Previous version: August 2006
This version: November 2006

Is Information Risk (PIN) Priced?

Partha Mohanram
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
Email: rajgopal@u.washington.edu

Abstract:

Several recent papers assume that private information (PIN), proposed by Easley, Hvidkjaer and O'Hara (2002, 2004), is a priced risk factor. In this paper, we formally test whether PIN is indeed a priced risk factor. We first replicate the 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. We divide the sample each year into three groups based on size, and further into three sub-groups based on PIN within each size group, and construct a PIN factor using a methodology similar to Easley, Hvidkjaer and O'Hara (2004) that goes long on high PIN stocks and short on low PIN stocks. We investigate the properties of the PIN factor by adding this factor to standard Fama-French 3-factor and 4-factor models tests of one-year-ahead monthly returns. We find that the average PIN factor loading for small firms is negative whereas loadings for large firms are positive, indicating counter-intuitively that the information risk component of the cost of capital is lower for small firms. Further, we find a strong correlation between PIN and loading on the PIN factor, necessitating the need for tests that isolate the impact of PIN loading after controlling for PIN characteristics. Following Daniel and Titman (1997), we further divide the sample into three sub-groups based on PIN loading estimated using past data and compare the associations of high-loading and low-loading firms. We find no evidence that PIN loadings predict returns after controlling for the PIN characteristic. In addition to the portfolio tests, we find similar results hold in cross-sectional tests. Overall, our findings cast doubt on whether PIN is a priced risk factor.

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

Is Information Risk (PIN) Priced?

Abstract:

Several recent papers assume that private information (PIN), proposed by Easley, Hvidkjaer and O'Hara (2002, 2004), is a priced risk factor. In this paper, we formally test whether PIN is indeed a priced risk factor. We first replicate the 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. We divide the sample each year into three groups based on size, and further into three sub-groups based on PIN within each size group, and construct a PIN factor using a methodology similar to Easley, Hvidkjaer and O'Hara (2004) that goes long on high PIN stocks and short on low PIN stocks. We investigate the properties of the PIN factor by adding this factor to standard Fama-French 3-factor and 4-factor models tests of one-year-ahead monthly returns. We find that the average PIN factor loading for small firms is negative whereas loadings for large firms are positive, indicating counter-intuitively that the information risk component of the cost of capital is lower for small firms. Further, we find a strong correlation between PIN and loading on the PIN factor, necessitating the need for tests that isolate the impact of PIN loading after controlling for PIN characteristics. Following Daniel and Titman (1997), we further divide the sample into three sub-groups based on PIN loading estimated using past data and compare the associations of high-loading and low-loading firms. We find no evidence that PIN loadings predict returns after controlling for the PIN characteristic. In addition to the portfolio tests, we find similar results hold in cross-sectional tests. Overall, our findings cast doubt on whether PIN is a priced risk factor.

Keywords: Information Risk, PIN, Risk, Mispricing

JEL Classification: G12, G14, M40

Is Information Risk (PIN) Priced?

1. Introduction

An influential set of recent papers by Easley, Hvidkjaer and O'Hara (2002, 2004) suggests that a risk factor 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 recently published papers in the finance and accounting literatures cite this interpretation (e.g., Bhattacharya and Daouk 2002, Amihud 2002, Grullon et al. 2003, Botosan et al. 2004, Brown et al. 2004, 2006, Aboody et al. 2005, Bushman et al. 2005, Francis et al. 2005, Hou and Moskowitz 2005, Ellul and Pagano 2006, and Vega 2006). However, whether PIN is really a priced risk factor is debatable. In particular, Spiegel and Wang (2005, footnote 6) suggest that PIN captures a stock's liquidity characteristics and whether liquidity is systematic risk is unclear.

We investigate whether PIN is indeed a priced risk factor. We begin by probing deeper into the properties of the PIN factor and the PIN characteristic. Because size and PIN are highly correlated, we form nine portfolios sorted first on size and then sequentially on PIN at the outset to conduct our tests.[1] Consistent with Easley et al. (2004), we find that the PIN characteristic predicts future returns for small stocks.

However, several observations from a deeper investigation of PIN and the PIN factor cast doubt on whether PIN is really a priced risk factor. First, the average PIN factor loading for small firms is negative while the corresponding loadings for large firms are positive. In particular, approximately 60% of the PIN factor loadings for small firms are negative relative to roughly 40% of corresponding loadings for large firms. This counter-intuitively suggests that the information risk component of cost of capital for large firms with the highest PIN is lower than that for small firms with the highest PIN. Recall, in contrast, that the PIN characteristic predicts future returns only for small firms. Hence, we would have expected to see more positive PIN factor loadings in small firms. Second, when we decompose the source of power in the construction of the PIN factor, we find that firms with the *lowest* PIN drive the average PIN factor loading for small firms while firms with the *highest* PIN drive the average PIN factor loading for large firms. We would have expected firms with the *highest* PIN to drive the average factor loading for small firms, especially given that return predictability for PIN is strongest among small firms with the highest PIN.

Next, we conduct formal tests for whether the PIN factor reflects rational risk premium associated with the information risk factor or whether the returns merely reflect firms that have similar characteristics such as transaction costs that may be correlated with PIN. To do so, we rely on an approach developed by Daniel and Titman (1997) to examine whether risk (covariance) or mispricing (characteristics) explains the size and

[1] We restrict the sorts to only three portfolios because we conduct triple sorts later on size, PIN and PIN factor loadings. Triple sorts, based on three, as opposed to say five, groups lead to a manageable 27 (3^3) sub-portfolios as opposed to an unwieldy 125 (5^3) sub-portfolios.

book-to-market effects in average returns. That is, we investigate whether high (low) returns to greater (smaller) information risk can be attributed to the PIN factor loadings.

Because the PIN factor is constructed from firm-specific PIN measures, the constructed risk measures (the PIN factor loadings) and the original characteristic (PIN) are likely highly correlated. Thus, finding a successful PIN factor that explains returns, as in Easley et al (2004), is a necessary but not sufficient condition for rational risk pricing of PIN. To distinguish risk from mispricing explanations for PIN factor, it is therefore essential to test whether variation in factor loadings that is unrelated to the PIN characteristic still predicts returns. Following Daniel and Titman (1997), Davis, Fama and French (2000), and Daniel, Titman and Wei (2001), we test whether characteristics associated with PIN or whether covariances of high (low) PIN stocks with other high (low) PIN stocks are priced.

To distinguish between loadings and characteristics effect on returns, we sort stocks based on both the level of PIN and the level of loadings on the PIN factor. This allows us to test whether, after controlling for the firm characteristic (PIN), a higher level of risk (PIN factor loading) is associated with higher average returns. We find that this is not the case. Specifically, when the PIN characteristic is held constant, increasing the PIN loading has no impact on average returns. If the PIN factor were a priced risk factor, we would expect increases in PIN factor loadings, holding the PIN characteristic constant, to be positively associated with stock returns.

To buttress these results, we follow Hirschleifer, Hou and Teoh (2006) and conduct tests of risk versus (mis) pricing using Fama and MacBeth (1973) cross-sectional regressions of returns on PIN, PIN loadings, and other average return predictors. The cross-sectional regression approach allows us to (i) integrate portfolio results in a parsimonious manner; and (ii) employ individual stocks in the asset pricing tests without imposing portfolio breakpoints; and (iii) introduce a richer set of asset pricing controls such as the CAPM beta, characteristics and factor loadings for size and book-to-market and lagged stock returns. We find that neither the PIN characteristics, nor the PIN factor loadings, exhibits strong statistical associations with returns after the introduction of such control variables under the cross-sectional regression approach. In sum, 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). Section 3 constructs a PIN factor in line with Easley et al. (2004), and probes its properties. Section 4 presents covariance versus characteristics tests inspired by Daniel and Titman (1997) in a portfolio form and in a cross-sectional regression approach. Section 5 concludes.

