82) | |
| Medium | LOW | 0.150 | 1.063 | 0.690 | 0.215 | -0.146 | -0.029 | 23.6% |
| | | (5.45) | (113.61) | (46.7) | (12.83) | (-12.64) | (-1.01) | |
| Medium | MEDIUM | 0.324 | 1.091 | 0.811 | 0.218 | -0.133 | 0.016 | 23.5% |
| | | (10.9) | (113.33) | (52.88) | (12.71) | (-10.92) | (0.53) | |
| Medium | HIGH | 0.492 | 1.014 | 0.841 | 0.313 | -0.055 | 0.068 | 21.8% |
| | | (16.03) | (101.33) | (57.05) | (18.58) | (-4.43) | (2.18) | |
| Big | LOW | 0.309 | 1.022 | -0.069 | 0.131 | -0.053 | 0.018 | 34.9% |
| | | (17.6) | (148.03) | (-6.95) | (11.13) | (-6.36) | (1.01) | |
| Big | MEDIUM | 0.295 | 1.102 | 0.143 | 0.181 | -0.092 | 0.082 | 32.2% |
| | | (14.63) | (146.64) | (13.1) | (14.1) | (-10.17) | (4.03) | |
| Big | HIGH | 0.504 | 1.114 | 0.323 | 0.226 | -0.084 | 0.207 | 28.9% |
| | | (20.42) | (128.75) | (25.6) | (16.1) | (-8.38) | (8.58) | |

Panel B: Analysis of Negative Loadings by Size and PIN groups

| | | Proportion of Loadings that are Negative | | | | |
|---|---|---|---|---|---|---|
| Size Group | PIN Group | $\beta_i$ | $s_i$ | $h_i$ | $m_i$ | $p_i$ |
| Small | LOW | 8.0% | 11.8% | 41.1% | 56.4% | 60.6% |
| Small | MEDIUM | 8.1% | 11.9% | 38.3% | 54.9% | 59.1% |
| Small | HIGH | 11.1% | 14.3% | 35.4% | 53.0% | 54.5% |
| Medium | LOW | 2.3% | 19.4% | 36.4% | 58.9% | 48.7% |
| Medium | MEDIUM | 2.3% | 15.1% | 37.2% | 59.0% | 46.4% |
| Medium | HIGH | 3.1% | 11.6% | 33.7% | 53.9% | 45.5% |
| Big | LOW | 0.5% | 59.3% | 38.2% | 51.8% | 48.7% |
| Big | MEDIUM | 0.4% | 42.4% | 37.1% | 55.6% | 47.2% |
| Big | HIGH | 1.3% | 32.3% | 35.8% | 56.1% | 41.7% |
| AVERAGE ACROSS GROUPS | | 4.0% | 24.7% | 37.1% | 55.5% | 50.2% |

## Table 6: Monthly Cross-Sectional Portfolio Regressions of Returns on Factor Loadings

As a first step, for each year, we run firm level regressions of monthly returns on the market (Rm – Rf) and the risk factors for size (SMB), book to market (HML), momentum (UMD) as well as PIN (PINF). The respective loadings are labeled LRMRF, LSMB, LHML, LUMD and LPIN. Portfolios are formed using a combination of large number of parameters – size, PIN and PIN Loadings. Portfolio factor loadings are calculated as the average of firm level loadings estimated using at least 24 and up to 60 prior months of prior returns. Factor loadings are winsorized at 1% and 99% to reduce the effects of outliers. The dependent variable in the regressions presented below is the average return for the portfolio in the year after portfolio formation. Cross-sectional regressions are estimated every month between January 1986 and December 2002 (204 months). We calculate time-series means of monthly regression coefficients as in Fama-Macbeth (1973). Mean coefficients and t-statistics are calculated using the precision of coefficients from regressions as weights, using the procedure from Litzenberger and Ramaswamy (L-R) (1979). In Panel B, we also include controls for the following characteristics in addition to the factor loadings: SIZE (log of market capitalization at prior year end), LBM (the log of the book-to-market ratio at prior year end), R1 (previous month's return), R2_12 (return from month –12 to month –2), R13_36 (return from month –36 to month –13) and PIN measured at prior year end. Similar to the loadings, characteristics are averaged across the portfolio being used.

