

私有信息风险被市场定价了吗^{*}

——来自中国股市的证据

刘莎莎 孔东民 邢精平

[摘 要] 近来,一些研究指出私有信息交易是股票收益的决定因素之一,国内学者在研究中对其也日渐接受并采用。但是,私有信息交易在中国股市是否可以有效地作为一个风险定价因子,并没有明确的答案。采用 Fama-MacBeth 两阶段横截面检验,并结合多因子模型,本文发现,除市场、公司规模、账面市值比和惯性因子以外,知情交易的因子负载无法预测资产收益。这意味着,在中国股市采用私有信息或知情交易概率作为定价因子或信息交易含量测度值得进一步商榷。

关键词: 知情交易概率测度 两阶段横截面检验 定价因子

JEL 分类号: G10 G11 G12

一、导 言

传统的资本资产定价模型假定信息风险是可以完全被分散化的,因此信息风险不应该对资产的期望回报有所影响,见 Fama(1970,1991)。但是,基于知情交易概率(probability of informed trading,简称 *PIN*)测度,Easley,Hvidkjaer et al.(2002)和 Easley,Hvidkjaer et al.(2004)研究发现,*PIN* 的增加会带来股票回报的提高,基于 *PIN* 构造的零成本套利组合(在公司规模因素之外)能够获得统计上显著的回报。结合这些积极的证据,他们指出,私人信息风险是一个风险定价因子^①。

然而,很多金融学研究者指出信息风险是潜在的个股特质风险,从而使可以被分散化的,这甚至可以回溯到二十世纪六十年代,见 Fama(1970)。事实上,虽然 Easley,Hvidkjaer et al. (2002,2004)的研究在学界有较大影响,不少学者也采用其研究方法,但是近来许多研究却在质疑 Easley 的结论,如 Hughes, Liu et al.(2005)研究了信息风险在多因子定价模型中的角色并发现,信息风险或者是可以被分散化的,或者是可以被现存的其他因子替代的。Lambert,Leuz et al.(2007)也指出,Easley,Kiefer et al.(1996);Easley and O'Hara(2004)的结论是值得商榷的。具体而言,他们发现在交易者数量足够大的时候,信息风险并不一定被定价,并且这种风险是可以被完全分散化的。最近一些经验研究文献也对信息风险是否可以被定价提出质疑(具体见本文第二部分)。因此,在 *PIN* 被频繁使用的今天,对于其代理信息风险的有效性却存在质疑。我们相信,对于 *PIN* 测度在中国是否合用的经验性检验是有着重要意义。

全文的安排如下:第二部分讨论相关文献;第三部分给出研究方法和数据说明;第四部分报告检验结果并分析;最后给出结论。

* 刘莎莎,北京大学光华管理学院,硕士研究生;孔东民,华中科技大学经济学院,硕士生导师,经济学博士;邢精平,深圳证券交易所,管理学博士。本文得到国家自然科学基金(70803013; 70703024)的资助,谨致谢意。

① Easley and O'Hara(2002,2004)指出,系统的、不可以被分散化的信息风险是来自于知情交易者和不知情交易者之间的信息不对称。他们基于如下直觉构造了一个理论模型来说明为什么信息风险可以被定价:即在知情交易者和不知情交易者共存的市场上,当资产收益的信息是私有而非公开时,不知情交易者无法完全弄清价格包含的私有信息,他们需要资产有更高的回报。为了完全避免这种风险,不知情交易者将不得不仅仅持有无风险资产,而这相对于持有一部分风险资产来说显然是次优决策。由于不知情交易者是理性的,他们将会选择持有最优的分散化投资组合,但是无论他们如何分散其投资组合,不知情交易者都会被知情交易者所利用。

二、文献回顾

对于私有信息交易而言, Easley, Kiefer et al. (1996) (简称EKOP) 主要基于序贯交易模型提出了估计知情交易概率 PIN 的模型, 用于衡量利用私有信息进行交易的概率。 PIN 反映了各类利用私有信息进行的交易 Jayaraman (2008), PIN 值越低, 知情交易概率越低, 信息对称程度越高。Easley, Hvidkjaer et al. (2002) 结合 Fama and French (1993) 有关资产定价体系的三因素模型, 按照 EKOP 提出的 PIN 模型计算出知情交易概率, 把它作为回归的第四个因子后, 发现知情交易概率与价格正相关而且显著性很高, 这说明知情交易概率越高, 所要求的风险补偿也越高, 所以, 他们认为 PIN 作为一种风险因子是有意义的。目前, 大量研究引用了这一发展较为成熟的测度代理知情交易概率。

自从 EKOP (1996) 提出知情交易概率的估计方法后, PIN 测度已经在文献中得到广泛运用。Mohanram and Rajgopal (2009) 对近期和 PIN 有关的一些相关研究进行了较好的总结。他们从三个方面进行了回顾, 一是将 PIN 的变动与资本成本的变动相连接的研究。二是认为 PIN 被股票价格所反应。三是引用 Easley et al. (2002) 的结果, 认为更高的 PIN 伴随着更高的资本成本。

例如, Duarte, Han et al. (2008) 认为公司的资本成本是有关公司贝塔值、规模、账面市值比以及 PIN 的函数, 他们认为在公平披露规则颁布后, 公司的资本成本变化就是公司 PIN 因子负载的变化。Brown, Hillegeist et al. (2004) 对公司电话会议与 PIN 的关系进行考察, 发现平均而言, 一次电话会议会使下一季度的 PIN 值减少 0.59%。这意味着, 电话会议使得 PIN 相对于均值水平 18.24 减少了 3.2%。从而, 他们认为每季度主持一次电话会议可以减少权益资本成本的 0.15%。Brown and Hillegeist (2007) 也得到了类似的结论, 他们从公司全年披露质量的拟合概率系数发现, 随着公司披露质量得分高于中位数水平的概率的增加, PIN 值会降低。具体而言, PIN 值每减少 1.4%, 将会导致资本成本减少 0.35%。

Pan and Poteshman (2006) 发现, 在考虑与 PIN 的影响时, 看涨看跌比率可以更好地预测第二天经风险调整之后的股票回报。与此同时, 考虑看涨看跌比率与 PIN 的相互影响, 大大提高了看涨看跌比率的预测能力。Ellul and Pagano (2006) 发现, PIN 与 IPO 价格的低估有关。他们提出, 随着 PIN 从平均水平 0.286 增加到 0.42, IPO 价格就会低估 16%。Odders-White and Ready (2006) 证明了, 标准普尔信用评级与组成 PIN 测度的逆向选择负相关。

Hilary (2006) 提到: “Easley, Hvidkjaer et al. (2002) 报告中指出, PIN 与买卖价差, 也就是资本成本正相关。” Francis, LaFond et al. (2004) 认为, 有关知情交易风险与资本成本之间的正相关的经验性检验, 使用了不同类型的知情交易风险。例如集中考虑了知情者和不知情者之间的信息不对称, 他们使用了知情交易 (PIN) 得分概率。Francis, LaFond et al. (2005) 进一步提到: “私人信息 (由 PIN 得分来测度, 知情交易的一种市场微观结构测度) 越多和公共信息越少的公司的预期收益越高, Easley, Hvidkjaer et al. (2002) 找到了与此预期相一致的证据。”

Huddart and Ke (2007) 提到: “Easley, Hvidkjaer et al. (2002) 认为, 知情交易概率越大的股票会提供更高的回报, 借此补偿因为逆向选择而增加的风险。采用与本文的样本相重叠的公司和年度, 通过纽约证券交易所提供的日交易数据, 他们计算出了知情交易 (PIN) 概率, 并证明了 PIN 是一种经市场定价的风险。” Botosan, Plumlee et al. (2004) 提到: “Easley, Hvidkjaer, and O’Hara (简称 EHO) 证明了已实现的平均回报和 PIN 之间有着强烈的正相关关系。” Ecker, Francis et al. (2006) 提到: “在 Easley and O’Hara (2004) 模型中, 与信息不确定相关联的风险溢价是私人信息 (属于信息不对称) 以及公共、私人信息准确度的函数。”