2. Return premium to the PIN characteristic

2.1 PIN: theory and estimation

The theoretical intuition for why PIN ought to be priced is derived in Easley and O'Hara (2001) and Easley et al (2002). In particular, Easley et al. (2002, 2004) use a structural 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, the market requires a greater expected excess return. When information is private rather than public and uninformed investors cannot perfectly infer such private information from prices, they view the asset as being more risky. Uninformed investors could avoid this risk, but they choose not to do so. To completely avoid this risk, uninformed traders would have to hold only the risk free asset. However, holding a risk free asset is not optimal because uninformed investors get higher expected utility from 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.

In a series of papers, Easley et al demonstrate how such models can be estimated using trade data to determine the probability of information-based trading, or PIN, for specific stocks. 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:

$$PIN = \frac{\alpha\mu}{\alpha\gamma + \varepsilon_s + \varepsilon_B} \tag{1}$$

where α is the probability that there is new information at the beginning of the trading day, μ is arrival rate of orders from informed traders, ε_S is arrival rate of orders from uninformed sellers and ε_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. (2004) estimate PIN for specific stocks using maximum likelihood estimation with trade and quote data for NYSE and AMEX stocks.

2.2 PIN characteristic

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 the New York Stock Exchange (NYSE) and the American Stock Exchange (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).

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 cut-off points 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 ρ = -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

[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.

correlated with analyst following, institutional ownership, share turnover and Tobin's q; and (ii) positively correlated with smaller firms, ROA, stock return volatility.

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

As mentioned before, our dataset and tests rely heavily on the working paper, Easley et al. (2004) which is based on PIN data from 1983-2001. However, 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 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 to, 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 given as the sum of the two coefficients. Next, 40 portfolios are sorted every 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, β_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 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*, γ_{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 in 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 correction). Both the magnitude of the mean coefficient on PIN as well as the level of significance are 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.

Panel B extends the sample period to cover PIN from 1983-2001 and hence returns for 1984-2002. It is interesting to note that the statistical significance related to the pricing of PIN is significantly weaker in the extended time period. In particular, the t-statistic related to PIN is 1.36 under the Fama-MacBeth methodology and 1.97 under the Litzenberger-Ramaswamy correction.

Panels C and D show that in both time periods, PIN, by itself, is not related to returns as the maximum t-statistic on PIN in these panels is only 0.91. In fact, PIN appears to load only when accompanied by other variables, especially SIZE, in panels A and B. Thus, the pricing of PIN is not robust to an extended time period and specification changes.

2.4 PIN-size portfolios-independent sorts

We begin by verifying that portfolios sorted on PIN earn differential returns. An important methodological issue deserves mention here. Traditionally, the asset-pricing literature has relied on independent sorts of the variables whose ability to predict returns is being tested (Fama and French 1993, 1996, Daniel and Titman 1997, Fama French and Davis 2000 and Hirschleifer, Hou and Teoh 2006). In keeping with this tradition, at the outset, we independently sort stocks on the3 basis of size and PIN. That is, 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)* and independently sort stocks into three equal-sized groups based on PINs estimated in the prior year (*t-1*). Panel A of Table 3 reports the resultant number of observations from such an independent sort into nine portfolios formed by the intersection of the above sorts. It is immediately obvious that most of the observations are concentrated in the off-diagonal cells on account of the high negative correlation between size and PIN (i.e. Small Size, High PIN and Big Size, Small PIN). As a result, the extreme diagonal cells have very few observations (i.e. Small Size, Low PIN and Big Size, High PIN). The smallest diagonal portfolio, the Big Size, High

PIN group, is left with only 770 firm-years. Recall that we would have to further sort these firm-years into three groups later in the paper based on factor loadings to conduct tests based on Daniel and Titman (1997). Thus, another sort on 770 firm-years will leave only about 256 firm-years in each of the three sub-portfolios or approximately 14 firms on average (256 firms spread over 19 years of PIN data from 1983-2001).

Panel C presents the mean stock returns for the sample sorted into 9 size-PIN groups based on independent sorts. For small firms, high PIN firms do earn higher returns than low PIN firms, with a return difference of 0.288%.[3] However, this difference is statistically insignificant, as is the difference for medium and big firms as well. The small number of observations in the extreme diagonal cells is potentially responsible for the inability of PIN to predict stock returns when independent sorts are considered. In particular, note that the spread in returns between Low PIN and High PIN groups is statistically insignificant in every size partition in panel C of Table 3. To avoid unfairly penalizing the ability of PIN to predict returns, due to its high correlation with size, we depart from tradition and rely instead on sequential sorts on PIN within a size group. This approach is also consistent with the way Easley et al. (2004) construct the PIN factor in their paper. We hasten to add that we have replicated all the tables in the paper using independent sorts of PIN and size. The fundamental inferences from such tables (related to whether or not PIN is robustly able predict returns or behave like a risk factor) remain similar to the ones reported in the paper based on sequential sorts.[4]

[3] This is consistent with the results shown in Easley et al. (2002) (Table III, Panel A). Using five groups, they find a return difference of 0.326% for the smallest size quintile and 0.281% for the next smallest quintile in 1984-1998. This corresponds to the 0.288% we find for the smallest tercile in 1984-2001. When we restrict our sample to 1984-1998, we find a difference of 0.310%. As Easley et al. (2002) do not report t-statistics for return differences, we cannot comment on whether the significance levels are similar.

[4] These tables are available on request from the authors, should readers be interested.

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. 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. 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 have roughly 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 4 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 4 reveals 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 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, 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 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 *Exret* consistent with the hypothesis that more information asymmetry (higher PIN) is associated with greater expected returns. In particular, S/Low PIN portfolio earns excess 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 4 is that we observe economically significant abnormal return to PIN only among small stocks.

3. Constructing the PIN Factor and Examining its Properties

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.[5]

The descriptive statistics reported in Table 5 related to the PIN factor and the other factors closely resemble those reported in Easley et al. (2004). The correlation table shows that the PIN factor returns exhibit modest correlation with SMB and the HML factor returns (ρ = -0.09 and 0.027 respectively) but reports a strong correlation with the momentum factor UMD (ρ = 0.576). At first blush, the low correlation between SMB and PIN factor appears 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. [6] 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.

[5] 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.

[6] 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.

3.2 Does the PIN Factor Load for the Entire Sample?

In panel A of Table 6, we investigate whether the PIN factor explains returns for the sample as a whole. 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:

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

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

The first row of panel A 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. Row 2 shows that the momentum factor loads weakly (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]

As an aside, it is worth thinking about why the PIN factor attains 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 5, 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.[8] 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. The latter point is clear from the results presented in panel B of Table 6.

3.3 Loadings on the PIN Factor by Size Groupings

In panel B of Table 6, we decompose the entire sample into three size groups and report the results of estimating the Fama-French three-factor model with UMD and PIN factors for each of these size groups. The first row shows that intercept terms are significant and negative in for the S and M size groups with the three-factor model (t-statistic of –2.75 and –2.12 respectively). Thus, there is room for improving on the three-factor model in two of the three size groups. Consistent with this intuition, Row 2 shows

[7] 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.

