Journal of Accounting and Economics 47 (2009) 226–243

Contents lists available at ScienceDirect

Journal of Accounting and Economics

journal homepage: www.elsevier.com/locate/jae

Is PIN priced risk?[☆]

Partha Mohanram[a,1], Shiva Rajgopal[b,*]

[a] Columbia Business School, 3022 Broadway, New York, NY 10027, USA
[b] University of Washington Business School, Box 353200, Seattle, WA 98195, USA

ARTICLE INFO

Article history:
Received 9 March 2007
Received in revised form
25 August 2008
Accepted 14 October 2008
Available online 29 October 2008

JEL classification:
G12
G14
M40

Keywords:
Information risk
PIN
Risk

ABSTRACT

Several recent papers assume that private information (PIN), proposed by Easley et al. [2002. Is information risk a determinant of asset returns? Journal of Finance 57, 2185–2221; 2004. Factoring information into returns. Working Paper, Cornell University], is a determinant of stock returns. We replicate Easley et al. (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. There is no evidence that PIN factor loadings predict returns or that PIN factor returns reflect future GDP growth. PIN exhibits no association with implied cost of capital derived from analysts' earnings forecasts. Overall, our findings cast doubt on whether PIN reflects information risk systematically priced by investors.

1. Introduction

An influential set of recent papers by Easley et al. (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 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 literature (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.'s (2002) result that higher PIN is associated with higher cost of capital (see the appendix for detailed citations). We investigate whether PIN reflects information risk that is systematically priced by investors for three reasons.

☆ We thank our respective schools for financial support. We acknowledge helpful written comments from the referee (Mozaffar Khan), the editor (Ross Watts), Hemang Desai, Gus DeFranco, Jefferson Duarte, Per Olsson, Stephen Penman, Darren Roulstone, Gil Sadka, Lakshmanan Shivakumar, K.R. Subramanyam, Rodrigo Verdi and workshop participants at Columbia University, MIT, University of Washington, Rutgers University, University of Southern California, Tel Aviv University, Hebrew University, the 2007 FARS Conference, 2007 JAAF conference, 2007 EFMA Meeting in Vienna, and the 2007 American Accounting Association (AAA) meetings. All errors are ours.
* Corresponding author. Tel.: +1 206 543 7913; fax: +1 206 685 9392.
E-mail addresses: pm2128@columbia.edu (P. Mohanram), rajgopal@u.washington.edu (S. Rajgopal).
[1] Tel.: +1 212 854 2561, fax: +1 212 316 9219.

0165-4101/$ - see front matter
doi:10.1016/j.jacceco.2008.10.001

First, there has been considerable skepticism regarding the pricing of information risk. Earlier work by Fama (1970) suggests that information risk is potentially fully diversifiable and hence does not have to be priced. More recently, Hughes et al. (2005) and Lambert et al. (2007) question the Easley et al. (2002) result and conclude that information risk is either diversifiable or subsumed by existing risk factors. Spiegel and Wang (2005, footnote 6) suggest that PIN captures a stock's liquidity characteristics and whether liquidity is a systematic risk is unclear.

Second, unlike Easley et al. (2002, 2004), we explicitly test whether PIN is a priced risk factor using a (i) a time-series regression of PIN factor returns on time-series measures of the three Fama–French factors (market factor, SMB and HML) and the momentum factor (UMD), and a (ii) two-pass cross-sectional regression of excess returns on risk factor betas. The two-pass cross-sectional test has been extensively used to evaluate whether a candidate variable is a priced factor as in testing the CAPM (Fama and MacBeth, 1973), the conditional CAPM (Jagannathan and Wang, 1996), the inter-temporal CAPM (Brennan et al., 2004; Petkova, 2006), and the two-beta model (Campbell and Vuolteenaho, 2004) and in evaluating whether accruals or earnings quality is a priced risk factor (Hirshleifer et al., 2006; Core et al., 2008; Khan, 2008).

Third, the empirical literature is beginning to raise questions on whether PIN captures priced information risk. Duarte and Young (2008) report that the PIN component related to asymmetric information is not priced, while the PIN component related to illiquidity is priced. Aktas et al. (2007) document that the PIN variable decreases before announcements of M&A transactions and increases after the announcement although there is considerable evidence of illegal insider trading or information leakage prior to such M&A announcements (Dennis and McConnell, 1986; Keown and Pinkerton, 1981; Meulbroek, 1992). Benos and Jochec (2007) find, counter-intuitively, that the PIN variable is lower in the periods before earnings announcements dates than in the periods after earnings announcements dates. This contradicts prior research which has consistently found that information asymmetry between investors and therefore opportunities for informed trading are the greatest prior to earnings announcements (Brooks, 1996; Christophe et al., 2004; Frazzini and Lamont, 2006). In general, considering (1) the critical role that the accounting process plays in the generation of information, (2) the reliance in extant 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.

To investigate this question, we conduct four tests. First, we investigate whether the Easley et al. (2002) result for the 1984–1998 period that PIN predicts future stock returns is robust to multiple periods of varying time-series length. Easley et al. (2002) use data ranging from 1984 to 1998, whereas we have access to four additional years of data, 1998–2002. This allows us to evaluate whether the pricing of PIN is a robust result.

Second, we conduct a time-series test of regressing returns to the PIN factor on the four standard factors (market factor, SMB, HML and UMD) and an intercept. If the PIN factor is priced, we expect to find a significant intercept term. Further, we estimate Fama and MacBeth (1973) type two-pass 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.

Third, we evaluate whether PIN is a risk factor by regressing returns to the PIN factor on future GDP growth, a proxy for future macro-economic activities, as these activities contain information about the future investment set of investors. Prior research including Chen (1991), Liew and Vassalou (2000) and Chordia and Shivakumar (2006) rely on similar intuition to test whether the Fama–French factors or price momentum proxy for risk factors. If PIN were a risk factor, we expect a positive association between PIN factor and future business conditions.

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 and higher PIN factor loadings 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 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. Further investigation reveals that this counter-intuitive result occurs because returns to the PIN factor are negatively correlated with returns for high PIN stocks although one would expect such correlation to be positive. In the time-series tests, we do not find a significant intercept in a regression of the PIN factor on the market factor, SMB, HML and UMD. In the two-pass cross-sectional tests, we find that the PIN factor loadings do not exhibit any statistical association with returns and a negative association with the ex-ante cost of capital. We are also unable to document any association between PIN factor and future GDP growth. 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, 1970, 1991) assumes that information risk is completely diversifiable and should hence have no effect on expected returns. A review of the extant literature in accounting and finance suggests that there are at least two notions of information risk: (i) information asymmetry and (ii) estimation risk or parameter uncertainty.

The first stream of work, represented by several papers by Easley and O'Hara argues that systematic, un-diversifiable information risk arises from information asymmetry between the informed and uninformed traders. The second stream of research stems from Bawa et al. (1979), among others (e.g., Jorion, 1985; Kandel and Stambaugh, 1996; Stambaugh, 1999; Barberis, 2000; Lewellen and Shanken, 2002), who consider the estimation risk stemming from investors' uncertainty about the true underlying parameters of the returns or the cash flow process of stocks when they have to make a portfolio allocation decision. When low information securities form a non-trivial portion of an investor's portfolio, estimation risk is likely to have an arguably non-diversifiable effect because the resolution of uncertainty about low information securities affects the return earned on the portfolios. Of course, whether either notion of information risk is un-diversifiable is open to debate.

The focus of our work is the first stream of papers related to information risk and, in particular, the recent stream of papers by Easley et al. (2002) and Easley and O'Hara (2004) which has developed the theoretical intuition for why information risk can potentially 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.

Researchers as far back as Fama (1970) have argued that information risk is potentially idiosyncratic and hence diversifiable. More recent theory work has questioned the specific findings of the Easley et al. papers. Hughes et al. (2005) study the role of information risk in a multi-factor asset-pricing model and conclude that information risk is either diversifiable or subsumed by existing risk factors. Lambert et al. (2007) argue that the Easley and O'Hara (2004) inference that information risk is 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 and is fully diversifiable. A small but growing empirical literature is beginning to cast doubt on whether PIN even captures information risk (see cites in the Introduction). Given (1) the vital role played by the financial reporting process in the generation of firm-value relevant information, (2) the general assumption in extant accounting literature that PIN is priced information risk, and (3) the lack of consensus on whether information risk, in general, and PIN in particular, commands a risk premium, we believe that it is important to empirically investigate whether 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

$$\mathrm{PIN} = \frac{\alpha\mu}{\alpha\gamma + \varepsilon_{\mathrm{S}} + \varepsilon_{\mathrm{B}}} \tag{1}$$

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, ε_{S} is the 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. (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.[2]

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.,

[2] Asquith et al. (2007) 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.