国内对知情交易概率的研究, 主要有杨之曙、姚松瑶 (2004) 利用 PIN 测度发现交易活跃的股

票的知情交易概率比不活跃的股票要低一些。私人信息对交易不活跃的股票影响更大。他们进一步发现交易活跃与交易不活跃股票之间的买卖价格差可以用 *PIN* 来解释。刘元海、陈伟忠(2004)在序贯交易模型基础上对中国股市知情交易概率进行测定。王春峰、董向征和房振明(2005)将知情交易概率作为衡量信息不对称程度的一种直接指标,发现我国股市知情交易概率具有和流动性正相关,波动性负相关的特性。攀登、施东晖(2006)新建了一个关于知情交易概率的模型,发现知情交易确实会影响股价波动;知情交易者会被好的市场质量吸引,但随后又会降低市场质量;知情交易概率的日内行态为倒 J 型。王展翔(2007)发现我国股市具有较高的信息不对称,但由于流动性较高且知情交易者对坏消息的处理能力有限,所以总体上知情交易比率并不高。韩立岩、郑君彦和李东辉(2008)发现 *PIN* 对价格变动存在着影响,但对收益却产生了负效应。张宗新(2008)也基于 *PIN* 测度,发现在我国股权分置改革基本完成的背景下,上市公司很多重大事件的背后都有着私人信息交易行为,存在内幕交易的可能性很高。

就整体而言,国内基于 *PIN* 的研究也发现了知情交易概率测度可以在一定程度上代理知情交易行为,目前大多数研究都是在基于金融市场微观结构的角度下进行的,而且均是在承认知情交易概率可以作为风险定价因子的前提下进行的。

但是,最近学术界出现了一些反对 *PIN* 可以测度知情交易概率的观点。例如,Mohanram and Rajgopal(2009)认为,他们只能在 EKOP(1996)的样本期间内发现同样的证据,但在其它样本期间,结论就未必尽然,他们对 *PIN* 和资产回报的关系持怀疑态度。Duarte and Young(2009)检验了 *PIN* 是否被信息不对称或者是流动性因素定价。他们首先建立一个模型,将 *PIN* 分解成两部分,一部分与信息不对称相关,另一部分与流动性相关。在 Fama-MacBech 回归中发现与信息不对称相关的部分没有被定价,而与流动性相关的部分被定价。因此他们推断,只有与信息不对称无关的流动性因素解释了 *PIN* 与预期回报横截面的关系。由此,*PIN* 在代理知情交易概率方面就显得效力低下了。此外,Benos and Jochev(2007)也发现盈余公告之前的知情交易概率要比公告之后的更低,这似乎也与 *PIN* 反映知情交易概率的原意相左。

孔东民、申睿(2008)也利用 *PIN* 测度发现中国股市信息效率较低股票有更高的知情交易。他们同时指出,在不同时期,*PIN* 的代理能力存在明显差异。王茂斌、彭晴、孔东民(2010)发现共同基金参与那些具有更高知情交易的股票反而给其带来显著的损失。这也在一定程度上说明,知情交易有可能并不是一个较好的风险定价因子。

整体而言,目前还没有看到探讨 *PIN* 是否可以作为一个合理的风险因子的国内研究。对于这个在研究中应用较为广泛的信息风险测度变量,本文的研究表明,知情交易概率在中国股市作为风险定价因子,需要进一步商榷。

三、数据来源与研究设计

(一)数据来源说明

本文的研究样本期间为 1998 年 1 月至 2009 年 6 月,因为要利用上一年的数据构造组合、并利用至少两年的历史交易数据来计算资产贝塔系数以及需要用到未来一年的收益,所以有效的检验期为 2000 年 1 月至 2008 年 6 月。所有数据来自于两个数据库,其中,1999 年 1 月至 2004 年 12 月样本期的高频分笔数据来自于北京大学中国经济研究中心(简称 CCER)色诺芬(SinoFin)数据库,2005 年 1 月至 2008 年 6 月的高频分笔数据来自于锐思(RESET)数据库。所有的日频率数据和财务指标均来自于 CCER 色诺芬(SinoFin)数据库的价格收益数据库和财务数据库。研究样本为所有的 A 股上市公司的月度数据。

为保证数据的有效性并消除异常样本对研究结论的影响,在选择样本时,本文按照以下原则

处理数据:为消除 IPO 效应,IPO 上市当月不计入样本中;剔除了所有 ST、PT 公司(仅限于其被 ST 或 PT 的年份);剔除了账面市值比为负的公司;因为需要用到账面市值比变量,本文排除了金融类公司;最后,本文还利用 Winsorized 的方法对 0.05%和 99.5%的异常值进行整理。

(二)研究设计

1. 私有信息交易测度的计算:知情交易概率 PIN

知情交易概率测度最早是由 EKOP(1996)所提出,本文按照 Easley, Kiefer et al.(2008)的方法进行计算,该测度的基本原理是,认为市场的交易是由知情交易引起,并且服从泊松分布。在一系列假设下,可以根据订单的买卖方向和订单数量情况,从金融微观结构的角度来推测私人信息引致交易的发生概率问题。具体而言,PIN 一般是按照下面的步骤进行参数估计。首先,确定单位时间的似然函数:

$$L(\theta|B, S) = (1-\alpha)e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!} + \alpha \delta e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-(\mu+\varepsilon_s)} \frac{(\mu+\varepsilon_s)^S}{S!} + \alpha(1-\delta) e^{-(\mu+\varepsilon_b)} \frac{(\mu+\varepsilon_b)^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!} \quad (1)$$

其中, B 和 S 分别表示单位时间内的买单和卖单数量,其余的 5 个参数均为待定变量,即 $\theta=(\alpha, \mu, \delta, \varepsilon_b, \varepsilon_s)$ 需要利用极大似然法进行估计。这里, α 是信息事件发生概率, δ 是坏消息的概率, $1-\delta$ 是好消息的概率, μ 表示知情者提交订单的到达率, ε_b 表示非知情者提交买单到达率, ε_s 表示非知情者提交卖单到达率。在推断订单买卖方向的时候,本文采用 Lee and Ready(1991)的方法进行确定,即如果当前交易价格大于前一买卖报价的中位数,则认为此交易属于买方发起的交易,否则,则认定为卖方发起的交易。如果在此准则下无法加以判断,则再往前追溯一笔交易直至可以加以判断为止;如果交易无法按照 Lee-Ready 算法识别,本文则将该类订单做抛弃处理,Campbell, Ramadorai et al.(2009)也采用了这样的处理方法,研究中发现,这一类的订单比例极小(低于 0.5%)。事实上, Lee and Radhakrishna(2000)发现 Lee-Ready 算法是极为精确的。

在假设每个交易日的消息相互独立的情况下,本文可以很容易地给出一段时期内(设为 I)的似然函数,如下:

$$L(\theta|M) = \prod_{i=1}^I L(\theta|B_i, S_i) \quad (2)$$

显然,对两边取对数,在不改变单调性的情况下,可以表达为求和的形式,在此基础上,本文根据优化程序很容易计算出 $(\alpha, \mu, \delta, \varepsilon_b, \varepsilon_s)$ 。然后求得 PIN 为:

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s} \quad (3)$$

需要指出的是,在计算中本文发现中国股市的单向订单数量较大,这使得最大似然估计存在着数值溢出的问题,这导致结果不稳定,为了解决这一问题,本文将中国股市的订单数量同比缩小 10 倍,这对 PIN 的结果没有影响,而计算的精度和计算拟合的稳定性则得到了极大的增强,本文在最大似然估计中,每一只股票都选了 27 对组合(即 $\alpha\mu, \varepsilon_b$ 和 ε_s 分别用[0.25, 0.5, 0.75]两两交叉)作为初始值,发现结果极为稳定,基本上不受初始值的影响。

在估计中,本文采用月度的高频分笔数据对 PIN 进行估计(需要该数据的读者,请通过 Email 与作者联系),尽管 Easley, Hvidkjaer et al.(2002)用了年度的 PIN,然而也有其他学者采用不同的期间估计 PIN 测度,例如 Zhao and Chung(2006)利用了 59 天以及 64 天的交易数据;Vega(2004)利用了 40 天的交易数据;Henry(2005)利用了月度交易数据估计;Fuller, Van Ness et al.(2010)则利用了季度交易数据估计。

2. 各检验因子的构造及说明

市场收益率(简称 MR)根据 A 股市场所有股票当月的流通市值计算加权市场平均收益率,其

中剔除了 IPO 当月的交易数据。

规模因子(简称 SMB)和账面市值比因子(简称 HML)根据 Fama and French(1993)的三因子模型计算。具体而言,首先以上年度末每股权益与收盘价的比值来度量账面市值比(B/M)的高低;以当年度 6 月底流通市值来度量规模的大小。然后按照 Fama and French(1993),在每年 6 月底按流通市值的中位数样本分为大、小两组(B 和 S),按 B/M 分成高、中和低三组,即 H(30%)、M(40%)和 L(30%);两者交叉分组形成 6 个组合,由此计算当年度 7 月到下年度 6 月每个组合价值加权月收益率。HML 和 SMB 两个因子分别定义如下: $HML = (Ret_{SH} + Ret_{BH} - Ret_{SL} - Ret_{BL})/2$,表示高 B/M 与低 B/M 组合的收益差; $SMB = (Ret_{SL} + Ret_{SM} + Ret_{SH} - Ret_{BL} - Ret_{BM} - Ret_{BH})/3$,表示小规模公司和大规模公司组合的收益差。

在构造惯性因子(简称 MOM)时,本文参照 Carhart(1997)的研究,对于月度 t 的 MOM 而言,本文根据月度 $t-12$ 到 $t-1$ 月度的个股收益率排序,将收益率最高的 30% 的股票分为赢家组合,将收益率最低的 30% 的股票归类于输家组合,然后利用赢家组合的平均收益率和输家组合的平均收益率之差作为月度 t 的 MOM 因子。

知情交易因子(简称 PINF),被定义为所有 SIZE 组别下通过构造零成本组合而得到收益率。对于月度 t 的 PINF 而言,本文根据上年度 12 月的流通市值大小将样本分为 10 个规模组合。然后在各个规模组合内,本文基于 PIN 值进行第二次排序,并将样本分为 High, Medium, Low 三组。本文根据 $t+1$ 月度到 $t+12$ 月度的个股收益率,分别计算出在各个规模组合中高 PIN 与低 PIN 组合以流通市值为权重的组合收益率,继而计算出二者的差值,记为 High-Low。最后再求出 10 个规模组合的 High-Low 的均值,即可得到该月度的 PINF。

四、实证结果分析

(一)描述性统计

本文首先考察各个变量的描述性统计,具体如表 1 所示。其中,Panel A 部分报告了模型中各变量的均值、标准差以及 t 统计值,由此可以看出 PINF 因子的 t 统计值不显著。Panel B 部分报告了各变量的相关系数,对角线以上(下)部分为 Pearson(Spearman-秩)相关系数。由表 1 可知,PINF 与市场的超额收益率、HML 和 MOM 等变量之间表现了弱相关性。PINF 因子与 SMB 因子之间却表现了一定的负相关性,这与本文的直觉相一致,即公司规模越大,PIN 值越低,知情交易概率越低,信息对称程度越高,PIN 因子被定价的部分也就越低。

表 2 报告了有效样本期内各年满足要求的样本总数。根据不同标准划分组合时,可以得到在每一年各组合平均含有的样本数目,保证本文的结论是在足够的样本数目下获得的。

(二)基本检验结果

1. Esaley, Hvidkjaer 的估计方法

因为需要用到个股的贝塔值进行排序,本文使用两年的月度收益数据来估算个股的排序前组合的贝塔值,见 Esaley, Hvidkjaer et al.(2002), Mohanram and Rajgopal(2009)。对于每一只股票,本文使用前两年该股票数据对同期市场收益和滞后一期的市场收益进行回归。通过加总两个估计系数(即当期和滞后一期市场收益变量的回归系数),即可得到个股排序前的组合贝塔。在此,用贝塔进行排序是参照金融学研究习惯而定,尽管理论上可以随机抽取不同股票作为不同组合,但是习惯上认为按照贝塔值排序是一个更好的选择。

在每年的任一月份,基于所估计的个股排序前组合贝塔,本文将满足要求的样本分为 40 个组合,将组合的月度收益率定义为该组合所含有的个股月度收益率的等权平均数。然后,基于全样本期间,本文将各组合的月度收益率与同期市场收益和滞后一期的市场收益进行回归,通过加总两个回归系数,得到各组合的贝塔。

表 1 各风险因子基本统计

风险因子	<i>MR-Rf</i>	<i>SMB</i>	<i>HML</i>	<i>MOM</i>	<i>PINF</i>
<i>Panel A: 描述性统计</i>					
均 值	0.003	0.003	0.005	0.005	0.0002
标准差	0.084	0.044	0.027	0.056	0.015
t-stat	0.386	0.787	1.923	0.854	0.150
<i>Panel B: 相关系数</i>					
MR-Rf		0.153	0.29	-0.066	0.070
SMB	0.227		0.267	-0.699	-0.144
HML	0.281	0.132		-0.270	-0.015
MOM	-0.242	-0.670	-0.247		0.033
<i>PINF</i>	0.034	-0.146	0.042	0.028	

注：本表报告了主要变量的描述性统计(*Panel A*)和相关系数(*Panel B*)。其中,*Panel B*的对角线以上(下)部分为 Pearson(Spearman-秩)相关系数。*MR*为市场收益率,根据 A 股市场所有股票当月的流通市值计算加权市场平均收益率,其中剔除了 IPO 当月的交易数据。*SMB*为规模因子,*HML*为账面市值比因子。这两个因子根据 Fama and French(1993)的三因子模型计算。*MOM*为惯性因子,按照 Carhart(1997)的研究方法计算。*PINF*为知情交易因子,定义为所有 *SIZE* 组别下通过构造零成本组合而得到收益率。对于月度 *t* 的 *PINF* 而言,我们根据上年度 12 月的流通市值大小将样本分为 10 个规模组合。然后在各个规模组合内,基于 *PIN* 值进行第二次排序,并将样本分为 High, Medium, Low 三组。我们根据 *t*+1 月度到 *t*+12 月度的个股收益率,分别计算出在各个规模组合中高 *PIN* 与低 *PIN* 组合以流通市值为权重的组合收益率,继而计算出二者的差值,记为 High-Low,最后再求出该 10 个规模组合的 High-Low 的均值,即可得到该月度的 *PINF*。

表 2 每年进入样本的股票数

年 份	2000	2001	2002	2003	2004	2005	2006	2007	2008
样本数	825	936	1011	1044	1083	1180	1131	1139	1247

注：2000~2007 年样本数是以全年为度量时间,2008 年样本数是以该年 01 月~06 月为度量时间。

对于公司水平的横截面回归,公司在各年的贝塔系数均为公司在当年所归属的组合贝塔值。在样本期 2000 年至 2008 年的各月,本文进行如下的横截面回归:

$$R_{im}=\gamma_{0t}+\gamma_{1t}\beta_{it}+\gamma_{2t}PIN_{it-1}+\gamma_{3t}SIZE_{it-1}+\gamma_{4t}LBM_{it-1}+\varepsilon_{it}$$

