



# 基于改进信息交易概率模型的信息风险测度研究

杨宝臣<sup>1</sup>, 郭 灿<sup>1</sup>, 常建勇<sup>1,2</sup>

1 天津大学 管理与经济学部, 天津 300072

2 首都师范大学, 北京 100048

**摘要:**准确测度信息不对称是研究金融市场信息不对称问题的基础。在详细考察经典的信息交易概率模型的建模过程和假设条件的基础上,针对信息交易者和非信息交易者在不同信息状态下的订单提交状况,提出考虑外生信息状态概率、异质期望和交易活跃度的改进的信息交易概率模型;通过带机制转换的向量自回归模型估计外生信息状态概率,综合考虑多方面的市场交易信息,提高模型估计的准确性。实证结果表明,改进的信息交易概率模型可以较好地刻画证券市场的信息不对称程度。从截面上看,信息交易概率与标的资产的垄断优势和话题性相关,垄断优势越强的股票信息交易概率越低,话题性越强的股票信息交易概率越高;从时间上看,信息交易概率的变化受公告信息的影响,公告信息发布前的信息交易概率较高,公告信息发布后的信息交易概率较低。因此,在防范信息风险时,应重点关注话题性较强的版块,特别是股价比较容易受舆论影响波动的个股;同时,在公告信息发布之前的时段,应对市场波动和信息动向给予更多关注。

**关键词:**信息风险;PIN模型;交易活跃度;异质期望;信息状态

**中图分类号:**F830.9

**文献标识码:**A

**doi:**10.3969/j.issn.1672-0334.2014.06.011

**文章编号:**1672-0334(2014)06-0121-11

## 1 引言

证券市场中的信息不对称水平是投资者和市场监管者都十分关心的问题。金融市场微观结构理论认为,信息交易者和非信息交易者是证券市场中普遍存在的交易主体<sup>[1]</sup>,信息交易者有准确认知信息的能力,而非信息交易者则没有这样的优势。信息不对称问题的存在,不仅给非信息交易者带来额外的或有损失,损害市场的流动性供应能力,从而影响市场的有效性和资源配置能力,制约证券市场对经济发展的促进作用;而且有可能引发证券市场的剧烈波动,甚至导致经济泡沫化,造成不可估量的损失。但是,由于制度设计的不健全和投资者的不成熟,中国证券市场的信息不对称程度较高,散户的价值发现过程有限,加上机构投资者的操控行为和媒体对散户的引导作用,使中国证券市场中的价值泡

沫和投机交易大量存在。

传统资产定价理论假设所有投资者都可以及时、免费地获得充分的市场信息,在这样的理想市场中,证券的市场价格就是其内在价值的完全反映,没人能利用信息获取超额收益。但在实际市场中,信息的搜集成本是不可忽略的,投资者把握信息的能力也是有限的。Grossman等<sup>[2]</sup>认为,只要信息成本存在,完全信息均衡市场就不会出现,由信息不对称带来的偏差就不可忽略。但是,迄今为止,学术界对于信息不对称水平的准确测度方法尚无定论,最为广泛应用的信息交易概率模型及其PIN指标<sup>[3-5]</sup>更是受到国内外不少学者的质疑。因此,本研究在详细考察信息交易概率模型的建模过程和假设条件的基础上,对该模型在模型设计和估计过程中存在的问题进行梳理,并提出相应的改进。

**收稿日期:**2014-07-17 **修返日期:**2014-10-07

**基金项目:**国家自然科学基金(71171144,71471129);高等学校博士学科点专项科研基金(20130032110016)

**作者简介:**杨宝臣(1966-),男,河北唐山人,毕业于天津大学,获博士学位,现为天津大学管理与经济学部教授、博士生导师,研究方向:经济系统建模与分析、金融市场定量分析、固定收益证券管理等。E-mail:bchyang@tju.edu.cn