Panel A: Factor Loadings Alone

| Portfolio | Intercept | LRMRF | LSMB | LHML | LUMD | LPIN | Adj. $R^2$ |
|---|---|---|---|---|---|---|---|
| 50 Size | 1.450 | -0.625 | -0.337 | 0.385 | -0.211 | -0.014 | 24.2% |
| | (3.57) | (-1.59) | (-1.98) | (1.93) | (-0.74) | (-0.09) | |
| 50 PIN | 1.790 | -0.988 | -0.262 | 0.157 | -0.565 | -0.125 | 16.9% |
| | (5.61) | (-2.68) | (-1.50) | (0.80) | (-2.41) | (-1.14) | |
| 10 Size x 10 PIN | 1.389 | -0.555 | -0.308 | 0.315 | -0.079 | -0.046 | 16.7% |
| | (5.63) | (-1.86) | (-1.80) | (1.73) | (-0.36) | (-0.40) | |
| 5 Size x 5 PIN x 5 PIN Loading | 1.204 | -0.318 | -0.294 | 0.331 | 0.005 | -0.032 | 15.1% |
| | (4.64) | (-1.04) | (-1.73) | (1.84) | (0.02) | (-0.67) | |

## Table 6: Monthly Cross-Sectional Portfolio Regressions of Returns on Factor Loadings (Cont'd)

Panel B: Factor Loadings and Characteristics

| Portfolio | Intercept | LRMRF | LSMB | LHML | LUMD | LPIN | SIZE | LBM | R1 | R2_12 | R13_36 | PIN | Adj. $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 Size | 2.146 | -0.250 | -0.345 | 0.076 | -0.173 | 0.072 | -0.097 | 0.231 | 13.844 | 2.973 | -0.288 | -3.233 | 34.9% |
| | (1.94) | (-0.80) | (-1.66) | (0.34) | (-0.63) | (0.59) | (-1.13) | (0.99) | (6.22) | (6.71) | (-1.63) | (-1.14) | |
| 50 PIN | 1.229 | -0.218 | -0.298 | -0.024 | -0.304 | -0.136 | 0.007 | 0.483 | 10.969 | 2.958 | -0.004 | -0.303 | 23.17% |
| | (1.30) | (-0.76) | (-1.61) | (-0.11) | (-1.15) | (-1.14) | (0.07) | (2.24) | (5.01) | (7.80) | (-0.02) | (-0.23) | |
| 10 Size x 10 PIN | 1.103 | 0.120 | -0.310 | -0.019 | -0.073 | -0.105 | -0.040 | 0.352 | 12.940 | 2.985 | -0.084 | -0.395 | 26.6% |
| | (2.38) | (0.55) | (-1.97) | (-0.11) | (-0.38) | (-1.15) | (-0.63) | (1.99) | (6.28) | (8.48) | (-0.67) | (-0.57) | |
| 5 Size x 5 PIN | 0.985 | 0.044 | -0.169 | 0.089 | -0.136 | -0.028 | -0.013 | 0.306 | 12.846 | 3.166 | -0.132 | -0.928 | 23.9% |
| x 5 PIN Loading | (2.15) | (0.21) | (-1.22) | (0.60) | (-0.75) | (-0.63) | (-0.22) | (2.25) | (6.29) | (10.17) | (-1.16) | (-1.39) | |

## Table 7: PIN and Implied Cost of Capital

At the beginning of each year from 1984–2002, stocks are sorted into three groups based on market capitalization at the end of the prior year, and within each size groups, three portfolios are formed based on the PINs estimated over the prior year. Implied cost of capital estimates are calculated using stock prices and earnings forecasts as of the end of the previous year, based on the Ohlson & Juettner-Nauroth (2005) OJ model as operationalized by Gode and Mohanram (2003). Risk Premia, $RP_{OJ}$, are calculated from implied cost of capital estimates by subtracting out the risk free rate. Panel A presents mean $RP_{OJ}$ by size and PIN groups. t-statistics for differences are in parentheses and are calculated using pooled estimate of standard error. Panel B presents correlations between $RP_{OJ}$, PIN and the following control variables: BETA, a portfolio beta based on 40 portfolios using the procedure described in section 2.3; DISP is the dispersion of analyst forecast errors defined as standard deviation of one year ahead forecasts scaled by stock price; LNUM is the log of the number of analysts issuing forecasts; SIZE is log of market capitalization at the end of the year; LDM is the log of the ratio of long term debt to market capitalization; LTG is the consensus estimate of the long term growth rate; and LBM is the log of the book-to-market ratio. Panel C regresses $RP_{OJ}$ on the control variables and PIN in a pooled regression. Panel D provides summary of annual regression using the Fama-Macbeth (1973) procedure. The t-statistics are derived from the distribution of annual coefficients after controlling for auto-correlation using the procedure in Bernard (1995).