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 1863 and 2414 stocks in the years 1983–2001. For further details on the construction and content of the dataset, see Easley et al. (2004).

Untabulated descriptive data confirms that our sample matches theirs. 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 $\rho = -0.660$).[3] In a recent paper, Aslan et al. (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 6 of Easley et al., 2002) that the PIN characteristic is priced. In particular, Easley et al. (2002) use PIN data over the years 1983–1997 and regress 1-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 2 years, when possible, 5 years before the test year. Thus, for each stock, we use at least 24 monthly return observations in the estimation. We run a regression of 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 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, $\beta_{\rm p}$, is then the sum of the two coefficients. For the firm-level cross-sectional regressions, we set each firm's beta to be the beta of the portfolio to which it belongs.

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:

$$R_{it} = \gamma_{\rm ot} + \gamma_{1t}\beta_{\rm p} + \gamma_{2t}\mathrm{PIN}_{it-1} + \gamma_{3t}\mathrm{SIZE}_{it-1} + \gamma_{4t}\mathrm{LBM}_{it-1} + \varepsilon_{it} \tag{2}$$

where R_{it} is the excess return of stock i in month m of year t, and 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 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 1 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 (L–R) 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. In untabulated work, we find that PIN, by itself, is not related to returns (t-statistic = 0.59 under Fama–MacBeth and 0.34 under L–R 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 B, 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, 1998–2002, 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 5-year windows (1984–1988, 1989–1993, 1994–1998, 1999–2002). The 1999–2002 window only has 4 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 private information because we rely on Fama–Macbeth tests that use between 48 and 60 months of data.

[3] 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.

Table 1
Replication of asset pricing tests for the PIN characteristics from Easley et al. (2002) with new data and across sub-periods.

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) 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)	

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.

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, Hvidkjaer and O'Hara result that PIN is priced appears to be restricted to the 1984–1988 period. The pricing of PIN does not appear to be robust in an extended time period.

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.[4] In particular, at the beginning of the year t, we sort stocks into 10 deciles based on market capitalization at the end of the prior year ($t-1$). Next, within each size decile, we sort into three equal-sized groups based on PINs from the prior year (labeled S for small, M for medium and B for big). The sequential sorting procedure yields 30 portfolios based on 10 levels of size and three levels of PIN (10*3). This sequential sorting process ensures that each of the 30 sub-portfolios has roughly an equal number of firm-year observations (approximately 1313 firm years). We rely on these sequentially sorted portfolios in the remainder of the paper.

Table 2 reports descriptive data on PIN, size and value-weighted as well as equally-weighted monthly returns in excess of 1-month T-bill rates (*Exret*) for each of these 30 portfolios are computed from January to December of year *t*. The data reveal several interesting patterns. First, sorting PIN into 10 portfolios, keeping size constant, does appear to capture reasonable variation in PIN independent of size. In particular, PIN spreads in each size group, reported in Panel B, are strongly significant at conventional levels. It is interesting to note that the spread in PINs of 19.5% is highest for the smallest size group and lowest for the largest size group (spread = 6.3%), suggesting that PIN is likely to have greater potential to explain returns for small stocks relative to large stocks.

Second, panel A shows that for a given size category, as PIN increases, the average size declines. That is, the largest firms fall in the size decile 10 but low PIN group (average market capitalization = $20.897 billion), while the smallest firms fall in the size decile 1 but high PIN group (average market capitalization = $7.9 million). This outcome seems intuitive because

[4] 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.

Table 2
Replication of the characteristics and returns of PIN portfolios based on sequential sorts on size and PIN as per Easley et al. (2004).

Panel A: Mean size (market capitalization) by size-PIN groups

Size decile	Low PIN	Medium PIN	High PIN	High−Low	*t*-stat
1	9.1	8.6	7.9	−1.2	−5.64
2	28.3	27.8	26.9	−1.3	−2.56
3	63.6	63.2	61.2	−2.4	−2.13
4	126.5	125.1	122.4	−4.1	−1.75
5	232.1	231.1	225.7	−6.3	−1.43
6	409.5	402.6	394.7	−14.9	−1.98
7	732.7	717.6	691.3	−41.5	−3.39
8	1339.8	1313.3	1277.9	−61.8	−2.75
9	2906.9	2797.0	2640.6	−266.3	−4.89
10	20897.0	11443.1	9693.0	−11204.0	−13.63

Panel B: Mean PIN (%) by size-PIN groups

Size decile	Low PIN	Medium PIN	High PIN	High-Low	*t*-stat
1	20.3	28.2	39.8	19.5	76.95
2	19.0	25.9	35.8	16.9	79.24
3	18.0	24.0	32.7	14.7	76.18
4	16.6	22.0	29.8	13.2	75.79
5	16.0	20.8	27.3	11.3	70.04
6	15.2	19.5	25.5	10.3	66.28
7	13.9	18.0	23.5	9.6	64.45
8	13.0	16.5	21.4	8.4	58.12
9	12.1	15.0	19.1	7.0	53.96
10	9.8	12.4	16.1	6.3	53.95

Panel C: Mean 1-year-ahead monthly stock return (%)

	Value-weighted returns					Equally-weighted returns				
Size decile	Low PIN	Medium PIN	High PIN	High−Low	*t*-stat	Low PIN	Medium PIN	High PIN	High−Low	*t*-stat
1	1.12	1.11	1.54	0.42	0.57	1.50	1.74	2.28	0.78	0.93
2	0.25	0.61	1.11	0.86	1.47	0.27	0.62	1.12	0.85	1.42
3	0.46	0.49	0.89	0.43	0.83	0.46	0.48	0.88	0.42	0.81
4	0.90	0.78	1.01	0.10	0.20	0.88	0.74	1.02	0.14	0.26
5	0.84	0.94	0.93	0.09	0.17	0.80	0.93	0.92	0.12	0.24
6	0.84	1.03	1.05	0.21	0.42	0.83	1.03	1.03	0.20	0.41
7	1.09	1.06	1.05	−0.04	−0.09	1.09	1.06	1.05	−0.04	−0.09
8	1.07	1.15	1.13	0.06	0.12	1.06	1.14	1.13	0.07	0.15
9	1.05	1.17	1.17	0.12	0.25	1.05	1.17	1.14	0.09	0.21
10	1.14	1.13	1.15	0.01	0.02	1.19	1.15	1.09	−0.11	−0.25
Average	0.88	0.95	1.10	0.22	2.12	0.91	1.01	1.17	0.25	2.39

At the beginning of each year from 1984 to 2002, stocks are sorted into deciles based on size (market capitalization at prior year end), and within each size decile, three groups are formed based on prior year PIN. There were 39,376 observations in total, or approximately 1313 per size-PIN group. The table presents mean size (Panel A), mean PIN (Panel B) and mean value-weighted and equally-weighted 1-year-ahead monthly return (Panel C) for each group. *t*-Statistics for difference between groups are calculated using a pooled differences-in-means test. Panel C also presents the equally-weighted average of returns across the size deciles. The equally-weighted mean across the size deciles of the difference in value-weighted returns between High PIN and Low PIN firms are used to calculate the PIN factor, PINF, used in subsequent tables.

the largest firms are most likely well followed by information intermediaries such as analysts and are likely to be associated with lower PIN relative to the smallest firms.

Third, somewhat remarkably, the spread in returns between the highest and lowest PIN groups is not statistically significant at conventional levels (panel C) for any given size decile, for either value-weighted returns or equally-weighted returns. The only *t*-statistic that approaches statistical significance relates to the value-weighted spread of 0.86% of size decile 2 (*t*-statistic = 1.47). Hence, the evidence in Easley et al. (2004) that the PIN factor earns a statistically significant spread among small firms is inconsistent with the evidence in Table 2. However, when the returns are aggregated across the 10 size deciles, it appears as though the highest PIN group does outperform the lowest PIN group by 0.22% using value-weighted returns (0.25% using equally-weighted returns) and the spread is statistically significant with a *t*-statistic of 2.12 (2.39 for equally-weighted). As explained in the following section, the 0.22% average for value-weighted returns can be interpreted as the returns to a PIN factor, PINF. Thus, PINF needs the statistical power of the entire sample to produce statistically significant returns.