[8] As an aside, we have repeated all the tests in the paper without the UMD factor. Untabulated inferences after the omission of the UMD factor are broadly similar to reported results in the text.

that the UMD momentum factor produces statistically significant loadings in only the S and M size groups and not in the B group (t-statistic = -1.4). Row 3 finds that the PIN factor attains a significant loading in the S group and surprisingly, in the B group.

Row 4 estimates a five-factor model with both UMD and the PIN factor and shows that the PIN factor loadings are statistically significant in all three size groups. Note that the PIN factor loadings for the three size groups, S, M, and B are –0.372, 0.275, and 0.192 respectively. 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.

3.4 Loadings on the PIN Factor by Size/PIN Groupings

We seek to understand the properties of the PIN factor in greater depth in this section. In particular, we sort each size group shown in panel A of Table 7 into three sub groups based on the PIN and report the results of estimating equation (2) for these nine sub groups (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) in panel A of Table 7. The last row in each size group reports the returns to a zero-investment portfolio where we go long in the High PIN group and short in the Low PIN group within a size partition.

The loadings on the PIN factor in each of the three size-based zero-investment PIN portfolios decrease with size (1.083, 0.567 and 0.441), suggesting intuitively that small firms are more sensitive to information risk than large firms. However, a closer look at the source of the PIN factor loadings reveals some anomalous patterns.

First, the PIN factor loading of 1.083 on the zero-investment portfolio of the smallest firms is driven by the negative loading of –0.86 on small firms with the lowest PIN (t-statistic = -6.74). Small firms with the highest PIN are sensitive to the PIN factor but the statistical significance of such sensitivity is modest (loading = 0.222, t-statistic = 1.91). This is especially surprising considering how small these firms really are (average market capitalization of such firms is only $29.37 million, see Table 4). In contrast, the PIN factor loadings of 0.567 and 0.441 on zero-investment PIN portfolios for medium (M) and large firms are driven by firms with the highest PIN. That is, M/High PIN group has a PIN factor loading of 0.608 (t-statistic = 6.53) and B/High PIN group has a PIN factor loading of 0.535 (t-statistic = 7.35) whereas the M/Low PIN and the B/High PIN groups report statistically insignificant PIN factor loadings (t-statistic = 0.42 and 1.45 respectively). In sum, one would have expected firms with the highest PIN to drive the average factor loading for small firms, especially given that return predictability for PIN is strongest among small firms with the highest PIN.

Second, the table again, counter-intuitively, suggests that the information risk component of the cost of capital is greater for larger firms than smaller firms. In particular, compare the average PIN factor loading for the B/High PIN portfolio (loading = 0.535, t-statistic = 7.35) with that of the S/High PIN portfolio (loading = 0.222, t-statistic = 1.91). This leads to a difference of 0.313 that is significant at the 5% level. Similarly, compare the average PIN factor loading for the B/Medium PIN portfolio (loading = 0.234, t-statistic = 3.63) with that of the S/Medium PIN portfolio (loading = -

0.154, t-statistic = -1.18). These data imply that the information risk component of cost of capital for large firms with the highest PIN is higher than that for small firms with the highest PIN. This observation, *prima facie*, seems to be odds with the generally accepted idea that stock prices of larger firms are informationally more efficient than stock prices of smaller firms. To put this statistic in perspective, recall from Table 4 that the average market capitalization of large firms with the highest PIN is $2.417 billion whereas the corresponding market capitalization of small firms with the highest PIN is only $29.37 million. Moreover, the average PIN for large firms with the highest PIN is itself much smaller (0.199) than the average PIN for small firms with the highest PIN (0.36).

4.0 Characteristics versus Covariances

4.1 Daniel-Titman tests

Although there appear to be inconsistencies in the descriptive statistics of the PIN factor loadings as discussed above, the fact remains that the PIN factor loads in a statistically significant manner in six out of the nine size-PIN portfolios in Table 7. These findings are potentially consistent with a rational model in which the PIN factor captures the risk factor underlying private information about the stock. However, as pointed out by Daniel and Titman (1997), in tests where factors are constructed from characteristics that are known return predictors, factor loadings can be found to predict returns even if risk is not priced.

In particular, PIN factor loadings and the PIN characteristics themselves are likely correlated (as confirmed later in Table 8). If markets are inefficient and investors misprice PIN characteristics perhaps because they proxy for transaction costs or liquidity, the factor loadings can pick up the mispricing that is correlated with such transaction cost characteristics. Daniel and Titman (1997) suggest that one way out of this deadlock is to identify variation in the PIN factor loading that is unrelated to the PIN characteristic and then evaluate whether the independent variation in PIN factor loadings is associated with spreads in average returns. The risk hypothesis predicts that PIN factor loadings will continue to predict returns after controlling for PIN characteristics. However, mispricing theory predicts that the PIN factor loading will have no incremental explanatory power after controlling for variation in the PIN characteristic.

4.2 Portfolio based Daniel-Titman Tests

We follow the methodology laid out in Daniel and Titman (1997), Davis, Fama and French (2000), and Daniel, Titman and Wei (2001) and triple-sort stocks into portfolios based on size, PIN characteristics, and PIN loadings. In particular, for each of the nine double-sorted size/ PIN portfolios studied in panel A of Table 7, we further divide it into three equally sized value-weighted portfolios (Low loading, Medium loading, and High loading) based on pre-formation PIN loading estimated over the previous 60 months (24 months minimum) using model (3). The resulting three sub-portfolios within each of the size/PIN category thus consist of stocks of similar size and PIN characteristics but different levels of PIN loading, and therefore should exhibit sufficiently low correlation between their PIN loading and PIN characteristic. We use these portfolios to test whether PIN factor loading can predict returns after controlling for variation in PIN characteristics.

Table 8 reports descriptive data on PIN factor loadings estimated at the firm-level for each of these nine portfolios. We need to estimate firm-level PIN factor loadings for use in the Daniel-Titman tests to follow later in this section. Panel A presents the summary of firm level regressions, grouped into the same nine size-PIN groups used in Table 7. Panel B examines the frequency of negative loadings for the PIN factor as well as other factors.

An examination of the factor loadings presented in Table 8 reveals several interesting findings. First, 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 a sorting on factor loadings keeping characteristics constant as suggested by Daniel and Titman (1997). Secondly, 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 41.7% 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.

On an overall basis, 50.2% of PIN factor loadings are negative for the entire sample. As a benchmark, the frequency of negative market betas for the entire sample is 4%, negative SMB loadings is 24.7% and negative HML loadings is 37.1%. Only UMD has a higher proportion of negative loadings (55.5%) than PIN but not many researchers claim that UMD is a priced risk factor.

Table 9 presents the summary statistics of the 27 triple-sorted portfolios as well as the five-factor model regression (2) results for these portfolios. The table confirms that the three-dimensional sort (size, PIN characteristic, and PIN loading) is effective in achieving considerable variation in PIN loadings that is unrelated to PIN characteristics. Within each of the nine size-PIN groups, the third dimensional sort on pre-formation PIN loadings produces a large spread in post formation PIN loadings while leaving the size and the PIN characteristic approximately constant.

The intercepts from the four-factor model, without the PIN factor, reported in panel A of Table 9 offer initial evidence that is inconsistent with rational factor pricing for PIN. If the PIN factor is indeed priced, the intercepts should be increasing with loadings on the PIN factor. Of the 27 intercepts, only one has a t-statistic in excess of the absolute value of two suggesting that there is no statistically significant spread between firms with high and low PIN factor loadings. Thus, it appears as though the three-factor Fama-French model (notably without the PIN factor) is sufficient to describe the cross-section of returns for a majority of these 27 portfolios.