### Panel A: Mean $RP_{OJ}$ by Size and PIN Groups

| Size Group | PIN Group | N | $RP_{OJ}$ | PIN |
|---|---|---|---|---|
| Small | Low | 653 | 10.86% | 18.41% |
| Small | Medium | 529 | 9.88% | 25.12% |
| Small | High | 272 | 9.97% | 32.77% |
| | High – Low | | -0.89% | 14.36% |
| | | | (-2.74) | (48.78) |
| Medium | Low | 2538 | 8.33% | 15.04% |
| Medium | Medium | 2520 | 8.14% | 19.81% |
| Medium | High | 2158 | 7.90% | 25.74% |
| | High – Low | | -0.42% | 10.71% |
| | | | (-3.36) | (100.13) |
| Large | Low | 3114 | 6.34% | 11.05% |
| Large | Medium | 3191 | 6.77% | 14.66% |
| Large | High | 3111 | 6.61% | 19.50% |
| | High – Low | | 0.26% | 8.45% |
| | | | (2.57) | (98.24) |

### Panel B: Correlations (above Diagonal Pearson, below Diagonal Spearman)

| | $RP_{OJ}$ | BETA | DISP | LNUM | SIZE | LDM | LTG | LBM | PIN |
|---|---|---|---|---|---|---|---|---|---|
| $RP_{OJ}$ | 1.000 | 0.141 | 0.393 | -0.187 | -0.306 | 0.305 | 0.047 | 0.335 | 0.152 |
| BETA | 0.146 | 1.000 | 0.077 | 0.072 | -0.039 | -0.005 | 0.281 | -0.010 | 0.041 |
| DISP | 0.418 | 0.103 | 1.000 | -0.118 | -0.280 | 0.240 | -0.050 | 0.304 | 0.178 |
| LNUM | -0.189 | 0.076 | -0.111 | 1.000 | 0.729 | -0.091 | -0.120 | -0.213 | -0.458 |
| LSIZE | -0.300 | -0.037 | -0.380 | 0.743 | 1.000 | -0.166 | -0.136 | -0.442 | -0.643 |
| LDM | 0.292 | -0.060 | 0.307 | -0.061 | -0.138 | 1.000 | -0.184 | 0.448 | 0.037 |
| LTG | 0.040 | 0.299 | -0.160 | -0.117 | -0.145 | -0.287 | 1.000 | -0.280 | 0.128 |
| LBM | 0.354 | -0.012 | 0.507 | -0.214 | -0.427 | 0.508 | -0.313 | 1.000 | 0.234 |
| PIN | 0.150 | 0.057 | 0.297 | -0.461 | -0.663 | 0.016 | 0.132 | 0.244 | 1.000 |