3. Time-series and cross-sectional tests of returns on PIN factor loadings

Easley et al. (2004, p. 3) believe that PIN might be a factor 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 a more rigorous test of whether stock prices reflect a systematic risk premium for PIN, we conduct two tests. First, we conduct a time-series regression of a factor based on PIN on the existing Fama–French factors, to test whether there is an incremental risk premium for the PIN factor. Second, we estimate a cross-sectional regression of returns on PIN factor loadings, and examine the statistical significance of the coefficients of these loadings. This test has been extensively used to evaluate whether a candidate variable is a priced factor as in testing the CAPM (Fama and MacBeth, 1973), the conditional CAPM (Jagannathan and Wang, 1996), the inter-temporal CAPM (Brennan et al., 2004; Petkova, 2006), and the two-beta model (Campbell and Vuolteenaho, 2004) and in evaluating whether accruals or earnings quality is a priced risk factor (Hirshleifer et al., 2006; Core et al., 2008; Khan, 2008). To conduct these tests, 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). The process is same as the one underlying returns described in panel C of Table 2 above. However, for completeness, we re-describe the process of creating the PIN factor, PINF, in detail. 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 size deciles.

The descriptive statistics reported in panel A of Table 3 related to the PIN factor, PINF, and the other factors closely resemble those reported in Easley et al. (2004). Note that the mean return to PINF is the same 0.22% found in Table 2.

Table 3
Time-series relationship between PIN factor and Fama–French factors.

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	0.224	1.592	2.12

Panel B: Correlations

	R_m–R_f	SMB	HML	UMD	PINF
R_m–R_f		0.164	−0.524	−0.087	−0.247
SMB	0.122		−0.452	0.108	−0.093
HML	−0.551	−0.292		−0.078	0.029
UMD	−0.072	−0.069	−0.029		0.576
PINF	−0.244	−0.057	0.042	0.381	

Panel C: Time-series regression of PINF on other factors
Model PINF $= \alpha + \beta(r_m - r_f) + s$ SMB $+ h$ HML $+ m$ UMD $+ \varepsilon_i$

Model	α	β	s	h	m	Adj. R^2 (%)
3-factor	0.317	−0.115	−0.059	−0.096		7.61
	(3.04)	(−4.42)	(−1.82)	(−2.45)		
4-factor	0.087	−0.082	−0.078	−0.060	0.199	38.45
	(1.00)	(−3.82)	(−2.94)	(−1.89)	(10.64)	

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 2. 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. Figures above (below) the diagonal are Pearson (Spearman rank-order) correlations.

The correlation table in panel B of Table 3 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 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. The high correlation between UMD and PINF may be related to the fact PIN is estimated essentially from order imbalance, and high momentum is also likely to be associated with strong order imbalance. Given this high correlation, we run all our tests both with and without the momentum factor UMD.

As an aside, it is important to note that the *t*-statistics in Panel A of Table 2 seem to suggest that the PINF portfolio formation procedure does not adequately control for size, in contrast to what is intended. For example, in size decile 1, the size difference between low and high PIN groups is significant. In contrast, the intention is to control for size differences between the low and high PIN groups. To address this persistent correlation between size and PIN, in untabulated work, we employed a methodology motivated by Penman (1983), a paper which analyzes the information content of management earnings forecasts incremental 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 deciles 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.

Note that statistically significant spreads on PINF are not *ipso facto* evidence that PINF is a priced risk factor. Firstly, the returns to PINF may be subsumed by the existing Fama–French risk factors. Secondly, one needs to verify that factor loadings for each variable predict returns, as is standard in the asset-pricing literature (e.g., Fama and MacBeth, 1973; Daniel and Titman, 1997; Davis et al., 2000). We turn to these tests next.

A researcher can use two methodologies to assess whether there is a risk premium to the PINF factor: (i) a time-series regression method and (ii) the two-pass cross-sectional regression method. We conduct both these tests to convince the reader that our results are not attributable to methodological differences.

3.2. Time-series regression of PINF on Fama–French factors

To conduct the time-series regression tests, we use the PINF factor, as described in Section 3.1, and regress the time series of PINF returns on the time-series of the three Fama–French (FF) factors. The idea is to test whether the intercept from such a regression is significantly different from zero. An insignificant intercept would suggest that PINF is not priced, and that the significant mean return to PINF of 0.224 reported earlier is simply due to its correlation with the FF factors. However, a significant intercept would suggest there is some premium for PINF exposure beyond that due to the correlation of PINF with the FF factors. The results from such a regression are presented in Panel C of Table 3.

The first regression uses the market (R_m–R_f), SMB and HML as dependent variables. The intercept is 0.317 with a significant *t*-statistic of 3.04. This indicates that the incremental risk premium for the PIN factor at 0.317% is actually higher than the unconditional risk premium of 0.22%. The second regression adds the momentum factor UMD as a dependent variable. When momentum is added, the intercept declines dramatically to 0.087 and is no longer significant (*t*-statistic = 1.00). Note that the coefficient on the UMD factor is highly significant (*t*-statistic = 10.64), causing the adjusted R^2 of the regression to increase from 7.61% for the 3-factor model to 38.45% for the 4-factor model. This analysis indicates that while PINF may be incrementally priced relative to the Fama–French 3-factor model, it is not incrementally priced when the 4-factor model is employed.[5] We turn to the two-pass cross-sectional tests next.

3.3. Two-pass cross-sectional tests of returns on PIN factor loadings

In this section, we investigate whether there is a risk premium for PINF using two-pass cross-sectional tests. We first estimate factor loadings for PINF and other risk factors at the portfolio level. We then estimate a cross-sectional regression of returns on factor loadings to test whether the factor loadings are predictably related to returns.

3.3.1. Estimating factor loadings

We estimate factor loading using the same portfolios (three PIN groups within each size decile) used in the creation of the PIN factor. Consistent with Easley and O'Hara (2004), we define the PIN factor differentially for each size decile as the equally weighted average of the value-weighted hedge returns (high PIN–low PIN) for all other deciles except the decile being considered. In other words, the PIN factor for size decile 1 is the average of the hedge returns for the 9 size

[5] Results are virtually identical if we use the equally weighted returns to calculate the PIN Factor or if we use hedge returns based on quintiles of PIN within size deciles instead of the 3 PIN groups used.

deciles 2–10, the PIN factor for size decile 2 is the average of the hedge returns for 9 size deciles 1, 3–10 and so on. We denote this modified PIN factor as PINF_{-1}.

We then run the following times-series regression at the portfolio level using the entire time period, i.e. 228 months from January 1984 to December 2002:

$$R_i - R_f = \alpha_i + \beta_i(r_m - r_f) + s_i\mathrm{SMB} + h_i\mathrm{HML} + p_i\mathrm{PINF}_{-1} + e_i \tag{3}$$

Table 4 (Panel A) reports the estimates for each of the coefficients for all 30 portfolios and their t-statistics, along with the adjusted R^2. An examination of the factor loadings reveals two interesting patterns. First, 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 (size decile 1), the average factor loading increases from -1.661 for low PIN firms to -0.550 for high PIN firms, while for large firms (size decile 10), the average factor loading increases from 0.086 for low PIN firms to 0.291 for high PIN firms. Thus, as expected, given the construction of the PIN factor, PIN factor loadings do increase across groups based on PIN.

Second, the PIN factor loading for small firms (size deciles 1 and 2) across all the three PIN groups is negative whereas for large firms (size deciles 9 and 10), most of the PIN factor loadings are 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 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.

One potential explanation for this perverse result is that the omission of the UMD factor distorts the loadings on the PINF factor. To investigate this, we report the loadings on PINF after inserting the UMD factor in panel B. However, the basic counter-intuitive finding of negative (positive) loadings on PINF for small (large) firms remains.

3.3.2. Second pass regressions

Next, we conduct monthly Fama and MacBeth (1973) cross-sectional regressions of returns on factor loadings 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 run these tests at the portfolio level instead of at the firm level. Khan (2008) 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 a combination of size and PIN. In the first stage, we estimate conditional factor loadings of portfolio returns on the Fama–French risk factors augmented by PINF. In the second stage, we estimate regressions of the portfolio returns on the estimated factor loadings (Table 5).