Following Daniel and Titman (1997), Davis, Fama and French (2000), and Daniel, Titman and Wei (2001) and Hirschleifer, Hou and Teoh (2006), we formally test the theory that PIN factor loadings capture information risk by forming a 'characteristic-balanced' portfolio within each size/PIN category. To do this, for each given size/PIN group, we form a portfolio long on the high PIN loading portfolio, and short on the low PIN loading portfolio. We label such portfolios as (H^L - L^L). The returns (HRET) on such characteristic-balanced portfolios therefore reflect the pure effect of varying factor loadings. To maximize power in an overall test, we also combine the nine characteristic-balanced portfolios into a single equally weighted portfolio. The average returns and intercepts from the following five-factor model regression for the nine characteristic-balanced portfolios and combined test portfolio are presented in Table 10:

$$HRET_t = a_i + b_i\,(R_{mt} - R_{ft}) + s_i\,SMB_t + h_i\,HML_t + m_i\,UMD_t + p_i\,PINF_t + \varepsilon_{it} \qquad (5)$$

Under the null hypothesis of rational factor pricing, the five-factor regression intercepts for each characteristic-balanced portfolio should be equal to zero. In contrast, under the alternative behavioral hypothesis, variation in PIN factor loading that is independent of the PIN characteristic should not be related to average returns.

The column labeled $t(p_i)$ in panel B of Table 10 indicates that none of the intercepts is associated with a t-statistic greater than two in absolute value. Thus, *prima facie*, it appears as though PIN could be rationally priced as a risk factor. However, note that only one of the loadings on the PIN factor in the nine sub portfolios is positive and significant at the 5% level (for S/Low PIN portfolio, t-statistic = 3.28). While the PIN factor loading for the S/Medium PIN group approaches significance (t-statistic = 1.93), none of the other seven loadings are significant. Thus, when the PIN characteristic is held constant, increasing the PIN loading is associated with greater average returns in only one case and is associated with no change in average returns in eight cases.

As an aside, it is interesting to ask which sub-portfolio in S/low PIN portfolio is responsible for the significant PIN factor loading. Ideally, one would expect the high factor-loading sub portfolio to contribute most to the characteristic-balanced portfolio's PIN factor loading. However, that does not turn out to be the case. Panel B of Table 9 reports the results for the high and low loading firms in the 27 portfolios underlying equation (5). That panel reveals that the PIN factor loading for the S/low PIN/low factor-loading portfolio is –1.355 (t-statistic = -7.11) whereas the PIN factor loading for the S/low PIN/high factor loading portfolio is –0.78 (t-statistic = -4.97). Thus, the low factor loading, as opposed to the higher loading, drives the positive PIN factor loading in the characteristic balanced S/low PIN portfolio.

Before we leave this section, it is important to note that 13 of the 27 PIN factor loadings attain t-statistics in excess of absolute value of two in panel B of Table 9. Further, when we consider the t-statistics on other factors for benchmarking purposes, we find that 27 of the 27 market betas and HML factor loadings and 23 of the 27 SMB loadings attain t-statistics in excess of two. Hence, it becomes harder to argue that statistical power is a problem in our analysis.

4.3 Cross-Sectional Daniel-Titman tests

The evidence thus far relies on sorts of firms into portfolios. The sorting portfolio based approach, however, is subject to several limitations. The cutoffs for assigning firms into portfolios, although consistent with convention, are necessarily arbitrary. Further, tests based on sorts of stocks into portfolios are not easily amenable to the introduction of control variables such as firm characteristics such as size, market-to-book or lagged stock returns accumulated over various intervals.

Hirschleifer, Hou and Teoh (2006) argue that one way to address these limitations and buttress the portfolio-based tests is to conduct monthly Fama and MacBeth (1973) cross-sectional regressions. The Fama-Macbeth regressions are useful because (i) they provide an integrated and parsimonious representation of the portfolio-based tests reported so far; (ii) they enable use of data on all individual stocks in the tests rather than relying on potentially arbitrary portfolios of sub-samples of stocks; and (iii) the Fama-Macbeth set up is flexible enough to accommodate several control variables at the same time. A potential cost of the Fama-Macbeth approach is that all stocks are assigned equal weights while the portfolio approach allows value-weighting of return observations.

To examine whether PIN loadings predict returns after controlling for the PIN characteristic, in Table 11, we regress monthly individual stock returns on the firm characteristics of SIZE (log of a firm's market capitalization at the end of previous year), LBM (log of the book-to-market ratio at the fiscal year end of the previous year), 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). PIN characteristics measured as at the fiscal year end of the previous year, and factor loadings with respect to the market factor $R_m – R_f$, SMB, HML, and PIN.

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 evidence that PIN is a priced risk factor. Hence, we follow Fama and French (1992) and Hou and Moskowitz (2005) and estimate factor loadings at the portfolio level and then assign the portfolio loadings to individual stocks within a portfolio in the firm-level cross-sectional regressions. We use the same 3x3 sorting on size and PIN within size as used earlier. We regress the value weighted returns of these nine portfolios in the previous five years on the market factor $R_m – R_f$, SMB, HML, and PIN to generate portfolio loadings, ensuring that at least 24 months of return data are available. Each individual stock is then assigned the portfolio factor loadings of the size/PIN group it belongs to at the end of December of each year. This procedure collapses each stock's individual factor loadings to the averages for stocks of similar size and PIN and thus mitigates measurement error relative to an empirical strategy where we would have used the actual factor loading.

The constraint of needing 24 months of return information and factors including the PIN factor causes us to drop observations for 1983 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 hence span the 204 months from January of 1986 to December of 2002.

Panel A of Table 11 reports time series averages of the monthly cross-sectional regression coefficients from January of 1986 to December of 2004 and their time series t-statistics using a Fama-Macbeth (1973) procedure. The regression reported in as per model 1 shows that the PIN factor loadings are insignificantly related to average returns (t-statistic = 0.39). Model 2 introduces factor loadings related to the CAPM beta, SMB, HML and UMD into the regression and finds that they are also generally insignificant. Model 3 introduces the PIN characteristic, by itself, and finds that PIN does not significantly explain cross-sectional variation in returns (t-statistic = 1.44). This is not surprising considering the evidence we presented before in section 2.4. When additional firm characteristics of size, book-to-market and past returns in the cross-sectional regressions are considered along with PIN, we find that book-to-market, R1, R2_12 and R13_36 are all significantly associated with returns.

The last regression performs a characteristics-versus-covariances test in the spirit of Daniel and Titman (1997). Specifically, when we introduce both the PIN characteristic and the PIN factor loading in the regression, neither the characteristic (t-statistic = 0.97) nor the loading (t-statistic = 0.04) is significant. Thus, the cross-sectional regression test appears to strongly reject the interpretation that PIN captures priced information risk. Among the other loadings, SMB loads strongly (t-statistic = -4.77) while all the characteristics with the exception of PIN are significant.

It is heartening to see that at least one factor loading (size) and two firm characteristics (size and book-to-market), besides return momentum, appear to explain cross-section of returns for our sample period. Otherwise, we run the risk that the dataset of stock returns under investigation suffers from low statistical power. Still, to ensure that the lack of power does not drive the weak results of the PIN characteristic and the PIN factor loading, we employ the Litzenberger and Ramaswamy (1979) correction. The precision-weighted Fama-Macbeth regression coefficients are presented in Panel B of Table 11. The results are largely similar to Panel A. While the coefficients and t-statistics on both the PIN factor and the PIN loading do increase, they continue to remain insignificant. For instance, in the final specification including all characteristics and loadings, the PIN factor has a t-statistic of 1.25, while the PIN loading has a t-statistic of 0.42. By contrast, all firm characteristics (Size and Book-to-Market), past returns are significant as are the loadings on both SMB as well as HML. Hence, we can conclude that our cross-sectional results indicate that neither the PIN factor nor the PIN characteristic is associated with firm-level returns.