## Table 7: PIN and Implied Cost of Capital (Cont'd)

Panel C: Pooled Regression of $RP_{OJ}$ on PIN and Control Variables

| Sample | Intercept | BETA | DISP | LNUM | LSIZE | LDM | LTG | LBM | PIN | Adj. $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Entire Sample N = 14060 | 0.084 | 0.009 | 1.490 | -0.002 | -0.004 | 0.023 | 0.074 | 0.011 | -0.043 | 26.4% |
| | (24.08) | (12.47) | (34.96) | (-3.17) | (-9.00) | (20.14) | (11.22) | (15.96) | (-5.37) | |
| | 0.060 | 0.010 | 1.551 | -0.006 | | 0.023 | 0.084 | 0.013 | -0.009 | 26.0% |
| | (26.08) | (12.93) | (36.74) | (-11.70) | | (19.87) | (12.93) | (21.00) | (-1.32) | |
| Small Firms N = 832 | 0.117 | 0.009 | 0.839 | -0.014 | -0.002 | 0.026 | 0.081 | 0.012 | -0.083 | 18.3% |
| | (5.43) | (2.58) | (6.05) | (-2.95) | (-0.64) | (5.99) | (2.59) | (3.43) | (-2.31) | |
| | 0.105 | 0.009 | 0.858 | -0.015 | | 0.026 | 0.081 | 0.013 | -0.077 | 18.2% |
| | (8.85) | (2.72) | (6.34) | (-3.16) | | (6.01) | (2.6) | (3.72) | (-2.22) | |
| Medium Firms N=5480 | 0.073 | 0.009 | 1.332 | 0.001 | -0.003 | 0.026 | 0.065 | 0.009 | -0.025 | 19.8% |
| | (9.50) | (7.78) | (22.21) | (0.50) | (-3.06) | (13.92) | (6.48) | (7.33) | (-1.98) | |
| | 0.052 | 0.010 | 1.368 | -0.000 | | 0.026 | 0.068 | 0.010 | -0.011 | 19.7% |
| | (14.39) | (8.35) | (23.25) | (-0.42) | | (13.83) | (6.81) | (8.71) | (-0.90) | |
| Large Firms N = 7748 | 0.051 | 0.011 | 2.673 | -0.001 | -0.000 | 0.015 | 0.066 | 0.010 | -0.061 | 24.1% |
| | (10.32) | (10.87) | (31.32) | (-1.45) | (-0.28) | (10.16) | (7.39) | (13.06) | (-5.63) | |
| | 0.050 | 0.011 | 2.676 | -0.001 | | 0.015 | 0.066 | 0.010 | -0.059 | 24.1% |
| | (16.09) | (10.87) | (31.56) | (-1.80) | | (10.19) | (7.45) | (14.02) | (-6.35) | |

Panel D: Summary of Annual Regressions of $RP_{OJ}$ on PIN and Control Variables

| Sample | Intercept | BETA | DISP | LNUM | LSIZE | LDM | LTG | LBM | PIN | Adj. $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Entire Sample N = 14060 | 0.073 | 0.010 | 2.172 | 0.001 | -0.004 | 0.019 | 0.085 | 0.012 | -0.019 | 29.7% |
| | (5.82) | (6.12) | (9.00) | (0.80) | (-2.86) | (13.29) | (3.23) | (4.56) | (-1.77) | |
| | 0.048 | 0.010 | 2.249 | -0.004 | | 0.019 | 0.097 | 0.015 | 0.002 | 28.9% |
| | (10.67) | (5.57) | (9.09) | (-9.80) | | (12.93) | (4.34) | (7.21) | (0.17) | |
| Small Firms N = 832 | 0.123 | -0.002 | 1.291 | -0.005 | -0.011 | 0.026 | 0.157 | 0.010 | 0.010 | 22.8% |
| | (2.51) | (-0.18) | (4.21) | (-0.69) | (-1.12) | (2.72) | (3.11) | (1.72) | (0.17) | |
| | 0.073 | -0.003 | 1.469 | -0.010 | | 0.023 | 0.158 | 0.011 | 0.033 | 20.5% |
| | (3.68) | (-0.31) | (3.89) | (-2.13) | | (2.93) | (3.51) | (2.42) | (0.62) | |
| Medium Firms N=5480 | 0.090 | 0.011 | 2.395 | 0.002 | -0.008 | 0.021 | 0.068 | 0.008 | -0.012 | 24.0% |
| | (6.60) | (4.27) | (6.46) | (1.48) | (-6.19) | (12.14) | (3.66) | (2.31) | (-0.49) | |
| | 0.043 | 0.011 | 2.469 | -0.001 | | 0.021 | 0.078 | 0.011 | 0.005 | 23.6% |
| | (6.91) | (4.06) | (6.31) | (-0.79) | | (12.50) | (4.36) | (3.23) | (0.28) | |
| Large Firms N = 7748 | 0.039 | 0.010 | 3.393 | 0.002 | -0.001 | 0.013 | 0.103 | 0.013 | -0.018 | 27.8% |
| | (3.81) | (8.68) | (8.04) | (0.87) | (-0.66) | (7.84) | (2.12) | (2.69) | (-1.08) | |
| | 0.033 | 0.009 | 3.417 | 0.001 | | 0.012 | 0.105 | 0.014 | -0.013 | 28.2% |
| | (9.64) | (7.56) | (8.23) | (0.70) | | (7.14) | (2.27) | (3.43) | (-0.66) | |