## 2 相关研究评述

Bagehot<sup>[6]</sup>、Glosten等<sup>[1]</sup>和 Kyle<sup>[7]</sup>最早关注投资者交易行为的信息现象,并提出序贯交易模型和批量交易模型等信息模型。由于知情交易很难从市场中直接观测,早期信息风险的实证研究大多从信息不对称的影响因素入手,通过买卖价差、成交量和换手率以及上市公司内部的持股比例等指标间接测度信息不对称的水平和程度<sup>[6,8]</sup>。但是,仅依靠间接指标无法准确测度信息交易的比例及其影响,不能直接反映投资者在市场中面临的信息不对称风险的大小。因此,许多学者更深入地分析信息交易的过程及其影响,并提出逆向选择成本<sup>[9-11]</sup>和信息交易概率<sup>[3-4,12]</sup>等直接度量信息不对称的指标。

Easley等<sup>[3-4,12]</sup>提出的信息交易概率模型(EKOP模型)和PIN指标受到最为广泛的关注。因为简单易用,许多研究都不加验证地直接使用PIN指标作为信息不对称程度的替代变量<sup>[13-15]</sup>。但信息交易概率指标能否合理衡量信息不对称也受到国内外众多学者的质疑<sup>[16-22]</sup>,除对于规模较大、交易频繁的样本无法估计以外,模型自身设计的不合理性<sup>[18]</sup>和模型对不同样本区间数据估计结果的不一致<sup>[16]</sup>是众多学者质疑的焦点。

信息交易概率模型刻画的是做市商从交易者提交的订单中获得信息的贝叶斯学习过程<sup>[4]</sup>,模型最初关注的是投资者行为对市场信息的反应<sup>[23]</sup>,而不是市场的信息不对称程度。信息交易概率模型中有两类待估参数,即信息状态概率和各类型订单的到达率。模型构建的依据是信息交易者和非信息交易者在不同信息状态下会有不同的决策表现,但仅通过不同类型订单的提交量很难同时准确估计出信息状态概率和信息交易者比例这两类参数。值得注意的是,从估计结果看,该模型对市场信息状态的估计往往显示出较大的误差<sup>[3]</sup>。Easley等<sup>[5,24-25]</sup>将基于买卖订单到达强度不平衡的信息交易概率(PIN)模型扩展为基于买卖交易量不平衡的信息交易概率(VPIN)模型时,甚至放弃了对不可观测的市场信息状态变量的估计。这也说明该模型在反映市场实际信息状态方面的局限。

事实上,市场盘面信息中的所有交易指标都是投资者行为的反映,都在一定程度上反映了市场的实际信息状态<sup>[23,26-27]</sup>。正常情况下,这些交易指标会按照一定的规则和模式在一定范围内相互影响变化。没有新信息进入市场,各指标间的相互影响关系就不会改变,而当市场信息状态发生改变时,新信息会随着投资者的交易行为汇入市场<sup>[28-31]</sup>,在此过程中,各交易指标间的相互影响关系有可能发生相应的改变。因此,本研究在对信息交易概率模型进行估计之前,先通过带机制转换的向量自回归模型(MSVAR模型)估计出信息状态概率,再将其代入信息交易概率模型估计其他参数。而且,这种外生确定信息状态概率的方法也可以减少信息交易概率模型本身的待估参数,提高模型估计的效率和准确性。