这里, R_{im} 为个股 *i* 在 *t* 年度第 *m* 月份里的收益率, LBM_{it-1} 为通过对个股 *i* 在上一年度末的账面市值比进行调整得到的对数值, $SIZE_{it-1}$ 为个股 *i* 在上一年度末流通市值的对数值。 PIN_{it-1} 为个股 *i* 在上一年度末的知情交易概率。通过使用标准的 Fama-MacBeth 两阶段横截面检验,本文可以得到时间序列上的各横截面回归系数,继而得到回归方程中的各系数估计值及相应的 *t* 统计量。

估计结果见表 3。此表格中,本文复制了 Esaley, Hvidkjaer et al.(2002)的研究方法。容易发现,在 2000 年 1 月至 2008 年 6 月全样本期内,*PIN* 因子估计系数的均值即使在 10%的显著性水平下也不显著异于 0。

进而,为了检验在不同的子样本期间内 *PIN* 是否被定价,本文采用分样本时期进行相同的检验,每两年为一估计窗口。由于暂时仅得到各公司在截至 2008 年 6 月的 *PIN* 值,故 2008 年估计窗口仅含有 6 个月份的样本数据,估计结果如表 3 所示。通过观测估计结果,本文发现:在所有的子样本时期内,即使在 10%的显著性水平下,本文也无法拒绝 *PIN* 因子估计系数均值为 0 的假设,也就是说,在各子样本期内 *PIN* 因子均没有被定价。

2. *PIN-SIZE* 分组排序估计

参照 Mohanram and Rajgopal(2009)的研究,接下来本文在 *SIZE* 分组的基础上,进一步分离出

表 3 各样本期间的检验结果

	<i>Intercept</i>	<i>Beta</i>	<i>PIN</i>	<i>SIZE</i>	<i>LBM</i>	<i>Adj-R²(%)</i>
2000.01~2008.06	0.036	-0.012	0.013	-0.001	0.008	5.773
	0.772	-1.470	1.039	-0.490	3.496	
2000.01~2001.12	0.210	-0.001	-0.005	-0.009	0.007	5.397
	2.333	-0.054	-0.284	-2.274	1.306	
2002.01~2003.12	0.119	0.019	0.011	0.006	0.009	5.843
	-1.403	-1.361	0.591	1.511	2.396	
2004.01~2005.12	0.052	0.028	0.016	0.002	0.010	5.700
	0.646	-1.293	-0.850	-0.395	3.079	
2006.01~2007.12	0.053	0.036	0.029	0.003	0.005	6.359
	-0.415	2.165	1.513	0.615	0.717	
2008.01~2008.06	0.146	0.087	0.077	0.006	0.014	5.355
	1.015	-3.137	0.927	-1.053	1.707	

注：针对不同的估计窗口，本文使用标准的 Fama-MacBeth 两阶段横截面检验，得到各月基于公司水平的横截面回归结果。表中报告了月度回归系数的时间序列均值以及相应的 t 统计值。

PIN 对收益率带来的净影响。

首先，本文基于个股 *SIZE* 进行排序，然后在每个 *SIZE* 组内基于 *PIN* 值进行第二次排序。具体来说，在 *t* 年度 1 月份，本文依据 *t*-1 年度的流通市值将样本等分为 10 个组合。然后，在每一个组合内，根据 *t*-1 年度的 *PIN* 值将样本等分为 3 组，分别为 Low, Medium, High(在本文的后面部分，均采用该方法进行分组)。基于这种双向排序，本文进一步考察在控制了 *SIZE* 之后，*PIN* 对收益的影响。

表 4 报告了排序后不同组合的主要变量的比较结果，其中 Panel A 给出了各组合 *SIZE* 均值的描述性统计，Panel B 给出了各组合 *PIN* 均值的描述性结果。基于 Panel A 的检验结果，可以看到，在同一 *SIZE* 组别内，拥有不同 *PIN* 水平的组合的规模差异并不是显著异于 0，这一结论在绝大多数的 *SIZE* 组别里都成立。Panel B 的结果则说明，在每一个 *SIZE* 组别内，*PIN* 值差异在常规的显著性水平下均显著地异于 0。

本文还发现，随着规模的逐渐增大，组内的 *PIN* 值差异也随之增大。这与我们的直觉相违背，即公司规模越大，披露信息的渠道也就越多，信息对称程度越高，*PIN* 值越低，知情交易概率越低。此外，这一现象也与 Mohanram and Rajgopal(2009)对美国股市的检验结果不大一致。从某种程度上讲，这也间接证明了 *PIN* 在中国不是一个很好的代理变量。

表 5 报告了排序后不同组合收益的比较结果，其中 Panel A 给出了以一年以后的流通市值为权重而得到的各组合一年以后的月度收益率平均值，Panel B 给出了各组合一年以后的月度收益率的等权平均值。基于 Panel A 的估计结果，可以发现，对于给定的某一 *SIZE* 组别，在常规的显著性水平下，高 *PIN* 和低 *PIN* 组合的价值加权预期收益率并没有显著的差异。进一步，在对所有组合的价值加权预期收益率进行平均时，虽然高 *PIN* 组合比低 *PIN* 组合获得约 0.022% 的超额收益，但是这一结果却并不是显著的。本文通过观察 Panel B 也可以得到类似的结果，不同的是此时高 *PIN* 组合看似比低 *PIN* 组合获得的收益更低，同样，这一结果并不显著。

在这里，本文定义表 5 中的 Hig-Low 组的收益率均值之差为 *PINF* 因子。在接下来的检验中，本文将继续使用这一因子。

表 4 不同组合排序的主要变量比较

基于 <i>SIZE</i> 的组合	<i>Low PIN</i>	<i>Medium PIN</i>	<i>High PIN</i>	<i>Hig-Low</i>	<i>t-stat</i>
<i>Panel A: 公司规模</i>					
1	4.03E+08	3.93E+08	3.97E+08	-6.17E+06	-0.890
2	5.35E+08	5.52E+08	5.70E+08	3.43E+07	4.022
3	6.58E+08	6.63E+08	6.65E+08	7.25E+06	0.668
4	7.92E+08	7.96E+08	7.96E+08	4.32E+06	0.313
5	9.26E+08	9.60E+08	9.38E+08	1.19E+07	0.777
6	1.12E+09	1.10E+09	1.11E+09	-1.07E+07	-0.597
7	1.37E+09	1.40E+09	1.39E+09	2.13E+07	0.898
8	1.75E+09	1.81E+09	1.80E+09	4.38E+07	1.319
9	2.55E+09	2.67E+09	2.63E+09	7.84E+07	1.384
10	6.46E+09	6.58E+09	7.02E+09	5.62E+08	1.758
<i>Panel B: PIN</i>					
1	0.119	0.118	0.127	0.008	4.573
2	0.117	0.116	0.125	0.008	5.091
3	0.114	0.117	0.125	0.010	6.316
4	0.114	0.117	0.123	0.009	5.780
5	0.114	0.116	0.124	0.010	6.387
6	0.109	0.114	0.132	0.023	14.834
7	0.111	0.114	0.133	0.022	13.624
8	0.108	0.113	0.135	0.027	15.897
9	0.109	0.114	0.134	0.024	15.210
10	0.109	0.117	0.140	0.031	18.226

注：该表报告的样本期为 2000 年 01 月至 2008 年 06 月，按照上一年的 *SIZE* 和 *PIN* 值先后进行两次排序分组，其中 *Panel A* 报告了各组合 *SIZE* 均值的描述性统计，*Panel B* 报告了各组合 *PIN* 均值的描述性结果。