4.0 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 risk factor. 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 or information risk is priced in expected returns.

However, a closer scrutiny of the properties of PIN factor reveals that such enthusiasm for the interpretation of PIN as a priced factor might be somewhat premature. In particular, 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.

A formal test based on the covariances (PIN factor loadings) versus characteristics (PIN characteristic) advocated by Daniel and Titman (1997) shows that keeping PIN characteristics constant, increases in PIN factor loadings are unrelated to increased average returns. 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 priced risk factor. 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 a priced systematic risk factor.

References

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

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

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.

Bhattacharya, U. and H Daouk. 2002. The World Price of Insider Trading. *Journal of Finance* 57: 75-108

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., S.A. Hillegeist and K. Lo. 2006. The Effect of Meeting or Missing Earnings Expectations on Information Asymmetry. Working paper, Emory University, INSEAD and University of British Columbia.

Bushman, R.M., J.D. Piotroski and A.J.Smith. 2005. Insider Trading Restrictions and Analysts' Incentives to Follow Firms. *Journal of Finance* 60(1): 35-66.

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

Daniel, K., S. Titman, and J. Wei. 2001. Cross-sectional Variation in Common Stock Returns in Japan. *Journal of Finance* 56: 743-766.

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

DeBondt, W. and R. Thaler. 1985. Does the Stock Market Overreact? *Journal of Finance* 40: 793-808

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.

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

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. 2005. The Market Pricing of Earnings Quality. *Journal of Accounting and Economics* 39: 295-327.

Grullon, G., G Kanatas and JP Weston. Advertising, Breadth of Ownership, and Liquidity. *Review of Financial Studies* 17(2): 439-461.

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.

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.

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.

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.

Vega, C. 2006. Stock Price Reaction to Public and Private Information. *Journal of Financial Economics.* (forthcoming).

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 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: Replication of Easley et al (2002) in longer time period							
1984-2002	Fama-Macbeth	0.999	-0.335	0.922	0.055	0.221	3.00%
		(2.07)	(-0.82)	(1.36)	(0.78)	(2.39)	
1984-2002	L-R WLS	0.653	-0.753	1.340	0.124	0.238	3.00%
		(1.55)	(-1.98)	(1.97)	(2.07)	(2.74)	
Panel C: 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 D: Regressions with just PIN in the 1984-2002 time period							
1984-2002	Fama-Macbeth	0.871		0.880			0.58%
		(2.47)		(0.91)			
1984-2002	L-R WLS	0.693		0.659			0.58%
		(2.03)		(0.75)			

Table 3: Characteristics and Returns of PIN Portfolios based on Independent Sorts on Size and PIN

At the beginning of each year from 1984–2002, stocks are sorted into three equal groups based on market capitalization at the end of the prior year, and independently into three equal groups based on PIN. Nine size-PIN groups were formed based on the intersection of the above groupings. There were 39,376 observations in total, or an average of approximately 4,375 per group based on size-PIN grouping. Panel A outlines the number of firm-years in each size-PIN group. Panel B presents the average PIN in each portfolio. Panel C presents the average firm size in each portfolio. Panel D contains the time series average of the monthly returns of each portfolio. Returns are weighted by the prior year-end market value.

Panel A: Number of Firm-Years in Each Group

	PIN Group		
Size Group	Low	Medium	High
Small	1187	3509	8424
Medium	2939	6259	3932
Big	8994	3362	770

Panel B: Mean PIN

	PIN Group				
Size Group	Low	Medium	High	High - Low	$t_{(High-Low)}$
Small	0.146	0.203	0.311	0.164	140.06
Medium	0.143	0.194	0.275	0.131	141.06
Big	0.130	0.191	0.264	0.134	94.43

Panel C: Mean Size

	PIN Group				
Size Group	Low	Medium	High	High - Low	$t_{(High-Low)}$
Small	41.5	47.0	35.8	-5.8	-5.05
Medium	408.5	360.6	271.6	-136.9	-22.27
Big	7007.3	2349.7	1820.8	-5186.5	-28.46

Panel C: Mean Returns

	PIN Group				
Size Group	Low	Medium	High	High - Low	$t_{(High-Low)}$
Small	0.572	0.432	0.860	0.288	0.49
Medium	0.947	0.994	0.927	-0.019	-0.04
Big	1.122	1.208	0.876	-0.245	-0.53

Table 4: 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, and Dif is the difference between high and low PIN portfolios. Panel B contain the time series average of the average firm size in each portfolio, and Dif is the difference between high and low PIN portfolios. Panel C contains the time series average of the monthly returns of each portfolio. Returns are weighted by the prior year-end market value. Dif is the average return difference between high and low PIN portfolios, and t(Dif) is the t-statistics of Dif. The last row provides return statistics on a portfolio, denoted PINF, which is equally invested in each of the individual portfolios in the 5 preceding rows. The remaining tables in the paper rely on PINF thus calculated.

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 5: 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) described in Table 3. 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

Table 6: Fama-French Regressions on the Entire Sample and Size Groupings

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. 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, along with the adjusted R^2. The sample period is 1984–2002.

Panel A: Fama-French Regressions for Entire Sample

α_i	β_i	s_i	h_i	m_i	p_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	$t(p_i)$	Adj. R^2
-0.0085	1.0037	-0.1284	0.3059			-0.12	57.71	-5.91	11.73			94.3%
0.0260	0.9988	-0.1256	0.3006	-0.0299		0.36	57.18	-5.8	11.54	-1.96		94.4%
-0.0361	1.0133	-0.1236	0.3139		0.0826	-0.51	56.18	-5.68	11.95		1.87	94.4%
0.0055	1.0153	-0.1100	0.3129	-0.0699	0.1999	0.08	57.99	-5.14	12.27	-3.85	3.81	94.7%

Panel B: Fama-French Regressions by Size Groupings

Size Group	α_i	β_i	s_i	h_i	m_i	p_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	$t(p_i)$	Adj. R^2
S	-0.448	0.988	0.877	0.622			-2.75	24.29	17.25	10.21			80.2%
M	-0.265	1.111	0.571	0.648			-2.12	35.53	14.61	13.83			87.0%
B	0.018	0.998	-0.176	0.284			0.26	58.04	-8.20	11.01			94.4%
S	-0.202	0.952	0.897	0.585	-0.214		-1.31	25.21	19.16	10.38	-6.49		83.3%
M	-0.088	1.085	0.586	0.621	-0.154		-0.73	36.94	16.09	14.17	-6.01		88.7%
B	0.042	0.994	-0.174	0.280	-0.021		0.59	57.35	-8.10	10.83	-1.40		94.5%
S	-0.247	0.917	0.841	0.564		-0.605	-1.60	23.45	17.81	9.89		-6.33	83.2%
M	-0.240	1.102	0.567	0.641		-0.075	-1.88	33.79	14.40	13.49		-0.95	87.0%
B	-0.013	1.008	-0.171	0.293		0.092	-0.19	56.69	-7.96	11.29		2.12	94.5%
S	-0.164	0.921	0.869	0.562	-0.139	-0.372	-1.08	24.13	18.59	10.10	-3.52	-3.25	84.0%
M	-0.116	1.108	0.607	0.638	-0.209	0.275	-0.98	37.22	16.68	14.71	-6.77	3.08	89.1%
B	0.022	1.010	-0.159	0.292	-0.060	0.192	0.32	58.02	-7.48	11.51	-3.31	3.68	94.8%

Table 7: Fama-French Regressions on Portfolios based 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. Monthly excess returns for these portfolios are regressed against $r_m - r_f$, SMB, HML , UMD and PINF using the following specification $R_i = \alpha_i + \beta_i(r_m - r_f) + s_i SMB + h_i HML + m_i UMD + p_i PINF + \varepsilon_i$ The table reports the point estimates for each of the coefficients and their t-statistics, along with the adjusted R^2. The sample period is 1984–2002.