另外,经典的信息交易概率指标PIN不但与信息交易者和非信息交易者的订单到达率有关,还与信息发生的概率有关,这就决定了不同信息条件下的信息交易概率并不具有可比性。其实,现有的信息不对称指标大都没有考虑实际信息状态对信息不对称的影响。需要强调的是,正如Easley等<sup>[5]</sup>所认为的,在证券市场中,市场所关心的信息不仅是与资产基础价值相关的公司和产品资讯,还包括市场整体的交易状况和流动性需求。这种广义的信息定义意味着信息事件的发生有可能是频率较高的事件。相对而言,市场中的信息不对称水平(信息交易者的比例)在一定时段内应该是相对稳定的。特别的,当以高频数据为研究对象时,没有剔除信息状态的信息不对称指标很大程度上反映的是实际信息状态的变化,而不是单纯的信息不对称水平。这也是在VPIN模型<sup>[5]</sup>中VPIN指标可以在一定程度上预测市场暴跌的原因,但这样的模型并不能保证准确测度信息不对称风险的有效性。其实,最早的信息交易概率模型(EKOP模型)的信息交易概率<sup>[23]</sup>原本与市场实际信息状态无关,由于未提交的潜在订单量很难直接观测,EKOP模型在实证中的应用不多。Easley等<sup>[4]</sup>曾尝试以5分钟无交易的时间间隔数作为未发生的潜在交易量的替代指标,但当时间间隔取不同的长度时,所得估计结果差异很大。因此,本研究考虑通过添加交易活跃度指标来解决对潜在订单量的估计问题。

因此,本研究在详细考察经典的信息交易概率模型的建模过程和假设条件的基础上,针对信息交易者和非信息交易者在不同信息状态下的订单提交状况,提出考虑外生信息状态概率、异质期望和交易活跃度的改进的信息交易概率模型,外生的信息状态概率通过带机制转换的向量自回归模型(MSVAR模型)估计得到,这样既综合考虑了多方面的市场交易信息,又减少了传统信息交易概率模型的待估参数,可以在一定程度上提高模型估计的准确性。

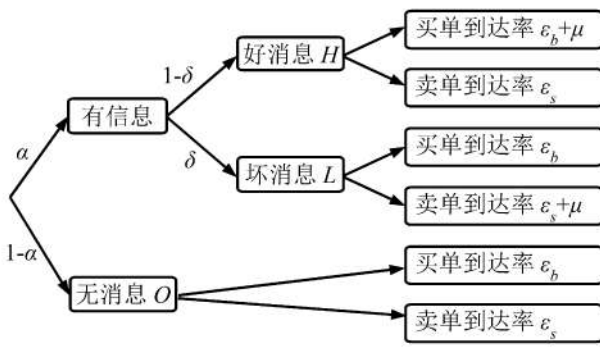
## 3 改进的信息交易概率模型的构建

### 3.1 经典的信息交易概率模型:EKOP模型

EKOP模型<sup>[3]</sup>假设市场中存在唯一的风险资产、单一的风险中性的做市商以及多个信息交易者和多个非信息交易者,做市商连续给出资产的买卖报价,并承诺在任意时间以所给报价与交易者达成交易。假设在 $T$ 个交易日的样本期内,每一个交易日中信息交易者和非信息交易者都与做市商进行以风险资产为标的的交易,每个交易日内时间连续,记为 $t \in [0, T]$ 。假设完全竞争的做市商期望利润为零,那么在任意时刻,做市商提出的报价应该等于其基于自身所知信息推断出的该风险资产的期望价值。图1给出经典的信息交易概率模型中做市商视角下的某一个股票在一个交易日内的交易过程。

在图1中,每个交易日开盘前(第1个节点时刻),自然(nature,指外生的决定随机变量概率分布的特殊机制,也可看作博弈分析中虚拟的参与者)就





注:  $\alpha$ 为某交易日信息事件发生概率,  $\delta$ 为坏消息发生概率,  $(1-\delta)$ 为好消息发生概率,  $\mu$ 为信息交易者买、卖订单到达率。

图1 EKOP模型中做市商视角下的资产日交易过程  
Figure 1 Daily Trading Process of Assets from the Perspective of Market Maker in EKOP Model

已经决定了该交易日是否有信息事件发生。假设所有信息事件的发生独立,且每天信息事件发生的概率都为 $\alpha$ 。在有信息事件发生的情况下,自然会进一步确定信息状态为好消息还是坏消息,坏消息发生的概率为 $\delta$ ,好消息发生的概率为 $(1-\delta)$ 。同时,假设所有新信息都会在当日完全汇入资产价格,而不会影响下一个交易日的资产价值。