We first conduct this analysis both for the Fama–French 3-factor model (R_m–R_f, SMB, HML) augmented with PINF, as well as for the Fama–French 4-factor model (R_m–R_f, SMB, HML, UMD) augmented with PINF. We use four different sets of portfolios: 10 size deciles, 10 size deciles with three PIN groups within each decile, 10 size deciles with 5 PIN groups within each decile, and finally 10 size deciles with 3 PIN groups within each decile and 3 further groupings based on firm-level PIN loading. Note that the second portfolio grouping corresponds to how the PINF factor was created. Finally, we run cross-sectional generalized least-squares regressions, weighting each portfolio by the inverse of the residual covariance matrix, consistent with Cochrane (2005). The cross-sectional regressions are run for each of the 228 months from January 1984 to December 2002, and the parameters are averaged and t-statistics estimated using the Fama and MacBeth (1973) procedure.

Panel A of Table 6 presents the results for the Fama–French 3-factor model, augmented with PINF. For the 10 size portfolios, the loading on PINF (LPIN_3) is significantly negative. If PINF were a risk factor, we would expect a positive coefficient. This anomalous result is probably related to the unexpected behavior of factor loadings across size groups discussed earlier. Specifically, small firms tend to have the most negative loadings, despite earning higher returns, and this leads to a negative coefficient on PINF factor loadings. When we use 10 size deciles and 3 PIN groups within deciles, the negative loading on PINF does disappear, but PINF fails to load significantly. Indeed the coefficient on LPIN_3 (0.067) is less than one-third of the unconditional mean of PINF (0.22%). The results are almost identical when we use 5 PIN groups within each size decile. Finally, we estimate factor loadings at the firm-level by running firm-level time-series returns regressions. Firms within each size and PIN grouping (10 size $\times$ 3 PIN) are divided into three further sub-groups based on PIN loading. Our final cross-sectional regression uses the portfolios based on size, PIN and PIN loading. Here too, the coefficient on PIN (LPIN_3) is an insignificant 0.143, which is slightly more than half the unconditional mean of PINF (0.22%).

In Panel B, we repeat the analysis for the Fama–French 4-factor model, augmented with PINF. The results remain essentially unchanged, indicating that the presence or absence of momentum is not responsible for the weak performance of the PIN factor.

Turning to the other factor loadings, the coefficient on HML loadings (LHML_3 or LHML_4) is generally positive and significant. The coefficient on market returns (LMKT_3 or LMKT_4) is insignificant in all specifications, consistent with a whole range of papers starting with Fama and French (1992). Finally, the coefficient on size (LSMB_3 or LSMB_4) is often negative. The negative loading on SMB may be related to the fact that the composition of firms for which PIN data is available tends to be large firms which have a negative loading on SMB by construction. More important, PIN factor

Table 4
Fama–French time-series portfolio regressions by size-PIN groupings.

Panel A: Model $R_i - R_f = \alpha_i + \beta_i(r_m - r_f) + s_i\text{SMB} + h_i\text{HML} + p_i\text{PINF}_{-1} + \varepsilon_i$

Size decile	PIN Group	α_i	β_i	s_i	h_i	p_i	$t(\alpha_i)$	$t(\beta_i)$	$t(s_i)$	$t(h_i)$	$t(p_i)$	Adj. R^2 (%)
1	Low	0.428	0.940	1.201	0.437	−1.661	0.98	8.55	8.99	2.68	−5.70	51.8
1	Medium	0.271	0.931	1.054	0.677	−1.367	0.72	9.80	9.14	4.80	−5.44	52.2
1	High	0.622	0.750	0.994	0.679	−0.550	1.97	9.39	10.24	5.72	−2.60	48.4
2	Low	−0.466	0.876	0.975	0.297	−1.492	−1.50	11.00	10.17	2.53	−7.20	63.2
2	Medium	−0.414	0.964	1.042	0.532	−0.399	−1.74	15.78	14.18	5.91	−2.51	70.7
2	High	0.079	0.820	0.693	0.586	−0.240	0.37	14.94	10.49	7.25	−1.68	61.7
3	Low	−0.503	1.000	0.937	0.578	−0.801	−2.39	18.76	14.45	7.37	−6.03	76.3
3	Medium	−0.615	1.009	0.873	0.575	−0.241	−3.35	21.67	15.41	8.38	−2.08	78.3
3	High	−0.156	0.818	0.788	0.573	0.059	−0.95	19.64	15.54	9.35	0.57	74.2
4	Low	−0.024	1.060	0.746	0.564	−1.048	−0.12	22.00	12.78	7.98	−8.49	80.5
4	Medium	−0.258	1.027	0.754	0.626	−0.530	−1.44	22.56	13.67	9.38	−4.54	79.0
4	High	−0.187	0.985	0.716	0.651	0.097	−1.13	23.51	14.11	10.58	0.90	77.7
5	Low	−0.290	1.107	0.711	0.690	−0.509	−1.66	25.04	13.24	10.67	−4.74	81.1
5	Medium	−0.295	1.103	0.775	0.656	−0.024	−1.88	27.68	16.02	11.27	−0.24	83.3
5	High	−0.291	1.015	0.719	0.621	0.172	−1.83	25.18	14.68	10.53	1.76	79.8
6	Low	−0.371	1.137	0.543	0.684	−0.264	−2.39	28.73	11.30	11.87	−2.81	82.9
6	Medium	−0.222	1.135	0.613	0.647	−0.044	−1.38	27.62	12.28	10.81	−0.45	81.8
6	High	−0.189	1.072	0.604	0.521	0.256	−1.37	30.58	14.18	10.20	3.06	84.7
7	Low	−0.088	1.055	0.343	0.677	−0.269	−0.62	29.02	7.81	12.85	−3.16	82.2
7	Medium	−0.278	1.186	0.393	0.683	0.087	−1.64	27.42	7.53	10.90	0.86	79.4
7	High	−0.196	1.078	0.411	0.523	0.182	−1.26	27.13	8.56	9.09	1.95	79.8
8	Low	−0.137	1.078	0.136	0.587	−0.127	−1.00	30.68	3.22	11.56	−1.58	83.0
8	Medium	−0.103	1.151	0.140	0.541	−0.037	−0.66	29.06	2.94	9.44	−0.41	81.4
8	High	−0.146	1.138	0.167	0.491	0.159	−1.02	30.95	3.77	9.24	1.88	83.0
9	Low	−0.124	1.055	0.044	0.630	−0.289	−0.97	32.23	1.13	13.40	−3.79	84.7
9	Medium	−0.084	1.121	0.075	0.517	0.050	−0.67	35.06	1.95	11.25	0.67	86.2
9	High	−0.040	1.130	0.105	0.329	0.147	−0.34	37.52	2.90	7.59	2.09	88.5
10	Low	0.037	0.947	−0.314	0.258	0.086	0.35	34.73	−9.56	6.55	1.39	87.3
10	Medium	−0.005	0.996	−0.241	0.258	0.131	−0.06	42.33	−8.49	7.59	2.45	90.9
10	High	−0.013	1.046	−0.118	0.165	0.291	−0.13	40.59	−3.79	4.44	4.97	90.3

Panel B: Model $R_i - R_f = \alpha_i + \beta_i(r_m - r_f) + s_i\text{SMB} + h_i\text{HML} + m_i\text{UMD} + p_i\text{PINF}_{-1} + \varepsilon_i$