Panel A: Results of Factor Regression

Size Group	PIN Group	α_i	β_i	s_i	h_i	m_i	p_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	$t(p_i)$	Adj. R^2
Small	LOW	-0.235	0.953	0.561	0.908	-0.150	-0.860	-1.39	22.37	9.04	17.43	-3.40	-6.74	84.2%
Small	MEDIUM	-0.347	0.971	0.555	0.884	-0.122	-0.154	-2.01	22.33	8.76	16.60	-2.71	-1.18	80.5%
Small	HIGH	0.126	0.816	0.593	0.791	-0.135	0.222	0.81	20.99	10.48	16.62	-3.36	1.91	76.8%
Small	HEDGE	0.361	-0.136	0.033	-0.118	0.015	1.083	3.06	-4.60	0.75	-3.24	0.49	12.18	61.9%
Medium	LOW	-0.092	1.120	0.670	0.538	-0.216	0.041	-0.71	34.05	13.99	13.37	-6.35	0.42	87.4%
Medium	MEDIUM	-0.104	1.122	0.641	0.648	-0.219	0.278	-0.83	35.57	13.96	16.79	-6.69	2.94	88.5%
Medium	HIGH	-0.180	1.066	0.582	0.660	-0.174	0.608	-1.46	34.31	12.86	17.35	-5.40	6.53	87.4%
Medium	HEDGE	-0.088	-0.055	-0.089	0.121	0.043	0.567	-1.02	-2.50	-2.79	4.54	1.89	8.67	48.6%
Big	LOW	0.071	0.957	0.284	-0.240	-0.065	0.095	0.83	44.15	8.98	-9.05	-2.91	1.45	91.6%
Big	MEDIUM	-0.064	1.077	0.330	-0.044	-0.016	0.234	-0.75	50.11	10.55	-1.66	-0.70	3.63	92.9%
Big	HIGH	-0.054	1.118	0.306	0.095	-0.087	0.535	-0.56	46.08	8.66	3.21	-3.46	7.35	91.8%
Big	HEDGE	-0.125	0.161	0.022	0.335	-0.022	0.441	-1.12	5.72	0.54	9.74	-0.74	5.22	44.9%

Panel B: Analysis of Factor Loading on PIN Factor

	PIN Group	
Size Group	LOW	HIGH
1	-0.860	0.222
3	0.095	0.535
diff	0.955	0.313
t stat	6.66	2.27

Table 8: Firm Level Regressions to Estimate PIN Loadings

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. 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, along with the average adjusted R^2.

Panel A: Summary of Firm-Level Factor Regressions

Size Group	PIN Group	α_i	β_i	s_i	h_i	m_i	p_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	$t(p_i)$	Adj. R^2
Small	LOW	-0.383	0.987	1.188	0.217	-0.149	-0.832	-9.93	81.82	60.87	10.68	-9.36	-18.99	16.2%
Small	MEDIUM	-0.201	0.940	1.148	0.286	-0.110	-0.742	-5.20	78.85	56.72	14.13	-6.89	-17.10	15.0%
Small	HIGH	-0.005	0.783	0.993	0.334	-0.064	-0.528	-0.14	67.63	50.78	17.79	-4.15	-12.82	12.6%
Medium	LOW	0.150	1.063	0.690	0.215	-0.146	-0.029	5.45	113.61	46.70	12.83	-12.64	-1.01	23.6%
Medium	MEDIUM	0.324	1.091	0.811	0.218	-0.133	0.016	10.90	113.33	52.88	12.71	-10.92	0.53	23.5%
Medium	HIGH	0.492	1.014	0.841	0.313	-0.055	0.068	16.03	101.33	57.05	18.58	-4.43	2.18	21.8%
Big	LOW	0.309	1.022	-0.069	0.131	-0.053	0.018	17.60	148.03	-6.95	11.13	-6.36	1.01	34.9%
Big	MEDIUM	0.295	1.102	0.143	0.181	-0.092	0.082	14.63	146.64	13.10	14.10	-10.17	4.03	32.2%
Big	HIGH	0.504	1.114	0.323	0.226	-0.084	0.207	20.42	128.75	25.60	16.10	-8.38	8.58	28.9%

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

		Proportion of Loadings that are Negative				
Size Group	PIN Group	β_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 9: Fama-French Regressions on Portfolios based on Size, PIN and PIN Loadings

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. We further divide each 3x3 grouping based on size and PIN into 3 groups based on the firms' loading on the PIN factor calculated using firm-level regressions described in Table 6. The sample is reduced to reduced to 32,630 as we lose observations for the first two years (1984 and 1985) in estimating factor loadings. Value weighted portfolios are formed for each of these groups. Monthly excess returns for each of 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, along with the adjusted R^2. Panel A excludes the PIN factor and also presents value weighted means of prior year end size (market capitalization), PIN and PIN loading for each group.