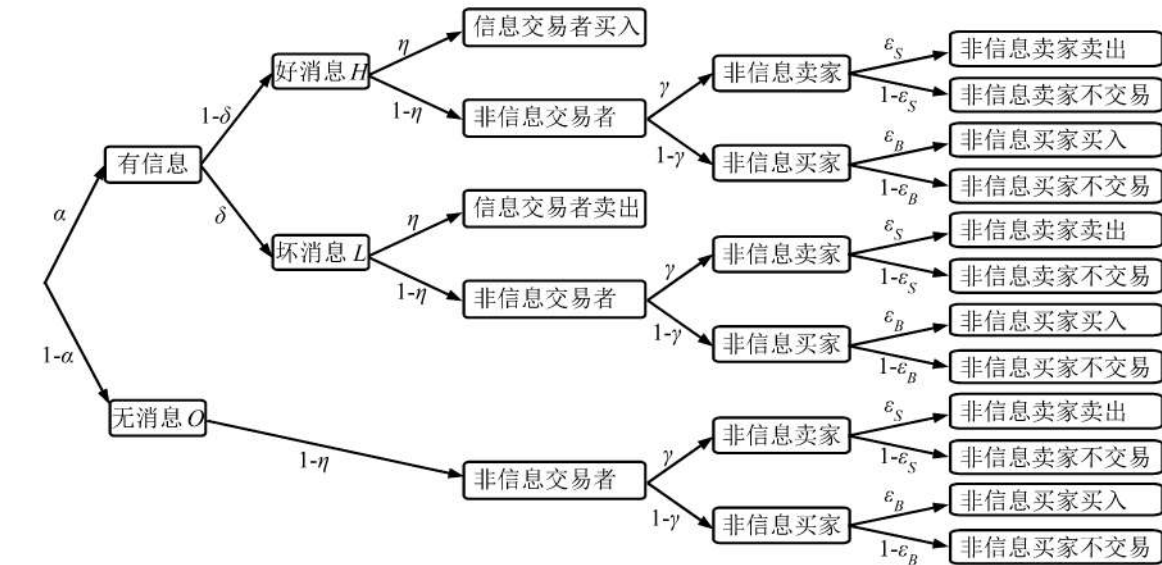
信息交易者和非信息交易者在不同信息状态下的交易决策不同,信息交易者是愿意并有能力依据资产真实信息做出买、卖决策的投资者。相对于非信息交易者,信息交易者拥有较大的信息优势,他们只在有信息的情况下交易,且交易方向与资产价格

的实际变动方向一致。非信息交易者或者不关心、或者没有能力正确把握资产价值的变动方向,所以他们的交易与资产实际价值变动关系不大。假设信息交易者和非信息交易者的订单到达率都服从泊松分布,且各交易日之间的买、卖订单到达率彼此独立。信息交易者的订单到达率服从期望值为 $\mu$ 的泊松分布,非信息交易者买入和卖出订单的到达率分别服从期望值为 $\epsilon_b$ 和 $\epsilon_s$ 的泊松分布。那么,好消息发生时,信息交易者会以 $\mu$ 的到达率提出买单,非信息交易者则分别以 $\epsilon_b$ 和 $\epsilon_s$ 的到达率提出买单和卖单,此时总买单到达率为 $\epsilon_b + \mu$ ,总卖单到达率为 $\epsilon_s$ ;同理,坏消息发生时,信息交易者会以 $\mu$ 的到达率提出卖单,非信息交易者的交易状态不变,此时的总买单到达率为 $\epsilon_b$ ,总卖单到达率为 $\epsilon_s + \mu$ ;没有信息发生时,只有非信息交易者交易,买单和卖单的到达率分别为 $\epsilon_b$ 和 $\epsilon_s$ 。