Size decile	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 (%)
1	Low	0.607	0.941	1.248	0.447	−0.285	−1.150	1.39	8.66	9.36	2.77	−2.51	−3.26	52.9
1	Medium	0.410	0.932	1.090	0.685	−0.222	−0.969	1.09	9.90	9.44	4.90	−2.26	−3.18	53.1
1	High	0.697	0.750	1.013	0.683	−0.120	−0.335	2.18	9.41	10.36	5.77	−1.44	−1.30	48.6
2	Low	−0.257	0.884	1.021	0.310	−0.299	−0.956	−0.84	11.41	10.85	2.72	−3.70	−3.85	65.2
2	Medium	−0.338	0.967	1.059	0.537	−0.109	−0.203	−1.40	15.89	14.35	5.98	−1.72	−1.04	70.9
2	High	0.241	0.826	0.728	0.597	−0.231	0.176	1.15	15.60	11.33	7.64	−4.20	1.04	64.4
3	Low	−0.358	0.999	0.973	0.574	−0.212	−0.454	−1.73	19.36	15.33	7.56	−3.98	−2.92	77.8
3	Medium	−0.540	1.009	0.892	0.573	−0.110	−0.062	−2.92	21.87	15.73	8.43	−2.31	−0.44	78.7
3	High	−0.104	0.818	0.801	0.572	−0.076	0.183	−0.62	19.72	15.71	9.37	−1.77	1.47	74.5
4	Low	0.091	1.064	0.784	0.564	−0.196	−0.707	0.49	22.83	13.70	8.25	−4.07	−4.85	81.8
4	Medium	−0.090	1.032	0.809	0.627	−0.287	−0.031	−0.54	24.78	15.82	10.26	−6.66	−0.24	82.4
4	High	−0.125	0.987	0.737	0.651	−0.106	0.281	−0.76	23.82	14.48	10.71	−2.47	2.17	78.2
5	Low	−0.174	1.110	0.749	0.685	−0.190	−0.191	−1.02	26.06	14.27	11.00	−4.27	−1.50	82.4
5	Medium	−0.182	1.106	0.812	0.652	−0.186	0.287	−1.20	29.02	17.29	11.70	−4.66	2.52	84.7
5	High	−0.182	1.018	0.755	0.617	−0.180	0.473	−1.17	26.28	15.82	10.89	−4.43	4.08	81.4
6	Low	−0.187	1.140	0.599	0.671	−0.274	0.160	−1.33	32.29	13.77	13.05	−7.63	1.59	86.4
6	Medium	−0.086	1.138	0.654	0.637	−0.204	0.271	−0.55	29.20	13.63	11.23	−5.14	2.44	83.6
6	High	−0.114	1.073	0.626	0.516	−0.111	0.428	−0.83	31.24	14.80	10.30	−3.18	4.36	85.3
7	Low	0.020	1.062	0.381	0.670	−0.183	0.030	0.14	30.77	9.01	13.41	−5.09	0.30	84.0
7	Medium	−0.129	1.195	0.447	0.673	−0.254	0.501	−0.81	29.74	9.06	11.57	−6.06	4.31	82.3
7	High	−0.078	1.086	0.453	0.516	−0.201	0.510	−0.52	28.79	9.79	9.45	−5.11	4.67	81.9
8	Low	−0.025	1.085	0.173	0.582	−0.182	0.158	−0.19	32.70	4.26	12.13	−5.34	1.70	84.8
8	Medium	−0.004	1.157	0.172	0.536	−0.160	0.215	−0.02	30.20	3.69	9.67	−4.08	2.00	82.6
8	High	−0.063	1.143	0.194	0.487	−0.134	0.369	−0.44	31.91	4.43	9.40	−3.64	3.67	83.9
9	Low	0.013	1.063	0.088	0.616	−0.213	0.044	0.11	35.85	2.43	14.44	−7.09	0.53	87.5
9	Medium	−0.001	1.126	0.101	0.508	−0.129	0.252	−0.01	36.41	2.69	11.43	−4.11	2.89	87.1
9	High	0.027	1.134	0.126	0.321	−0.104	0.310	0.23	38.55	3.52	7.60	−3.49	3.74	89.0
10	Low	0.095	0.951	−0.294	0.254	−0.093	0.225	0.91	35.68	−8.99	6.62	−3.47	3.10	87.9
10	Medium	0.008	0.997	−0.236	0.257	−0.020	0.162	0.08	42.31	−8.18	7.56	−0.87	2.52	90.8
10	High	−0.059	1.044	−0.134	0.168	0.072	0.184	−0.59	41.08	−4.30	4.58	2.81	2.65	90.6

This table using the 30 size-PIN portfolios described in Table 2. Value-weighted monthly excess returns for these portfolios (R_i-R_f) are regressed against factors for the market (r_m–r_f), size (SMB), book-to-market (HML), momentum (UMD) and PIN (PINF_{-1}). In Panel A, we supplement a Fama–French 3-factor model (r_m–r_f, SMB, HML) with PINF. In Panel B, we supplement a Fama–French 4-factor model (r_m–r_f, SMB, HML and UMD) with PINF_{-1}. PINF_{-1} is the equally-weighted average of the value-weighted hedge returns (high PIN–low PIN) using all deciles except the decile being analyzed (see page 15 for details). The time period is 1984–2002.

Table 5
Monthly cross-sectional portfolio regressions of returns on factor loadings.

Panel A: Regressions using loadings from Fama–French 3-factor model with PINF added

Portfolios	Intercept	$LMKT_3$	$LSMB_3$	$LHML_3$	$LPIN_3$	Adj. R^2 (%)
10 Size	0.031 (0.03)	−0.073 (−0.05)	−1.140 (−2.85)	1.774 (1.69)	−0.588 (−2.25)	59.1
10 Size∗3 PIN	1.073 (1.93)	−0.658 (−1.06)	−0.419 (−1.54)	0.637 (1.66)	0.067 (0.55)	42.9
10 Size∗5 PIN	0.944 (2.08)	−0.574 (−1.08)	−0.449 (−1.65)	0.757 (2.04)	0.062 (0.51)	33.0
10 Size∗3 PIN∗3 PIN loading	1.435 (3.56)	−1.011 (−2.01)	−0.375 (−1.38)	0.707 (2.09)	0.143 (1.21)	20.9

Panel B: Regressions using loadings from Fama–French 4-factor model with PINF added

Portfolios	Intercept	$LMKT_4$	$LSMB_4$	$LHML_4$	$LUMD_4$	$LPIN_4$	Adj. R^2(%)
10 Size	−2.810 (−1.61)	3.662 (1.70)	−0.352 (−0.73)	0.004 (0.00)	1.839 (1.07)	−0.869 (−2.72)	60.6
10 Size∗3 PIN	1.046 (1.92)	−0.656 (−1.07)	−0.443 (−1.63)	0.780 (1.93)	0.269 (0.54)	0.059 (0.48)	43.1
10 Size∗5 PIN	0.867 (1.95)	−0.488 (−0.92)	−0.444 (−1.62)	0.919 (2.40)	0.666 (1.41)	0.066 (0.54)	33.8
10 Size∗3 PIN∗3 PIN loading	1.370 (3.68)	−0.864 (−1.77)	−0.352 (−1.26)	0.525 (1.61)	0.273 (0.72)	0.132 (1.17)	20.8

Value weighted monthly excess returns for these portfolios are regressed against factors for the market (r_m–r_f), size (SMB), book-to-market (HML), momentum (UMD) and PIN (PINF). The factor loadings for the Fama–French 3-factor model (r_m–r_f, SMB, HML) with PINF are labeled as $LMKT_3$, $LSMB_3$, $LHML_3$ and $LPIN_3$. The factor loadings for the Fama–French 4-factor model (r_m–r_f, SMB, HML and UMD) with PINF are labeled as $LMKT_4$, $LSMB_4$, $LHML_4$, $LUMD_4$ and $LPIN_4$. Loadings are estimated for a variety of portfolios: 10 size, 10 size∗3 PIN portfolios, 10 size∗5 PIN portfolios and finally 10 size∗3 PIN∗3 PIN loading portfolios, where PIN loadings are first estimated at the firm level. Cross-sectional regressions are run for each of the 228 months across the portfolios, using loadings estimated over the entire period. The table presents summary of these regressions with the dependent variable being the value weighted monthly excess returns for these portfolios using the Fama and MacBeth (1973) method. All regressions are run using the generalized least squares (GLS) methodology, where the portfolios are weighted by the inverse of the inverse of the residual covariance matrix.

loadings are unrelated to returns.[6] Overall, the results from the two-pass cross-sectional regressions are consistent with our multivariate time-series regressions using the Fama–French 4-factor model and cast doubt on whether PIN is a priced information risk factor.

3.4. PIN factor and the macro-economy

If PIN is a risk factor in an inter-temporal asset-pricing model such as Merton (1973), we would expect payoffs to PIN to reflect future macro-economic activities as these activities contain information about the future investment opportunity set of investors. Several papers such as Chen (1991), Liew and Vassalou (2000) and Chordia and Shivakumar (2006) rely on this intuition to test whether the Fama–French factors or price momentum are indeed risk factors. Using this approach, we

[6] As mentioned before, one potential econometric explanation for the results in Table 6 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, in untabulated work, we supplement the factor loadings with size, B/M and PIN characteristics but the coefficient on PIN factor loadings continues to be insignificant.

Table 6
Future GDP Growth on Fama–French factors and the PIN factor.