表 5 不同组合排序的收益

基于 <i>SIZE</i> 的组合	<i>Low PIN</i>	<i>Medium PIN</i>	<i>High PIN</i>	<i>Hig-Low</i>	<i>t-stat</i>
<i>Panel A: value-weighted (10⁻³)</i>					
1	-1.809	1.219	1.920	3.729	1.146
2	-1.276	4.725	-2.643	-1.367	-0.463
3	-1.108	-2.240	0.995	2.104	0.643
4	3.345	1.500	2.835	-0.510	-0.184
5	0.890	2.853	-0.227	-1.117	-0.373
6	1.996	-2.012	-1.024	-3.019	-1.122
7	1.923	2.019	-0.106	-2.029	-0.666
8	1.961	2.866	4.107	2.145	0.610
9	3.087	0.020	2.191	-0.900	-0.310
10	0.403	-0.185	3.538	3.136	0.674
Average	0.941	1.077	1.159	0.218	0.150
<i>Panel B: equal-weighted (10⁻³)</i>					
1	-8.363	-5.473	-5.868	3.729	1.146
2	-7.517	-3.925	-8.964	-1.367	-0.463
3	-6.432	-8.198	-6.407	2.104	0.643
4	-4.097	-6.355	-6.208	-0.510	-0.184
5	-6.611	-4.841	-7.897	-1.117	-0.373
6	-5.963	-8.365	-7.931	-3.019	-1.122
7	-5.850	-5.918	-7.863	-2.029	-0.666
8	-4.774	-4.816	-5.012	2.145	0.610
9	-5.398	-7.226	-8.598	-0.896	-0.310
10	-6.467	-7.356	-7.583	3.136	0.674
Average	-6.147	-6.247	-7.233	-1.086	-0.913

注：该表报告的样本期为 2000 年 01 月至 2008 年 06 月，其中 *Panel A* 给出了以一年后的流通市值为权重而得到的各组合一年以后的月度收益率平均值及相应的 *t* 统计值，*Panel B* 给出了各组合一年以后的月度收益率的等权平均值及相应的 *t* 统计值。

(三)时间序列检验

通过时间序列进行考察 *PINF* 是否可以被常见的因子所解释。具体而言,使用前面所描述的 *PINF* 因子,分别将其与简单的 CAPM 模型和 Fama–French 模型中的因子进行回归。需要做的是,检验回归的截距项系数是否显著地异于 0,不显著的截距项意味着 *PINF* 并没有被定价,显著的截距项则意味着,在剔除 *PINF* 与其他因子的相关关系以后,*PINF* 仍会导致一些溢价。本文按照如下模型进行检验:

$$PINF_{i,t}=\alpha_i+\beta_i(MR_t-R_{f,t})+s_iSMB_t+h_iHML_t+m_iMOM_t+e_{i,t}$$

表 6 基于各因子模型的时间序列检验

因子模型	<i>Intercept</i>	<i>MR-Rf</i>	<i>SMB</i>	<i>HML</i>	<i>MOM</i>	<i>Adj-R</i> ² (%)
CAPM	0.0002	0.012				-0.01
	0.123	0.698				
3-Factor	0.000	0.016	-0.053	0.0002		-0.02
	0.230	0.900	-1.530	0.0031		
4-Factor	0.001	0.018	-0.086	-0.008	-0.037	0.02
	0.440	0.990	-1.820	-0.140	-1.020	

注:该表报告了基于各种不同因子模型的时间序列检验,样本期为 2000 年 01 月至 2008 年 06 月。估计结果的第一行为回归系数,第二行为相应的稳健 *t*-统计值。

回归结果如表 6 所示,无论是采用 CAPM 模型中的市场超额收益(*R_m-R_f*)作为解释变量,还是采用 FF 三因子抑或是四因子作为解释变量,截距项的估计结果虽说为正,但是,在常规的显著性水平下这一结果并不显著。这意味着,*PINF* 可以被常见的因子所解释,也说明其本身并未体现出显著的溢价。

(四)Fama–MacBeth 两阶段横截面检验(two-pass cross-sectional test)

在这一部分,本文使用两阶段横截面检验来考察 *PINF* 的存在是否能够导致风险溢价。主要按照如下两步进行研究:第一步,在组合水平下估计出 *PINF* 和 FF 模型因子的因子负荷;第二步,采用横截面回归,将组合收益与上一阶段估计出得组合水平下的因子负荷进行回归,以检验因子负荷是否会影响投资收益。

1. 因子负荷估计

本部分依旧根据前文得到的分组进行估计 *PINF* 等各因子的因子负荷。在这里,与 Mohanram and Rajgopal(2009)和 Easley, Hvidkjaer et al.(2002)一致,将每一 *SIZE* 组别的 *PIN* 因子定义为其其他 9 个 *SIZE* 组别的价值加权的套利收益率(hedge return,即 High *PIN* – Low *PIN*)的等权平均数,并将修改以后的 *PIN* 因子标记为 *PINF*-1。即公司规模组合 1 的 *PINF* 为公司规模组合 2 至 10 的价值加权的套利收益率之等权平均。

然后,在全样本期内,也就是 2000 年 01 月至 2008 年 06 月,本文在组合水平下采用三因子模型进行如下时间序列回归分析:

$$R_i-R_f=\alpha_i+\beta_i(MR-R_f)+s_iSMB+h_iHML+p_iPINF_{-1}+e_i$$

回归结果如表 7 所示,该表报告了 30 个组合的回归系数、相应的 *t* 估计值以及调整以后的 *R*²。其中 Port 表示不同的组合,第一位数表示第几个 *SIZE* 组合,第二位数表示不同的 *PIN* 组合。例如 5L,表示为按照 *SIZE* 分组的第五组投资组合,且其 *PIN* 为 Low 的分组。本文定义,各解释变量的因子负载为该阶段估计所得到的各解释变量的回归系数。