根据以上分析可知,信息交易占总交易的比例为  $PIN, PIN = \frac{\alpha\mu}{\alpha\mu + \epsilon_b + \epsilon_s}$ , 通过极大似然估计可以得到待估参数向量  $\theta, \theta = (\alpha, \delta, \mu, \epsilon_b, \epsilon_s)$ , 进而得到信息交易概率指标。从其计算公式可以看出,它与知情交易的买、卖订单的到达率以及非知情交易的买、卖订单的到达率有关,也与信息事件发生的概率有关,而与所发生的信息事件是好消息还是坏消息没有直接的关系。

3.2 最早的信息交易概率模型:EO模型

实际上,最早的信息交易概率模型并不是EKOP模型,而是Easley等<sup>[12, 23]</sup>提出的EO模型,EO模型的基本假设与EKOP模型一致。图2给出EO模型的资产日交易过程。



注:  $\eta$ 为市场中信息交易者买、卖订单所占比例,  $(1-\eta)$ 为市场中非信息交易者买、卖订单所占比例,  $\gamma$ 为非信息交易者倾向于卖出的概率,  $(1-\gamma)$ 为非信息交易者倾向于买入的概率,  $\epsilon_b$ 为所有倾向于买入的非信息交易者中实际发生交易的比例,  $\epsilon_s$ 为所有倾向于卖出的非信息交易者中实际发生交易的比例。

图2 EO模型中做市商视角下的资产日交易过程  
Figure 2 Daily Trading Process of Assets from the Perspective of Market Maker in EO Model

然而,如图2所示,EO模型并没有添加订单到达率的分布条件假设,而是直接假设市场中由信息交易者提交订单的比例为 $\eta$ ,好消息状态下以 $\eta$ 比例买入订单,坏消息状态下以 $\eta$ 比例卖出订单;非信息交易者的交易与实际信息状态无关,假设任意时刻都会有 $\gamma$ 比例的非信息交易者有卖出股票倾向,但其中只有 $\varepsilon_s$ 比例实际提交了卖出订单,而剩下的 $(1 - \varepsilon_s)$ 比例考虑卖出股票的非信息交易者却因各种原因没有卖出;同样,更倾向于买入股票的 $(1 - \gamma)$ 比例的非信息交易者中也只有 $\varepsilon_B$ 比例提交了买入订单。通过构建图2所示的日交易模型,就可直接估计出信息交易者提交订单的比例 $\eta$ 。

EO模型的不交易订单所反映的状况在证券市场中大量存在,除了相当比例以长期持有为目的的不活跃证券外,大多数证券不交易的原因与证券持有者对市场信息状态的无信息判断有关。与实际信息状态的无信息状态一致,投资者的不交易决策意味着投资者认为市场尚未出现足以影响他们对资产价值判断的新信息。对于非信息交易者,他们在所有信息状态下都可能对市场看平,而无信息状态下决定不交易的信息交易者依然关注着市场,且有提交订单的可能,只是出于信息优势,他们会对市场走势看平,并做出不交易的决策。由于不交易量难以估计,后来的EKOP模型通过订单到达率泊松分布的假设回避了对不交易量的估计问题。但本研究认为,可以通过考虑异质期望和交易活跃度构造更符合实际的信息交易概率模型,同时也可以解决不交易量的估计问题。

3.3 改进的信息交易概率模型的构建

与经典的信息交易概率模型一样,假设每交易日的信息状态由自然在每交易日开盘前决定,信息状态依然是无信息(O)、好消息(H)和坏消息(L)3种状态之一,各信息状态概率取值也与经典模型一致,分别为 $P(O) = 1 - \alpha, P(H) = \alpha(1 - \delta), P(L) = \alpha\delta$ 。由于本研究的信息状态概率是外生给定的,在模型