Panel A: Regressions without PIN factor

Size/PIN/ Loading	Mean SIZE	Mean PIN	Mean Loading	α_i	β_i	s_i	h_i	m_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	Adj. R^2
S/L/L	44	0.185	-3.65	-0.435	1.102	1.000	0.646	-0.335	-1.52	16.01	11.97	6.26	-5.64	69.6%
S/L/M	55	0.187	-0.65	0.055	0.948	0.923	0.675	-0.325	0.26	18.45	14.80	8.76	-7.33	75.8%
S/L/H	53	0.187	1.79	-0.259	0.986	1.015	0.630	-0.249	-1.16	18.38	15.57	7.82	-5.38	76.2%
S/M/L	34	0.255	-3.52	-0.372	1.045	1.103	0.503	-0.136	-1.29	15.02	13.05	4.81	-2.26	69.4%
S/M/M	45	0.257	-0.55	-0.200	0.991	0.827	0.768	-0.116	-1.05	21.72	14.91	11.20	-2.93	77.5%
S/M/H	44	0.256	1.83	-0.306	0.942	0.874	0.756	-0.073	-1.33	17.00	12.99	9.09	-1.52	69.0%
S/H/L	24	0.357	-3.08	-0.009	0.883	0.883	0.720	-0.054	-0.04	15.71	12.94	8.53	-1.11	66.8%
S/H/M	32	0.367	-0.31	0.044	0.743	0.718	0.603	-0.043	0.23	16.28	12.94	8.79	-1.08	67.7%
S/H/H	32	0.362	1.80	0.394	0.780	0.784	0.616	-0.132	1.87	15.42	12.75	8.10	-3.01	66.8%
M/L/L	412	0.152	-1.80	-0.277	1.158	0.576	0.703	-0.190	-1.62	28.10	11.51	11.36	-5.34	83.1%
M/L/M	445	0.150	0.03	-0.025	1.027	0.436	0.748	-0.121	-0.17	28.97	10.14	14.05	-3.95	82.6%
M/L/H	428	0.153	1.67	0.037	1.156	0.646	0.740	-0.250	0.22	28.21	12.98	12.03	-7.06	83.9%
M/M/L	355	0.199	-1.91	-0.033	1.149	0.659	0.673	-0.145	-0.20	29.12	13.75	11.36	-4.27	84.6%
M/M/M	365	0.200	0.11	-0.065	1.074	0.604	0.780	-0.095	-0.39	27.12	12.56	13.12	-2.77	81.6%
M/M/H	362	0.199	1.84	0.005	1.091	0.684	0.626	-0.184	0.03	27.02	13.94	10.31	-5.29	83.4%
M/H/L	282	0.262	-1.83	-0.352	1.083	0.711	0.666	-0.076	-1.90	24.30	13.14	9.95	-1.98	79.7%
M/H/M	302	0.268	0.15	-0.011	0.960	0.555	0.713	-0.002	-0.07	24.12	11.47	11.92	-0.05	77.5%
M/H/H	303	0.268	1.87	-0.143	1.004	0.595	0.589	-0.051	-0.84	24.60	12.00	9.60	-1.46	79.5%
B/L/L	12430	0.110	-1.15	0.207	0.903	-0.357	0.215	-0.092	1.26	22.75	-7.41	3.61	-2.69	76.8%
B/L/M	10936	0.110	0.02	0.039	0.972	-0.171	0.373	-0.004	0.33	33.40	-4.83	8.55	-0.16	86.0%
B/L/H	10426	0.113	1.17	-0.011	0.976	-0.148	0.358	-0.030	-0.10	37.83	-4.73	9.25	-1.34	88.9%
B/M/L	3775	0.147	-1.21	-0.073	1.044	-0.004	0.296	0.024	-0.58	34.36	-0.11	6.49	0.91	87.0%
B/M/M	4182	0.146	0.08	0.046	1.064	-0.116	0.443	0.030	0.35	33.79	-3.04	9.36	1.11	85.9%
B/M/H	4077	0.147	1.36	-0.025	1.079	-0.020	0.395	0.020	-0.18	31.39	-0.47	7.66	0.69	84.3%
B/H/L	2590	0.195	-1.28	-0.362	1.114	0.105	0.347	0.069	-2.28	29.23	2.27	6.07	2.10	82.9%
B/H/M	2733	0.195	0.23	0.168	1.055	-0.029	0.442	0.042	1.07	27.99	-0.63	7.81	1.28	80.6%
B/H/H	2520	0.196	1.67	0.097	1.102	0.059	0.324	-0.066	0.52	24.46	1.07	4.79	-1.71	77.6%

Table 9: Fama-French Regressions on Portfolios based on Size, PIN and PIN Loadings (Cont'd)

Panel B: Regressions with PIN factor

Size/PIN/ Loading	α_i	β_i	s_i	h_i	m_i	p_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	$t(p_i)$	Adj. R^2
S/L/L	-0.366	0.998	0.901	0.582	-0.064	-1.355	-1.43	15.78	11.85	6.27	-0.97	-7.11	75.7%
S/L/M	0.087	0.901	0.878	0.646	-0.202	-0.615	0.42	17.70	14.36	8.65	-3.83	-4.02	77.5%
S/L/H	-0.219	0.926	0.958	0.593	-0.093	-0.780	-1.04	17.77	15.28	7.75	-1.73	-4.97	78.7%
S/M/L	-0.351	1.014	1.073	0.483	-0.054	-0.409	-1.22	14.27	12.56	4.63	-0.74	-1.91	69.8%
S/M/M	-0.190	0.977	0.813	0.759	-0.078	-0.188	-1.00	20.86	14.44	11.04	-1.61	-1.33	77.5%
S/M/H	-0.308	0.944	0.876	0.757	-0.078	0.025	-1.33	16.53	12.77	9.04	-1.32	0.14	68.9%
S/H/L	-0.020	0.899	0.899	0.730	-0.097	0.216	-0.09	15.59	12.97	8.62	-1.63	1.24	66.9%
S/H/M	0.030	0.765	0.738	0.616	-0.099	0.282	0.16	16.42	13.18	9.01	-2.05	2.01	68.2%
S/H/H	0.385	0.794	0.797	0.624	-0.167	0.177	1.83	15.28	12.75	8.18	-3.11	1.13	66.8%
M/L/L	-0.274	1.153	0.571	0.700	-0.177	-0.063	-1.59	27.18	11.20	11.23	-4.04	-0.50	83.1%
M/L/M	-0.024	1.025	0.435	0.747	-0.117	-0.019	-0.16	28.08	9.91	13.93	-3.10	-0.17	82.5%
M/L/H	0.029	1.170	0.659	0.748	-0.284	0.171	0.17	27.82	13.03	12.13	-6.54	1.35	84.0%
M/M/L	-0.044	1.165	0.675	0.683	-0.189	0.216	-0.27	28.90	13.92	11.54	-4.52	1.78	84.8%
M/M/M	-0.077	1.093	0.622	0.792	-0.144	0.245	-0.47	27.05	12.81	13.35	-3.44	2.01	81.9%
M/M/H	-0.008	1.110	0.702	0.637	-0.233	0.245	-0.05	26.93	14.16	10.53	-5.47	1.97	83.6%
M/H/L	-0.381	1.127	0.753	0.694	-0.191	0.574	-2.14	25.68	14.28	10.76	-4.21	4.34	81.4%
M/H/M	-0.047	1.014	0.606	0.746	-0.141	0.696	-0.31	26.96	13.40	13.51	-3.63	6.14	81.0%
M/H/H	-0.175	1.052	0.640	0.618	-0.176	0.621	-1.10	26.68	13.51	10.67	-4.31	5.23	81.9%
B/L/L	0.199	0.916	-0.345	0.223	-0.126	0.166	1.21	22.49	-7.05	3.74	-2.98	1.35	76.9%
B/L/M	0.039	0.972	-0.171	0.374	-0.004	0.001	0.32	32.42	-4.73	8.49	-0.14	0.01	85.9%
B/L/H	-0.011	0.977	-0.148	0.359	-0.032	0.008	-0.10	36.74	-4.62	9.19	-1.15	0.10	88.8%
B/M/L	-0.084	1.060	0.012	0.307	-0.020	0.219	-0.67	34.36	0.32	6.76	-0.62	2.35	87.3%
B/M/M	0.029	1.089	-0.092	0.458	-0.037	0.334	0.23	34.62	-2.42	9.92	-1.13	3.53	86.7%
B/M/H	-0.033	1.091	-0.009	0.403	-0.010	0.150	-0.23	30.94	-0.20	7.77	-0.26	1.41	84.4%
B/H/L	-0.384	1.147	0.137	0.368	-0.017	0.429	-2.50	30.22	3.00	6.60	-0.43	3.75	83.9%
B/H/M	0.142	1.094	0.009	0.466	-0.062	0.518	0.95	29.71	0.20	8.62	-1.63	4.67	82.4%
B/H/H	0.074	1.137	0.092	0.346	-0.158	0.458	0.40	25.18	1.70	5.22	-3.38	3.36	78.7%

Table 10: Fama-French Regressions on Characteristic Balanced Hedge Portfolios based on PIN Loadings