可以看到,我国的情况表现为,*PIN* 的因子负载在每一个 *SIZE* 组别内并没有表现出随着 *PIN*

表 7 *PINF* 的因子负荷估计：三因子

<i>Port</i>	<i>Inter</i>	<i>MR</i>	<i>SMB</i>	<i>HML</i>	<i>PINF</i> ₋₁	<i>t(Inter)</i>	<i>t(MR)</i>	<i>t(SMB)</i>	<i>t(HML)</i>	<i>t(PINF</i> ₋₁)	<i>AdjR</i> ²
1L	-2.49	0.94	0.99	-0.04	-0.22	-0.93	28.55	15.55	-0.35	-1.26	0.93
1M	-2.68	1.00	1.08	-0.10	-0.34	-1.09	33.11	18.53	-1.09	-2.12	0.95
1H	-3.72	0.90	0.97	0.15	-0.51	-1.36	26.80	14.89	1.43	-2.84	0.92
2L	-5.16	1.00	0.93	0.06	-0.65	-1.83	28.64	13.91	0.56	-3.46	0.93
2M	-5.28	0.97	0.99	0.25	-0.71	-1.75	26.02	13.78	2.11	-3.53	0.92
2H	-1.10	0.97	0.92	0.00	-0.33	-0.47	33.98	16.62	0.02	-2.16	0.95
3L	-6.17	0.99	0.91	0.21	-0.31	-2.68	34.82	16.60	2.30	-2.06	0.95
3M	-3.65	1.00	0.78	0.34	-0.39	-1.52	33.91	13.68	3.66	-2.46	0.95
3H	-3.89	0.96	0.72	0.23	-0.62	-1.27	25.24	9.78	1.94	-3.03	0.90
4L	-4.43	0.95	0.75	0.36	-0.41	-1.83	31.88	12.97	3.83	-2.61	0.94
4M	-2.98	1.03	0.67	0.27	-0.43	-1.47	41.14	13.75	3.47	-3.29	0.96
4H	-4.49	0.99	0.62	0.37	-0.56	-1.63	29.30	9.45	3.46	-3.18	0.93
5L	-2.72	1.03	0.62	0.07	-0.61	-1.17	35.89	11.31	0.77	-4.04	0.94
5M	0.008	1.02	0.49	0.10	-0.56	0.04	39.14	9.72	1.22	-4.05	0.95
5H	-1.41	0.97	0.59	0.05	-0.35	-0.55	30.73	9.82	0.48	-2.13	0.93
6L	-4.22	1.00	0.54	0.42	-0.28	-1.84	35.38	9.92	4.79	-1.81	0.95
6M	-4.05	1.03	0.43	0.38	-0.54	-1.65	33.94	7.45	4.00	-3.28	0.94
6H	-2.70	1.01	0.45	0.13	-0.51	-1.08	32.73	7.52	1.37	-3.04	0.93
7L	-4.92	1.03	0.43	0.36	-0.54	-1.84	31.30	6.81	3.53	-2.99	0.93
7M	-1.62	1.04	0.26	0.23	-0.21	-0.79	41.57	5.27	2.87	-1.54	0.95
7H	-0.766	0.93	0.32	0.09	-0.20	-0.35	34.28	6.15	1.00	-1.32	0.93
8L	-6.95	1.04	0.29	0.43	-0.32	-2.74	33.28	4.81	4.41	-1.77	0.93
8M	-3.63	1.03	0.27	0.41	-0.40	-1.42	32.55	4.38	4.14	-2.24	0.93
8H	1.32	0.98	0.13	0.05	-0.37	0.54	32.61	2.28	0.56	-2.14	0.92
9L	-1.59	1.01	0.04	0.29	-0.18	-0.69	35.64	0.70	3.20	-1.14	0.94
9M	0.829	1.04	0.05	0.33	-0.22	0.38	38.65	0.98	3.90	-1.52	0.95
9H	3.48	1.00	-0.13	0.09	-0.29	1.21	28.43	-1.91	0.82	-1.51	0.90
10L	-1.74	1.04	-0.36	0.15	0.12	-0.78	37.54	-6.64	1.77	0.83	0.94
10M	1.18	1.01	-0.37	0.15	-0.11	0.61	42.44	-8.03	2.00	-0.87	0.95
10H	7.28	0.94	-0.46	-0.43	-0.64	3.46	36.30	-9.27	-5.22	-4.52	0.93

注：该表报告的样本期为 2000 年 01 月至 2008 年 06 月，给出了在组合水平下各组合的 *PINF* 和 FF 三因子的因子负载估计结果。其中每一 *SIZE* 组别的 *PINF*₋₁ 变量定义为其他 9 个 *SIZE* 组别的价值加权的套利收益率 (hedge return, 即 High *PIN*-Low *PIN*) 的等权平均数。

的增加而逐渐增加的特征，相反在大多数 *SIZE* 组别下，二者表现出了一定的负相关性，这与 Mohanram and Rajgopal (2009) 的研究不一致。与此同时，该现象与 Easley, Hvidkjaer et al. (2002) 的研究相违背，他们认为知情交易概率越大的股票会提供更高的回报，以补偿因为逆向选择而增加的风险，然而，本文从数据研究中并没有得到上述结论。

其次，本文发现，无论公司规模的大小，*PIN* 的因子负载的估计系数总是为负（除在第 10 个 *SIZE* 组别下的 Low *PIN* 组合下符号为正），它说明了无论是大公司还是小公司对信息风险因子有着负的敏感性，它们均倾向于规避知情交易，即知情交易风险不但没有带来更多的溢价，反而会表现为被更多的投资者追逐，从而提高了当期价格并同时降低了股票收益率。这一结果再次与

表 8 *PINF* 的因子负荷估计:四因子

<i>Port</i>	<i>Inter</i>	<i>MR</i>	<i>SMB</i>	<i>HML</i>	<i>MOM</i>	<i>PINF</i> ₋₁	<i>t_{Inter}</i>	<i>t_{MR}</i>	<i>t_{SMB}</i>	<i>t_{HML}</i>	<i>t_{MOM}</i>	<i>t_{PINF-1}</i>	<i>AdjR</i> ²
1L	-2.35	0.94	0.98	-0.04	-0.02	-0.23	-0.85	28.29	11.15	-0.38	-0.24	-1.27	0.93
1M	-2.56	1.00	1.07	-0.11	-0.01	-0.35	-1.01	32.80	13.33	-1.11	-0.22	-2.12	0.95
1H	-3.32	0.90	0.92	0.14	-0.05	-0.52	-1.19	26.65	10.40	1.32	-0.67	-2.88	0.92
2L	-5.44	1.00	0.96	0.07	0.03	-0.64	-1.88	28.32	10.45	0.61	0.45	-3.39	0.93
2M	-5.75	0.97	1.03	0.26	0.05	-0.69	-1.86	25.74	10.54	2.18	0.71	-3.44	0.92
2H	-1.39	0.97	0.94	0.01	0.03	-0.32	-0.58	33.62	12.49	0.10	0.56	-2.09	0.95
3L	-6.63	0.99	0.96	0.22	0.05	-0.30	-2.81	34.49	12.70	2.41	0.90	-1.92	0.95
3M	-3.86	1.00	0.81	0.34	0.02	-0.38	-1.56	33.52	10.20	3.66	0.39	-2.38	0.95
3H	-5.88	0.94	0.92	0.28	0.23	-0.54	-1.95	25.82	9.52	2.44	3.06	-2.74	0.91
4L	-4.87	0.95	0.79	0.37	0.05	-0.39	-1.96	31.55	10.00	3.90	0.80	-2.50	0.94
4M	-2.68	1.03	0.64	0.26	-0.03	-0.44	-1.29	40.86	9.56	3.33	-0.65	-3.33	0.96
4H	-5.53	0.99	0.73	0.39	0.12	-0.53	-1.99	29.24	8.15	3.71	1.73	-3.00	0.93
5L	-2.17	1.03	0.57	0.05	-0.06	-0.63	-0.91	35.83	7.50	0.61	-1.07	-4.15	0.95
5M	0.29	1.02	0.46	0.09	-0.02	-0.57	0.13	38.80	6.73	1.14	-0.45	-4.06	0.95
5H	-2.49	0.96	0.70	0.07	0.12	-0.31	-0.96	30.76	8.56	0.75	1.92	-1.87	0.93
6L	-4.37	1.00	0.55	0.43	0.02	-0.27	-1.85	35.00	7.43	4.76	0.30	-1.77	0.95
6M	-5.10	1.02	0.54	0.41	0.12	-0.51	-2.05	34.05	6.87	4.30	1.96	-3.14	0.94
6H	-4.66	1.00	0.64	0.18	0.22	-0.45	-1.93	34.29	8.41	1.98	3.75	-2.89	0.94
7L	-4.46	1.03	0.39	0.35	-0.05	-0.55	-1.62	31.15	4.45	3.37	-0.78	-3.04	0.93
7M	-2.20	1.04	0.31	0.24	0.07	-0.20	-1.06	41.38	4.76	3.03	1.30	-1.44	0.95
7H	-1.23	0.93	0.37	0.10	0.05	-0.18	-0.55	33.99	5.16	1.13	0.96	-1.24	0.93
8L	-7.05	1.04	0.30	0.43	0.01	-0.31	-2.70	32.91	3.60	4.36	0.18	-1.73	0.93
8M	-3.95	1.02	0.30	0.42	0.04	-0.39	-1.50	32.20	3.57	4.17	0.57	-2.14	0.93
8H	-0.25	0.97	0.29	0.09	0.18	-0.30	-0.10	33.46	3.82	1.02	3.05	-1.80	0.93
9L	-1.81	1.01	0.06	0.29	0.02	-0.17	-0.76	35.26	0.80	3.22	0.43	-1.08	0.94
9M	1.21	1.05	0.01	0.32	-0.04	-0.24	0.53	38.45	0.19	3.74	-0.77	-1.59	0.95
9H	0.71	0.99	0.15	0.16	0.31	-0.20	0.27	30.89	1.75	1.59	4.83	-1.16	0.92
10L	-1.40	1.04	-0.39	0.15	-0.04	0.12	-0.61	37.34	-5.36	1.65	-0.71	0.80	0.94
10M	0.98	1.01	-0.35	0.16	0.02	-0.11	0.49	42.03	-5.57	2.04	0.49	-0.86	0.95
10H	6.52	0.94	-0.39	-0.41	0.09	-0.63	3.06	36.30	-5.77	-4.98	1.67	-4.51	0.93