估计时不需要对参数 $\alpha$ 和 $\delta$ 进行估计。

假设某具体时段内,各类型交易者的比例及其分别可操作的订单量相对稳定。记某证券被投资者关注且有可能被交易的比例为该证券的活跃度 $\rho$ ,其中被信息交易者和非信息交易者的交易行为影响的比例,即信息交易者和非信息交易者买卖订单所占比例分别为 $\eta$ 和 $(1 - \eta)$ 。在此,本研究通过添加活跃度指标 $\rho$ ,剔除了相当一部分以长期持有为目的的不活跃证券,因为这些订单的不交易与市场实际的信息状态无关,不需要在模型中加以考虑。假设信息交易者提交订单的交易方向与实际信息状态一致,即好消息时买入、坏消息时卖出、无信息时不交易;非信息交易者在所有信息状态下,主动买入、主动卖出和不交易的订单比例分别为 $\theta_B, \theta_S$ 和 $(1 - \theta_B - \theta_S)$ 。那么,一个交易日内的交易过程可以用图3的交易结构图来表示。

从图3的交易结构图中可以看出,好消息H、坏消息L和无信息O共3种信息状态下的买、卖订单和不交易订单的比例分别满足

$$P(B, S, N | H) = [\eta + (1 - \eta)\theta_B]^B [(1 - \eta)(1 - \theta_B - \theta_S)]^N [(1 - \eta)\theta_S]^S$$

$$P(B, S, N | L) = [(1 - \eta)\theta_B]^B [(1 - \eta)(1 - \theta_B - \theta_S)]^N \cdot [\eta + (1 - \eta)\theta_S]^S$$

$$P(B, S, N | O) = [(1 - \eta)\theta_B]^B [\eta + (1 - \eta)(1 - \theta_B - \theta_S)]^N [(1 - \eta)\theta_S]^S$$

则,第 $t$ 个交易日的交易行为可以通过以下似然函数表示,即

$$L(\theta | B_t, S_t, N_t) = \alpha\delta[\eta + (1 - \eta)\theta_B]^{B_t} [(1 - \eta)(1 - \theta_B - \theta_S)]^{N_t} [(1 - \eta)\theta_S]^{S_t} + \alpha(1 - \delta) \cdot [(1 - \eta)\theta_B]^{B_t} [(1 - \eta)(1 - \theta_B - \theta_S)]^{N_t} [\eta + (1 - \eta)\theta_S]^{S_t} + (1 - \alpha)[(1 - \eta)\theta_B]^{B_t} [\eta + (1 - \eta)(1 - \theta_B - \theta_S)]^{N_t} [(1 - \eta)\theta_S]^{S_t}$$

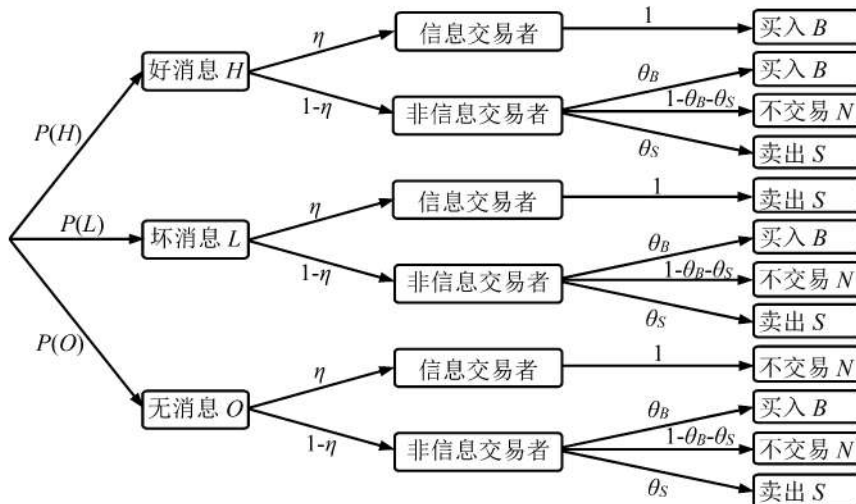


图3 改进信息交易概率模型中做市商视角下的资产日交易过程

Figure 3 Daily Trading Process of Assets from the Perspective of Market Maker in the Improved PIN Model