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 PIN estimated over the prior year. We further divide each 3x3 grouping based on size and PIN into 3 groups based on the firms' loading on the PIN factor calculated using firm-level regressions described in Table 6. The sample is reduced to reduced to 32,630 as we lose observations for the first two years (1984 and 1985) in estimating factor loadings. Value weighted portfolios are formed for each of these groups. For each of the nine size/PIN groups, a characteristic-balanced zero-investment portfolio (H^L–L^L) is formed by taking a long position in the highest PIN loading portfolio and a short position in the lowest PIN loading portfolio. Finally, a combined characteristic-balanced portfolio is formed by equal-weighting the above nine characteristic-balanced portfolios. The returns on the characteristic-balanced portfolios (HRET) are regressed on $R_M - R_F$, SMB, HML, and PINF from January 1986 to December 2002 using the following specification. (Panel A excludes PINF). $HRET_t = \alpha_i + \beta_i(r_m - r_f) + s_i SMB + h_i HML + m_i UMD + p_i PINF + \varepsilon_i$

Panel A: Regressions without PIN factor

Size/PIN	Avg. HRET	t(HRET)	α_i	β_i	s_i	h_i	m_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	Adj. R^2
S/L	0.197	0.85	0.176	-0.115	0.015	-0.016	0.086	0.73	-1.99	0.21	-0.19	1.71	2.3%
S/M	0.158	0.51	0.065	-0.103	-0.229	0.253	0.063	0.22	-1.42	-2.58	2.31	1.00	13.4%
S/H	0.243	1.12	0.403	-0.102	-0.099	-0.104	-0.078	1.78	-1.88	-1.50	-1.27	-1.65	1.9%
M/L	0.261	1.57	0.314	-0.001	0.070	0.038	-0.060	1.79	-0.03	1.36	0.59	-1.64	0.2%
M/M	-0.047	-0.27	0.038	-0.057	0.025	-0.047	-0.039	0.21	-1.29	0.46	-0.71	-1.02	-0.6%
M/H	0.171	0.93	0.209	-0.079	-0.117	-0.078	0.025	1.09	-1.72	-2.08	-1.12	0.62	1.7%
B/L	-0.081	-0.46	-0.218	0.073	0.209	0.143	0.062	-1.24	1.73	4.08	2.25	1.71	8.2%
B/M	0.094	0.57	0.047	0.036	-0.015	0.099	-0.004	0.27	0.87	-0.31	1.59	-0.10	0.0%
B/H	0.310	1.27	0.459	-0.012	-0.046	-0.023	-0.135	1.80	-0.19	-0.62	-0.25	-2.55	1.6%
Combined	0.145	1.65	0.166	-0.040	-0.021	0.029	-0.009	1.83	-1.84	-0.79	0.89	-0.47	3.6%

Panel B: Regressions with PIN factor

Size/PIN	α_i	β_i	s_i	h_i	m_i	p_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(m_i)$	$t(p_i)$	Adj. R^2
S/L	0.147	-0.071	0.057	0.011	-0.030	0.575	0.62	-1.23	0.81	0.13	-0.49	3.28	6.9%
S/M	0.043	-0.070	-0.197	0.274	-0.024	0.434	0.14	-0.94	-2.20	2.50	-0.31	1.93	14.6%
S/H	0.405	-0.105	-0.102	-0.106	-0.070	-0.039	1.78	-1.87	-1.51	-1.29	-1.20	-0.23	1.4%
M/L	0.302	0.017	0.087	0.049	-0.107	0.235	1.73	0.38	1.68	0.77	-2.39	1.80	1.3%
M/M	0.037	-0.055	0.027	-0.046	-0.045	0.028	0.20	-1.21	0.49	-0.69	-0.95	0.20	-1.1%
M/H	0.206	-0.076	-0.113	-0.076	0.015	0.047	1.07	-1.59	-1.98	-1.08	0.31	0.33	1.3%
B/L	-0.210	0.061	0.198	0.135	0.094	-0.158	-1.19	1.41	3.79	2.13	2.10	-1.21	8.4%
B/M	0.051	0.031	-0.020	0.096	0.010	-0.069	0.29	0.72	-0.40	1.53	0.23	-0.53	-0.4%
B/H	0.457	-0.010	-0.044	-0.022	-0.141	0.029	1.78	-0.15	-0.58	-0.23	-2.15	0.15	1.2%
Combined	0.160	-0.031	-0.012	0.035	-0.033	0.120	1.77	-1.39	-0.45	1.07	-1.42	1.79	4.6%

Table 11: Monthly Cross-Sectional Regressions of Stock Returns on Characteristics and Factor Loadings

This table presents results from firm-level Fama-MacBeth (1973) cross-sectional regressions estimated every month between January 1986 and December 2002. Monthly individual stock returns (RET) are regressed on 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), PIN measured at prior year end and 5-year pre-ranking portfolio factor loading with respect to the market factor (LRMRF), SMB (LSMB), HML (LHML) and PINF (LPIN). Portfolio factor loadings are calculated using 60 prior months of returns using nine portfolios, based on three groups of size and three groups of PIN with size groups. Time-series means of monthly regression coefficients are reported with their time-series t-statistics below in parentheses. In Panel B, 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).

Panel A: Fama-Macbeth Regressions

Model	Intercept	SIZE	LBM	R1	R2_12	R13_36	PIN	LRMRF	LSMB	LHML	LUMD	LPIN	Adj. R^2
ONLY PIN	1.097											0.138	0.9%
LOADING	(3.29)											(0.39)	
ALL FACTOR	2.701							-1.819	-0.362	1.099	-0.213	0.176	1.7%
LOADINGS	(3.01)							(-1.89)	(-0.97)	(1.38)	(-0.20)	(0.56)	
ONLY PIN	0.678						1.724						0.6%
CHARACTERISTIC	(1.77)						(1.44)						
ALL FIRM	1.100	-0.011	0.109	-4.999	0.843	-0.167	0.565						3.7%
CHARACTERISTICS	(1.50)	(-0.15)	(2.32)	(-7.97)	(3.83)	(-1.84)	(0.71)						
ALL LOADINGS AND	3.016	-0.309	0.115	-4.935	0.871	-0.146	0.979	0.450	-1.658	0.693	0.252	0.011	4.1%
CHARACTERISTICS	(2.70)	(-3.07)	(2.54)	(-7.99)	(3.98)	(-1.65)	(0.97)	(0.56)	(-4.50)	(1.02)	(0.26)	(0.04)	

Panel B: L-R Precision Weighted Fama-Macbeth Regressions

Model	Intercept	SIZE	LBM	R1	R2_12	R13_36	PIN	LRMRF	LSMB	LHML	LUMD	LPIN	Adj. R^2
ONLY PIN	0.874											0.240	0.9%
LOADING	(2.73)											(1.02)	
ALL FACTOR	2.029							-1.193	-0.496	1.160	0.347	0.335	1.7%
LOADINGS	(2.74)							(-1.41)	(-1.87)	(1.85)	(0.38)	(1.41)	
ONLY PIN	0.597						0.877						0.6%
CHARACTERISTIC	(1.61)						(0.84)						
ALL FIRM	0.305	0.069	0.145	-4.628	0.869	-0.147	1.129						3.7%
CHARACTERISTICS	(0.48)	(1.12)	(3.62)	(-8.19)	(4.81)	(-2.06)	(1.57)						
ALL LOADINGS AND	1.943	-0.171	0.149	-4.584	0.889	-0.134	1.060	0.238	-1.360	0.875	0.455	0.094	4.1%
CHARACTERISTICS	(2.03)	(-2.08)	(3.83)	(-8.22)	(4.94)	(-1.93)	(1.25)	(0.34)	(-4.77)	(1.76)	(0.56)	(0.42)	