注:该表报告的样本期为 2000 年 01 月至 2008 年 06 月,给出了在组合水平下各组合的 *PINF* 和 FF 四因子的因子负载估计结果。其中每一 *SIZE* 组别的 *PINF*₋₁ 变量定义为其其他 9 个 *SIZE* 组别的价值加权的套利收益率(hedge return,即 High *PIN*-Low *PIN*)的等权平均数。

Mohanram and Rajgopal(2009)的研究不相一致,即规模大的公司对信息风险因子应具备正的敏感性,以补偿大公司较大的资本成本。整体而言,*PINF* 的因子负载表现为,随着组合的增加,其因子负载呈现出 U 型特征(如果剔除组合 10L),这说明,规模最小的公司和最大的公司,似乎更大程度上体现出了更大的信息风险敏感性。为了考察 *MOM* 变量是否影响了 *PIN* 因子负载的估计结果,本文在全有效样本期内,采用 FF 四因子模型进行相类似的时间序列回归分析,得到的结果如表 8 所示。该表所反映的情况如表 7 相类似,在此不再赘述。此外,从表 7 和表 8 的结果也可以看到,中国股市的 *HML* 和 *SMB* 的因子负载也同样没有呈现出明显的规律。这意味着中国股市有些未观测的因素扭曲了因子负载的表现。

2. 第二阶段回归

接下来,本文将组合月度收益与上一阶段估计得到的因子负载进行 Fama–MacBeth 横截面检验,以检验 *PIN* 因子负载能否对收益起到预测作用。如果 *PINF* 真的是风险因子并被定价,本文预期第二阶段回归应该得到正的估计系数。

为了消除个股的因子负载对回归结果造成的噪声影响,本文在做以下检验时均以组合水平代替公司水平。本文将构造四种不同的组合集分别进行检验,即:(1)10 个 *SIZE* 组合;(2)先划分为 10 个 *SIZE* 组别,然后,在每一组别内按照 *PIN* 值排序等分为 3 个组合;(3)先划分为 10 个 *SIZE* 组别,然后,在每一组别内按照 *PIN* 值排序等分为 5 个组合;(4)先划分为 10 个 *SIZE* 组别,然后,在每一组别内按照 *PIN* 值排序等分为 3 个组合,再基于公司水平的 *PINF* 的因子负载将每一组合细化为 3 个组合。

对于不同的组合,本文在样本期间(2000 年 01 月至 2008 年 06 月)中的每一月度分别进行横截面回归,并在组合水平下采用多因子模型进行如下时间序列回归分析:

$$R_i=\alpha_i+\beta_iLMKT_i+s_iLSMB_i+h_iLHML_i+m_iLMOM_i+p_iLPIN_i+\varepsilon_i$$

这里, *LMKT_i*, *LSMB_i*, *LHML_i*, *LMOM_i* 和 *LPIN_i* 分别为组合 *i* 在第一阶段多因子模型回归中所得到的市场超额收益率(MR–Rf)因子、SMB 因子、HML 因子、MOM 因子以及 *PINF* 因子的因子负载。然后按照 Fama–MacBeth 方法进行相关参数的均值及其 *t* 统计值的估计。考虑到可能存在异方差影响,本文在进行横截面回归时采用经 White 调整的 *t* 统计量。

第二步的估计结果如表 9 所示。其中, *Panel A* 给出了在 *PINF* 因子和 FF 三因子分析下的估计结果, *Panel B* 给出了在 *PINF* 因子和 FF 四因子分析下的估计结果。由 *Panel A* 可知, *PINF* 的因子负载的估计系数均为非负,并且在前三种组合集下该结果在常规显著性水平下均不显著,而基于 *SIZE*、*PIN* 以及 *PINF* 因子负载进行划分的组合所得到的 *PINF* 因子负载的估计系数仅在 10% 的显著性水平下显著异于 0。通过观察 *Panel B*,可以得到有关 *PINF* 因子负载相类似的估计结果,表明惯性变量并没有对 *PIN* 因子的表现施加过多的影响。

表 9 第二阶段的横截面检验结果

Panel A: 基于三因子组合的估计结果							
Portfolios	<i>Intercept</i>	<i>LMKT</i>	<i>LSMB</i>	<i>LHML</i>	<i>LPIN</i>	<i>Adj-R²</i>	
10 <i>SIZE</i>	-0.029	0.035	0.001	-0.005	0.006	0.708	
	-0.537	0.629	0.267	-1.011	0.822		
10 <i>SIZE</i> *3 <i>PIN</i>	0.010	-0.006	-0.001	-0.004	0.000	0.466	
	0.905	-0.408	-0.252	-0.847	-0.093		
10 <i>SIZE</i> *5 <i>PIN</i>	0.010	-0.006	-0.001	-0.003	0.001	0.373	
	1.094	-0.444	-0.270	-0.775	0.272		
10 <i>SIZE</i> *3 <i>PIN</i> *3 <i>PIN</i> loading	-0.002	0.000	0.001	0.001	0.003	0.320	
	-0.254	1.070	0.196	0.297	1.690		
Panel B: 基于四因子组合的估计结果							
Portfolios	<i>Intercept</i>	<i>LMKT</i>	<i>LSMB</i>	<i>LHML</i>	<i>LMOM</i>	<i>LPIN</i>	<i>Adj-R²</i>
10 <i>SIZE</i>	-0.027	0.034	0.001	-0.006	0.001	0.006	0.724
	-0.501	0.591	0.257	-0.985	0.073	0.810	
10 <i>SIZE</i> *3 <i>PIN</i>	0.008	-0.004	-0.001	-0.004	0.004	0.000	0.492
	0.792	-0.301	-0.212	-0.888	0.484	0.004	
10 <i>SIZE</i> *5 <i>PIN</i>	0.009	-0.004	-0.001	-0.003	0.003	0.001	0.394
	1.024	-0.347	-0.245	-0.825	0.403	0.332	
10 <i>SIZE</i> *3 <i>PIN</i> *3 <i>PIN</i> loading	-0.002	0.000	0.001	0.001	0.000	0.003	0.336
	-0.305	1.049	0.182	0.278	-0.031	1.685	

注:该表报告了在四种不同的组合集下的第二阶段横截面检验结果,其中 *Panel A* 给出了在 *PINF* 因子和 FF 三因子分析下的估计结果, *Panel B* 给出了在 *PINF* 因子和 FF 四因子分析下的估计结果,其间采用了经 White 调整的 *t* 统计量。四种组合集的具体分组规则参见正文。估计结果的第一行是回归系数,第二行为异方差稳健的 *t* 检验值。

五、结 论

本文考察了私有信息交易(采用知情交易概率测度作为代理变量)在中国股市是否可被有效地作为一个风险定价因子。虽然,有不少研究使用了这一变量,但是,目前并没有研究对这一问题进行明确的回答。本文采用时间序列分析、Fama-MacBeth 两阶段横截面分析,并结合多因子模型,发现在市场、公司规模、账面市值比和惯性因子之外,知情交易无法预测资产收益。这意味着,在中国股市采用知情交易概率作为定价因子或信息交易含量测度值得进一步商榷。