Model: $\text{RealGDPGrowth}_{t+1,t+12} = \alpha + \beta(r_m - r_f)_{t-11,t} + \text{sSMB}_{t-11,t} + \text{hHML}_{t-11,t} + \text{mUMD}_{t-11,t} + \text{pPINF}_{t-11,t} + \varepsilon_i$							
	α	β	s	h	M	P	Adj. R^2 (%)
Fama–French 3 factor	0.02565 (8.50)	0.04543 (3.18)	−0.0197 (−1.04)	0.02032 (1.29)			17.98
Fama–French 4 factor	0.03001 (11.75)	0.04527 (3.40)	−0.0127 (−0.63)	0.01135 (0.77)	−0.0313 (2.09)		24.69
PINF	0.03042 (13.12)					−0.0069 (−0.20)	−1.32
Fama–French 3 factor and PINF	0.02558 (8.97)	0.04558 (3.29)	−0.0195 (−1.00)	0.02047 (1.35)		0.00213 (0.07)	16.78
Fama–French 4 factor and PINF	0.02967 (11.75)	0.04639 (3.65)	−0.0109 (−0.49)	0.01205 (0.85)	−0.0328 (−2.24)	0.01594 (0.52)	24.03

This table presents the regression coefficients from regressing 12-month-ahead growth in real GDP on the Fama–French factors and the PIN factor. RM_RF (market), SMB (size), HML (book-to-market), UMD (momentum) and PINF (PIN) are annually compounded from the monthly factors. The regression uses quarterly data, since data on GDP is available only on a quarterly basis. GDP data is obtained from the US Bureau of Economic Analysis. The dependent variable is the continuously compounded growth in real GDP over months t to t+12. Since the regressions use overlapping data, the t-statistics, which are reported in parentheses, are based on Newey–West standard errors. We have PIN Data for 4 quarters of 19 years, however we lose the first three observations, as we do not have adequate information for annual compounding. Our regression hence uses 73 quarterly observations from December 1984 (Q4, 1984) to December 2002 (Q4, 2002).

run a regression of future GDP growth on lagged values of the Fama–French factors and the PIN factor. The dependent variable is the continuously compounded growth in real GDP over months t+1 through t+12, while the explanatory variables are the value-weighted excess market return, SMB, HML and PINF, compounded over months $t-11$ through t. As GDP data are available only at a quarterly frequency, our regressions use quarterly data. Due to overlapping data, t-statistics are based on the autocorrelation-consistent Newey–West standard errors.

Table 6 presents the results. Over the sample period December 1984 to December 2002, the coefficient on PINF is insignificant in every specification reported. Consistent with Chordia and Shivakumar (2006) the coefficient on value-weighted market return is the only significant coefficient on a purported risk factor in the regressions. Thus, PINF appears to fail a macro-economic test of whether it is a risk factor.

4. Ex-ante cost of capital

4.1. PIN and ex-ante cost of capital

Another potential way to assess 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 cost of capital seems reasonable in our context for at least three reasons. First, one can 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 et al., 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 and Juettner-Nauroth (2005) model, commonly referred to as the OJ model, to infer the implied cost of capital. This model parsimoniously provides us with closed form estimates of the implied cost of capital without any assumptions about long run industry profitability. The OJ model calculates the cost of capital (r_e) as follows:

$$r_{\mathrm{e}} = A + \sqrt{A^2 + \frac{\mathrm{eps}_1}{P_0}(g_2 - (\gamma - 1))}$$

where

$$A = \frac{1}{2}\left((\gamma - 1) + \frac{\mathrm{dps}_1}{P_0}\right) \quad \text{and} \quad g_2 = \frac{(\mathrm{eps}_2 - \mathrm{eps}_1)}{\mathrm{eps}_1}$$

where eps_1 and eps_2 are consensus estimates of 1-year-ahead and 2-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 γ is the estimate of the long run economy-wide growth rate. Consistent with the Gode and Mohanram (2003) implementation of the OJ model, we use the average of short-term growth rate ($\mathrm{eps}_2/\mathrm{eps}_1-1$) and analyst 5-year earnings growth forecasts as our measure of g_2. Further, we set $\gamma-1 = r_{\mathrm{f}}-3\%$, 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 $\mathrm{RP_{OJ}}$.[7]

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 18,702 observations, or around 47% of the sample. Panel A of Table 7 provides descriptive statistics for the sub-sample with valid information, partitioned first into three groups of size and within each size group into three groups based on PIN. If PIN was indeed associated with the implied cost of capital, we would expect to see a positive trend in $\mathrm{RP_{OJ}}$ as PIN increases within each size group. However, we fail to see any evidence of a positive trend in $\mathrm{RP_{OJ}}$ across PIN groupings in any of the size terciles. The difference in $\mathrm{RP_{OJ}}$ between high PIN and low PIN stocks is actually negative for small and medium-sized firms. We do see an insignificant positive trend for large firms, but the trend is neither significant nor monotonic. These results are inconsistent with other earlier univariate results pertaining to returns, where we see weak evidence of a spread for small firms and no spread for large firms. These results are also counter-intuitive, as one would expect PIN to matter more for small stocks.

Panel B of Table 7 presents the correlations of $\mathrm{RP_{OJ}}$ with PIN and other measures of risk and information asymmetry 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 1-year-ahead EPS estimates scaled by price; analyst following, proxied for by the log of the number of analysts issuing 1-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 (LDM); growth, proxied for by analysts' consensus long-term growth forecasts (LTG), and finally, the log of the book-to-market ratio (LBM). Given the strong negative correlation between PIN and size, we present the correlations by size tercile.

Consistent with prior research, $\mathrm{RP_{OJ}}$ is positively correlated with factors related to systematic risk (BETA), unsystematic risk (DISP), leverage (LDM) and book-to-market (LBM) in each of the three size terciles. However, $\mathrm{RP_{OJ}}$ is uncorrelated with PIN in each of the three size terciles. One potential concern is that this lack of a relationship is driven by the fact that our sample is restricted to firms for whom analyst forecasts are available. This means that we are forced to eliminate precisely the sub-sample of firms where the impact of information asymmetry on stock prices is likely to be the strongest, leading to concerns about low power. While this is undeniably true, this does not explain why $\mathrm{RP_{OJ}}$ shows appropriate correlation with other measures of information asymmetry such as forecast dispersion (DISP) and analyst following (LNUM).

Panel C reports the results of firm-level regressions of our ex-ante cost of capital measures on PIN and the control variables. We run this regression both as a pooled panel regression as well as an annual cross-sectional regression using the Fama and MacBeth (1973) procedure. Given the strong correlation between firm size (LSIZE) and PIN, we run the regressions both with and without PIN. When we include firm size, we find an inverse relationship between PIN and ex-ante cost of capital in that the coefficient on PIN is −4.31 (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 firm size (LSIZE). When firm 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). 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 accruals-based accounting quality (AQ) measures that are gaining popularity in the accounting literature. Francis et al. (2004) demonstrate and Core et al. (2008) confirm that AQ measures are reliably

[7] 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 $\mathrm{re} = \sqrt{(g_2/(P_o/\mathrm{eps}_1))}$, 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.

Table 7
PIN and implied cost of capital.

Panel A: Mean RP_{OJ} by size and PIN groups

	Small size firms			Medium size firms			Large size firms		
PIN group	*N*	RP_{OJ} (%)	PIN (%)	*N*	RP_{OJ} (%)	PIN (%)	*N*	RP_{OJ} (%)	PIN (%)
Low	2069	9.88	16.40	2074	7.20	13.87	2072	5.65	10.56
Medium	2082	9.70	21.62	2087	7.02	17.78	2084	6.10	13.78
High	2078	9.75	28.93	2078	6.96	22.95	2078	5.80	18.15
High–Low		−0.13	12.53		−0.25	9.09		0.15	7.59
		(−0.68)	(98.60)		(−1.75)	(81.23)		(1.42)	(73.51)

Panel B: Correlation of RP_{OJ} with PIN and other risk factors

Size group	Correlation	BETA	DISP	LNUM	LSIZE	LDM	LTG	LBM	PIN
Small	Pearson	0.131	0.348	−0.026	−0.236	0.292	0.041	0.263	0.009
	Spearman	0.132	0.379	−0.011	−0.202	0.263	0.035	0.288	−0.006
Medium	Pearson	0.159	0.388	0.08	−0.074	0.254	0.002	0.241	−0.001
	Spearman	0.163	0.366	0.065	−0.034	0.26	0.004	0.251	−0.03
Large	Pearson	0.158	0.466	0.035	−0.115	0.271	−0.053	0.317	0.046
	Spearman	0.183	0.387	0.028	−0.103	0.303	−0.06	0.327	0.036