就其本质而言, PIN 只是根据订单的买卖进行估计,从而也大致体现了市场上的股票买卖情况,这可以类似地看作订单不平衡(order imbalance)变量。事实上,这样的测度很可能有着先天的不足。一方面,这个测度本身无法剔除掉流动性交易者的情况(尽管在 PIN 的计算中,采用了流动性交易者或者知情交易者在一段时间稳定的假设,但是这样的前提在现实中很难完全满足,尤其是在投机性较强、热点概念转变较快的中国股市);另一方面,人们对“知情投资者”的购买行为本身究竟是否真的是因为知情交易,也存在疑问。一些基于行为金融的研究(如 Barber and Odean (2000)Barber and Odean,2000)指出,当投资者交易次数越多时(或者更为过度自信时),其收益反而会下降,这会随之带来如下质疑,即如果个人交易越多(买卖行为越多),亏损越严重的话,那么他们有什么理由被称为知情交易者呢?如果这个质疑是合理的,那么 PIN 显然将难以胜任知情交易的代理变量。

参考文献:

- 韩立岩、郑君彦、李东辉(2008):《沪市知情交易概率特征与风险定价能力》,《中国管理科学》,第1期。
- 王茂斌、彭晴、孔东民(2010):《参与知情交易提高了基金业绩吗》,《南方经济》,第6期。
- 孔东民、申睿(2008):《R2、异常收益与交易的知情概率成分》,《中大管理研究》,第3期。
- 刘元海、陈伟忠(2004):《中国股市知情交易概率的测定》,《同济大学学报》,第1期。
- 攀登、施东晖(2006):《知情交易概率的测度模型及其影响因素分析》,《管理世界》,第6期。
- 攀登、施东晖和曹敏(2003):《中国个人投资者采用股价趋势交易策略的经验研究》,《世界经济》,第11期。
- 王春峰、董向征、房振明(2005):《信息交易概率与中国股市价格行为关系的研究》,《系统工程》,第2期。
- 王展翔(2007):《关于我国股市知情交易水平及其性质的实证研究》,《当代财经》,第2期。
- 杨之曙、姚松瑶(2004):《沪市买卖价差和信息性交易实证研究》,《金融研究》,第4期。
- 张宗新(2008):《内幕交易行为测定:理论模型与实证分析》,《管理世界》,第4期。
- Barber, B. and T. Odean (2000): "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors", *Journal of Finance*, 55, 773-806.
- Benos, E. and M. Jochev (2007): "Testing the PIN Variable", Working Paper, University of Illinois at Urbana Champaign.
- Botosan, C., M. Plumlee and Y. Xie (2004): "The Role of Information Precision in Determining the Cost of Equity Capital", *Review of Accounting Studies*, 9, 233-259.
- Brown, S. and S. Hillegeist (2007): "How Disclosure Quality Affects the Level of Information Asymmetry", *Review of Accounting Studies*, 12, 443-477.
- Brown, S., S. Hillegeist and K. Lo (2004): "Conference Calls and Information Asymmetry", *Journal of Accounting and Economics*, 37, 343-366.
- Campbell, J., T. Ramadorai and A. Schwartz (2009): "Caught on Tape: Institutional Trading, Stock Returns, and Earnings Announcements", *Journal of Financial Economics*, 92, 66-91.
- Carhart, M. (1997): "On Persistence in Mutual Fund Performance", *Journal of Finance*, 52, 57-82.
- Duarte, J., X. Han, J. Harford and L. Young (2008): "Information Asymmetry, Information Dissemination and the Effect of Regulation FD on the Cost of Capital", *Journal of Financial Economics*, 87, 24-44.

- Duarte, J. and L. Young (2009): "Why is PIN Priced?", *Journal of Financial Economics*, 91, 119–138.
- Easley, D., Hvidkjaer and M. O'Hara (2002): "Is Information Risk a Determinant of Asset Returns", *Journal of Finance*, 57, 2185–2221.
- Easley, D., S. Hvidkjaer and M. O'Hara (2004): "Factoring Information into Returns", Working Paper, Cornell University.
- Easley, D., N. Kiefer, M. O'Hara and J. Paperman. (1996): "Liquidity, Information and Infrequently Traded Stocks", *Journal of Finance*, 51, 1405–1436.
- Easley, D., N. Kiefer, M. O'Hara and Wu (2008): "Varying Arrival Rates of Informed and Uninformed Trades", *Journal of Financial Econometric*, 4, 1–37.
- Easley, D. and M. O'Hara (2004): "Information and the Cost of Capital", *Journal of Finance*, 59, 1553–1583.
- Ecker, F., J. Francis, I. Kim, P. Olsson and K. Schipper (2006): "A Returns-based Representation of Earnings Quality", *Accounting Review*, 81, 749–780.
- Ellul, A. and M. Pagano (2006): "IPO Underpricing and after-Market Liquidity", *Review of Financial Studies*, 19, 381–421.
- 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. and K. French (1993): "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33, 3–56.
- Fama, E. and J. MacBeth (1973): "Risk, Return and Equilibrium: Empirical Tests", *Journal of Political Economy*, 81, 607–636.
- Francis, J., R. LaFond, P. Olsson and K. Schipper (2005): "The Market Pricing of Accruals Quality", *Journal of Accounting and Economics*, 39, 295–327.
- Francis, J., R. LaFond, P. Olsson and K. Schipper (2004): "Costs of Equity and Earnings Attributes", *Accounting Review*, 79, 967–1010.
- Fuller, K., B. Van Ness and R. Van Ness (2010): "Is Information Risk Priced for NASDAQ-listed Stocks?", *Review of Quantitative Finance and Accounting*, 34, 301–312.
- Henry, T. (2005): "Constrained Short Selling and the Probability of Informed Trade", *Working Paper*, University of Washington.
- Hilary, G. (2006): "Organized Labor and Information Asymmetry in the Financial Markets. *Review of Accounting Studies*, 11, 525–548.
- Huddart, S. and B. Ke (2007): "Information Asymmetry and Cross-sectional Variation in Insider Trading", *Contemporary Accounting Research*, 24, 195–232.
- Hughes, J., J. Liu and J. Liu (2005): "Information, Diversification, and Cost of Capital", Working Paper, University of California at Los Angeles.
- Jayaraman, S. (2008): "Earnings Volatility, Cash Flow Volatility, and Informed Trading", *Journal of Accounting Research*, 46, 809–851.
- Lambert, R., C. Leuz and R. Verrecchia (2007): "Accounting Information, Disclosure, and the Cost of Capital", *Journal of Accounting Research*, 45, 385–420.
- Lee, C. and B. Radhakrishna (2000): "Inferring Investor Behavior: Evidence from TORQ Data", *Journal of Financial Markets*, 3, 83–111.
- Lee, C. and M. J. Ready (1991): "Inferring Trade Direction from Intraday Data", *Journal of Finance*, 46, 733–746.
- Mohanram, P. and S. Rajgopal (2009): "Is PIN Priced Risk?", *Journal of Accounting and Economics*, 47, 226–243.
- Odders-White, E. and M. Ready (2006): "Credit Ratings and Stock Liquidity", *Review of Financial Studies*, 19, 119–157.
- Pan, J. and A. Poteshman (2006): "The Information in Option Volume for Future Stock Prices", *Review of Financial Studies*, 19, 871–908.
- Vega, C. (2004): "Are Investors Overconfident?", Working Paper, University of Rochester.
- Zhao, X. and K. Chung (2006): "Decimal Pricing and Information-based Trading: Tick Size and Informational Efficiency of Asset Price", *Journal of Business Finance & Accounting*, 33, 753–766.

(责任编辑:周莉萍)