Panel C: Regression of RP_{OJ} on PIN and control variables

Method	Intercept	BETA	DISP	LNUM	LSIZE	LDM	LTG	LBM	PIN	Adj. R^2 (%)
Pooled	8.340	0.94	149.00	−0.22	−0.38	2.28	7.74	1.08	−4.31	26.4
	(24.08)	(12.47)	(34.96)	(−3.17)	(−9.00)	(20.14)	(11.22)	(15.96)	(−5.37)	
	6.05	0.97	155.06	−0.62		2.25	8.40	1.32	−0.94	26.0
	(26.08)	(12.93)	(36.74)	(−11.70)		(19.87)	(12.93)	(21.00)	(−1.32)	
Annual cross-sectional	7.28	1.00	217.24	0.12	−0.44	1.87	8.49	1.23	−1.88	29.7
	(5.82)	(6.12)	(9.00)	(0.80)	(−2.86)	(13.29)	(3.23)	(4.56)	(−1.77)	
	4.77	1.03	224.87	−0.40		1.89	9.70	1.47	1.18	28.9
	(10.67)	(5.57)	(9.09)	(−9.80)		(12.93)	(4.34)	(7.21)	(0.17)	

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 and 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. We are able to compute RP_{OJ} for 18,702 observations, or around 47% of the complete sample of 39,376 firm-years. Within the subset of data with valid RP_{OJ} information, 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. 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 1-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. The correlations are presented separately for each of the three size partitions. Panel C regresses RP_{OJ} on the control variables and PIN in both a pooled regression, as well as annual regressions using the Fama and MacBeth (1973) procedure. The requirement that at least two analysts follow a firm to calculate forecast dispersion as well as the data requirements for the control variables reduces the sample size to 15,319 for the regressions. The *t*-statistics for the annual regression are derived from the distribution of annual coefficients after controlling for auto-correlation using the procedure in Bernard (1995).

associated with ex-ante cost of capital estimates. In contrast, we show that PIN is either insignificantly related or negatively related to ex-ante cost of capital.

4.2. PIN factor loading and implied cost of capital

Finally, we run a regression of the implied cost of capital on factor loadings on PINF and the Fama–French factors and report the results in Table 8. We rely on the same methodology used earlier in the paper to generate conditional factor loadings both for the Fama–French 3-factor model augmented with PINF as well as the Fama–French 4-factor model augmented with PINF, using size deciles and 3 PIN groups within each decile. The factor loadings are calculated using only

Table 8
PIN factor loadings and implied cost of capital.

Panel A: Correlation of RP_{OJ} with factor loadings

Correlation	$LMKT_3$	$LSMB_3$	$LHML_3$	$LPIN_3$	$LMKT_4$	$LSMB_4$	$LHML_4$	$LUMD_4$	$LPIN_4$
Pearson	0.02	0.70	0.20	-0.32	0.05	0.71	0.20	−0.24	−0.20
Spearman	0.03	0.72	0.21	-0.25	0.07	0.72	0.21	−0.28	−0.09

Panel B: Regression of RP_{OJ} on factor loadings from Fama–French 3-factor model with PINF

Method	Intercept	$LMKT_3$	$LSMB_3$	$LHML_3$	$LPIN_3$	Adj. R^2 (%)	N
Pooled	3.744	2.305	2.656	−0.470	−0.916	40.99	510
	(5.80)	(3.82)	(16.89)	(−2.54)	(−4.77)		
Annual cross-sectional	5.129	0.944	2.688	0.039	−1.210	54.71	30 a year,
	(6.26)	(1.18)	(8.36)	(0.08)	(−6.80)		17 years

Panel C: Regression of RP_{OJ} on factor loadings from Fama–French 4-factor model with PINF

Method	Intercept	$LMKT_4$	$LSMB_4$	$LHML_4$	$LUMD_4$	$LPIN_4$	Adj. R^2 (%)	N
Pooled	3.940	2.043	2.662	−0.878	−2.380	−0.728	43.94	510
	(6.21)	(3.41)	(16.64)	(−4.25)	(−5.00)	(−3.72)		
Annual cross-sectional	5.047	0.829	2.389	0.182	−2.668	−1.102	56.25	30 a year,
	(5.81)	(1.04)	(8.07)	(0.40)	(−2.62)	(−9.22)		17 years

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 and 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. This information is available along with PIN for 14,060 firm-years between 1984 and 2002. At the beginning of each year from 1984 to 2002, these stocks are sorted into 10 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. For each group in each year, factor loadings are estimated using a Fama–French 3-factor model with PINF, a Fama–French 4-factor model with PINF and PINF alone. Factor loadings are estimated using at least 2 years and up to 5 years of lagged information. The factor loadings for the Fama–French 3-factor model (r_m–r_f, SMB, HML) with PINF are labeled as $LMKT_3$, $LSMB_3$, $LHML_3$ and $LPIN_3$. The factor loadings for the Fama–French 4-factor model (r_m–r_f, SMB, HML and UMD) with PINF are labeled as $LMKT_4$, $LSMB_4$, $LHML_4$, $LUMD_4$ and $LPIN_4$. Panel A presents the time-series averages of annual cross-sectional correlations between RP_{OJ} and factor loadings. Panels B and C present results of regressions of portfolio mean RP_{OJ} on factor loadings for the 3-factor model with PINF and 4-factor model with PINF respectively. Regressions are run both for the entire panel of data as well as annual cross-sectional regressions using the Fama and 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).

the subset of data for which ex-ante cost of capital can be estimated. Further, given that RP_{OJ} is measured every year, we estimate factor loadings at the portfolio level using up to 5 years of lagged returns, ensuring that at least 2 years of information are available.

Panel A of Table 8 presents the correlations between the mean portfolio-level RP_{OJ} and factor loadings. At the univariate level, RP_{OJ} is *negatively correlated* with PIN factor loading for both the 3-factor plus PINF model ($LPIN_3$) as well as the 4-factor plus PINF model ($LPIN_4$). Panel B presents regressions of mean portfolio-level RP_{OJ} on factor loadings for the 3-factor plus PINF model. The regressions are run both on the pooled panel of data, as well as annual cross-sectional Fama and MacBeth (1973) style regressions. In both specifications, the coefficient on the PINF factor loading ($LPIN_3$) is significantly negative, counter to what would be expected if PINF were priced. However, the coefficients on the market beta ($LMKT_3$) and factor loading related to SMB ($LSMB_3$) are positive and significant in both specifications while the coefficient on HML factor loading is negative for the pooled and insignificant for the annual regressions. Panel C repeats this analysis using the 4-factor plus PINF model, which also includes loadings for momentum ($LUMD_4$). While the momentum loadings also have a significantly negative coefficient, the strong inverse relationship between RP_{OJ} and PINF loading ($LPIN_4$) is unaltered. In summary, a regression of implied cost of capital on PIN factor loading does not support the hypothesis that PIN is a risk factor.

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 risk factor. These papers have been very influential in that empirical researchers in finance and especially in 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 findings of Easley et al. (2002) 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, unexpectedly, are unrelated (negatively associated) with returns (ex-ante cost of capital constructed from analysts' forecasts of earnings). Finally, there appears to be no robust relationship between PIN and ex-ante measures of cost of capital and between PIN factor returns and future GDP growth.

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.

A caveat is however in order. While our paper suggests PIN is not priced risk, it is difficult to make more general statements about the pricing of information risk since information risk can represent different notions to different researchers (e.g., estimation risk) and be proxied by different empirical variables. More work is clearly needed to address this open issue.

Appendix A. 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.'s (2002) 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:

Assertions that changes in PIN directly map into changes in cost of capital

- Duarte et al. (2008 JFE) in their Eq. (2) on page 36 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 Eq. (4) of their paper on page 38, 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 et al. (2004 JAE, p. 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 reduction in the *annual cost of equity capital*." (emphasis added).
- Brown and Hillegeist (RAST, 2007, p. 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 *reduction in the cost of capital*." (emphasis added).
- Chen et al. (2007a JAR, p. 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 et al. (RFS 2007b, p. 632) state: "PIN captures private firm-specific information impounded in stock price."
- In Table 6, Pan and Poteshman (RFS 2006, p. 900) find that put-call ratio 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, p. 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.'s (2002) result that PIN affects cost of capital

- Hilary (RAST 2006, p. 534): "For example, Easley et al. (2002) report that PIN is positively associated with spreads, the cost of capital."
- Francis et al. (2004 TAR, p. 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 et al. (JAE 2005, p. 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, p. 218, 219): "Easley et al. (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 et al. (JAR 2005, p. 653): "To be consistent with the idea of diversification in neoclassical asset-pricing theory, following Francis et al. (2005) and Easley et al. (2002), we estimate cost of capital using a factor model based on APT from ex post regressions."
- Botosan et al. (RAST, 2004, p. 235): "EHO document a strong positive association between averaged realized returns and PIN."
- Ecker et al. (2006 TAR, p. 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."

References

Aboody, D., Hughes, J., Liu, J., 2005. Earnings quality, insider trading, and cost of capital. Journal of Accounting Research 43, 651–673.
Aktas, N., de Bodta, E., Declerck, F., Oppens, H.V., 2007. The PIN anomaly around M&A announcements. Journal of Financial Markets 10 (2), 169–191.
Aslan, H., Easley, D., Hvidkjaer, S., O'Hara, M., 2006. The determinants of informed trading: implications for asset pricing. Working Paper, University of Houston, Cornell University and University of Maryland.
Asquith, P., Oman, R., Safaya, C., 2007. Short sales and trade classification algorithms. Working Paper, Massachusetts Institute of Technology.
Barberis, N., 2000. Investing for the long run when returns are predictable. Journal of Finance 55, 225–264.
Bawa, V., Brown, S., Klein, R., 1979. Estimation Risk and Optimal Portfolio Choice. North-Holland, Amsterdam.
Benos, E., Jochec, M., 2007. Testing the PIN variable. Working Paper, University of Illinois at Urbana Champaign.
Botosan, C.A., Plumlee, M.A., Xie, Y., 2004. The role of information precision in determining the cost of equity capital. Review of Accounting Studies 9 (2–3), 233–259.
Brennan, M.J., Wang, A., Xia, Y., 2004. Estimation and test of a simple model of intertemporal capital asset pricing. Journal of Finance 59 (4), 1743–1776.
Brooks, R., 1996. Changes in asymmetric information at earnings and dividend announcements. Journal of Business Finance and Accounting 23 (3), 359–378.
Brown, S., Hillegeist, S.A., 2007. How disclosure quality affects the level of information asymmetry. Review of Accounting Studies 12 (2–3), 443–477.
Brown, S., Hillegeist, S.A., Lo, K., 2004. Conference calls and information asymmetry. Journal of Accounting and Economics 37 (3), 343–366.
Campbell, J.Y., Vuolteenaho, T., 2004. Bad beta, good beta. American Economic Review 94, 1249–1275.
Chen, N.F., 1991. Financial investment opportunities and the macroeconomy. Journal of Finance 46, 529–554.
Chen, S., Shevlin, T., Tong, Y., 2007a. Does the pricing of financial reporting quality change around dividend changes? Journal of Accounting Research 45 (1), 1–40.
Chen, Q., Goldstein, I., Jiang, W., 2007b. Price informativeness and investment sensitivity to stock price. Review of Financial Studies 20 (3), 619–650.
Christophe, S., Ferri, M., Angel, J., 2004. Short-selling prior to earnings announcements. Journal of Finance 59, 1845–1875.
Chordia, T., Shivakumar, L., 2006. Earnings and price momentum. Journal of Financial Economics 80, 627–656.
Claus, J., Thomas, J., 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.
Cochrane, J.H., 2005. Asset Pricing. Princeton University Press, Princeton, NJ.
Core, J., Guay, W., Verdi, R., 2008. Is accruals quality a priced risk factor? Journal of Accounting and Economics 46, 2–22.
Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross-sectional variation in common stock returns. Journal of Finance 52, 1–33.
Davis, J., Fama, E.F., French, K.R., 2000. Characteristics, covariances, and average returns: 1929–1997. Journal of Finance 55, 389–406.
Dennis, D., McConnell, J., 1986. Corporate mergers and security returns. Journal of Financial Economics 16, 143–187.
Duarte, J., Young, L., 2008. Why is PIN priced? Journal of Financial Economics, forthcoming.
Duarte, J., Han, X., Harford, J., Young, L., 2008. Information asymmetry, information dissemination and the effect of regulation FD on the cost of capital. Journal of Financial Economics 87, 24–44.
Easley, D., O'Hara, M., 2004. Information and the cost of capital. Journal of Finance 59, 1553–1583.
Easley, D., Hvidkjaer, S., O'Hara, M., 2002. Is information risk a determinant of asset returns? Journal of Finance 57, 2185–2221.
Easley, D., Hvidkjaer, S., O'Hara, M., 2004. Factoring information into returns. Working Paper, Cornell University.
Easton, P., Monahan, S., 2005. An evaluation of accounting-based measures of expected returns. Accounting Review 80, 501–538.
Ecker, F., Kim, I., Olsson, P., Schipper, K., 2006. A returns-based representation of earnings quality. Accounting Review 81 (4), 749–780.
Ellul, A., Pagano, M., 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., 1970. Efficient capital markets: a review of theory and empirical work. Journal of Finance 25, 383–417.

Fama, E., 1991. Efficient markets: II. Journal of Finance 46, 1575–1617.
Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
Fama, E.F., MacBeth, J., 1973. Risk, return and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.
Francis, J., LaFond, R., Olsson, P., Schipper, K., 2004. Cost of equity and earnings attributes. Accounting Review 79 (4), 967–1010.
Francis, J., LaFond, R., Olsson, P., Schipper, K., 2005. The market pricing of accruals quality. Journal of Accounting and Economics 39, 295–327.
Frazzini, A., Lamont, O., 2006. The earnings announcement premium and trading volume. Working Paper, University of Chicago.
Gebhardt, W., Lee, C., Swaminathan, B., 2001. Toward an implied cost of capital. Journal of Accounting Research 39 (1), 135–176.
Gode, D., Mohanram, P., 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., Hou, K., Teoh, S.H., 2006. The accrual anomaly: risk or mispricing. Working Paper, University of California at Irvine.
Huddart, S., Ke, B., 2007. Information asymmetry and cross-sectional determinants of insider trading. Contemporary Accounting Research 24 (1), 195–232.
Hughes, J., Liu, J., Liu, J., 2005. Information, diversification, and cost of capital. Working Paper, University of California at Los Angeles.
Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. Journal of Finance 51, 3–53.
Jorion, P., 1985. International portfolio decisions with estimation risk. Journal of Business 58, 259–278.
Kandel, S., Stambaugh, R., 1996. On the predictability of stock returns: an asset-allocation perspective. Journal of Finance 51 (2), 385–424.
Keown, A., Pinkerton, J., 1981. Merger announcement and insider trading activity: an empirical investigation. Journal of Finance 36, 855–869.
Khan, M., 2008. Are accruals mispriced? Evidence from tests of an intertemporal capital asset pricing model. Journal of Accounting and Economics 45, 55–77.
Lambert, R., Leuz, C., Verrecchia, R., 2007. Accounting information, disclosure, and the cost of capital. Journal of Accounting Research 45, 385–420.
Lewellen, J., Shanken, J., 2002. Learning, asset-pricing tests, and market efficiency. Journal of Finance 57, 1113–1145.
Liew, J., Vassalou, M., 2000. Can book-to-market, size and momentum be risk factors that predict economic growth? Journal of Financial Economics 57, 221–245.
Litzenberger, R., Ramaswamy, K., 1979. The effect of personal taxes and dividends on capital asset prices: theory and empirical evidence. Journal of Financial Economics 7, 163–196.
Meulbroek, L., 1992. An empirical analysis of illegal insider trading. Journal of Finance 47, 1661–1699.
Merton, R., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.
Odders-White, E., Ready, M.J., 2006. Credit ratings and stock liquidity. Review of Financial Studies 19, 119–157.
Ohlson, J., Juettner-Nauroth, B., 2005. Expected EPS and EPS growth as determinants of value. Review of Accounting Studies 10, 349–365.
Pan, J., Poteshman, A.M., 2006. The information in option volume for future StockPrices. 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.
Petkova, R., 2006. Do the Fama–French factors proxy for innovations in predictive variables? Journal of Finance 61, 581–612.
Spiegel, M., Wang, X., 2005. Cross-sectional variation in stock returns. Liquidity and idiosyncratic risk. Working Paper No. 05-13, Yale University.
Stambaugh, R.F., 1999. Predictive regressions. Journal of Financial Economics 54 (3), 375–421.