其中,  $B_t$  为第  $t$  个交易日活跃证券中总买入的数量,  $S_t$  为第  $t$  个交易日活跃证券中总卖出的数量,  $N_t$  为第  $t$  个交易日活跃证券中不交易的数量,  $\theta = (\alpha, \delta, \eta, \theta_B, \theta_S)$ 。假设证券交易在  $T$  个交易日的样本期内互相独立, 那么样本期内交易过程的似然函数可表示为

$$L(\theta | M) = \prod_{i=1}^T L(\theta | B_i, S_i, N_i)$$

其中,  $M = ((B_1, S_1, N_1), \dots, (B_T, S_T, N_T))$ ;  $i$  为样本期中的第  $i$  个交易日,  $i = 1, 2, \dots, T$ 。通过对以上似然函数的估计, 可以得到参数向量  $\theta$ 。与EO模型相同, 模型估计所得的参数  $\eta$  就是信息交易概率指标。

#### 4 数据来源和样本选取

本研究从沪、深两市2010年之前上市且一直处于正常交易状态的股票中随机选取30支股票为研究样本, 其中沪市16支, 深市14支, 对这些股票在2010年1月至2012年6月的日交易数据进行分析。数据来源于新华社金融数据平台, 除基本价量信息之外, 日交易数据中还包括内、外盘指标。内盘和外盘分别是交易量中以主动卖出价格和主动买入价格成交的数量, 这两个指标通常可以被看作多、空双方总体实力的表现。内、外盘数量差距较大, 说明市场多、空双方力量比较悬殊, 如当总体内盘数量远远高于总体外盘数量时, 股价往往会下行, 这是很多人积极抛出的结果。

在估计信息交易概率模型时, 本研究以月度为单位进行计算。以月度作为单位的主要原因在于, 股票市场本身是一个充满变化的市场, 过长时间间隔内的信息状态和交易活动都有可能较大的差异, 相对而言, 1个月的时间较短, 可以认为模型参数在这段时间内比较稳定。

为了使读者了解模型估计的具体过程, 本研究以上海交易所上市的交易代码为600000的浦发银行股票的日交易数据为例演示信息交易概率模型的构建和估计过程。然后, 用同样的方法估计得到所有样本的信息不对称指标, 并对其进行统计分析。

### 5 改进的信息交易概率模型的估计

#### 5.1 外生信息状态概率估计

通过对价量指标的分析 and 考察, 综合考虑可得性、相关性和可比性等因素, 本研究选取日均价和日收益率作为被解释变量, 选取日内波幅、换手率和订单不平衡程度作为解释变量, 构建MSVAR模型, 并对外生信息状态概率进行估计。各指标的计算方式如下。

(1) 日均价 ( $p_t$ )。取每日开盘价与收盘价的报价中点为当天的日均价指标, 即第  $t$  天的日均价可以表示为  $p_t = \frac{P_{open,t} + P_{close,t}}{2}$ ,  $P_{open,t}$  为第  $t$  天的开盘价,  $P_{close,t}$  为第  $t$  天的收盘价。

(2) 日收益率 ( $r_t$ )。取对数形式的收益率指标可通过相临两天的收盘价计算得到, 即  $r_t = \ln(1 + R_t) =$

$$\ln\left(\frac{P_{close,t}}{P_{close,t-1}}\right), \text{ 其中 } R_t \text{ 为第 } t \text{ 天的相对收益率, } R_t = \frac{P_{close,t} - P_{close,t-1}}{P_{close,t-1}}。$$

(3) 日内波幅 ( $v_t$ )。通过某交易日的最高价和最低价可以求得相对价格形式的日内波幅,  $v_t = \frac{2(P_{max,t} - P_{min,t})}{P_{max,t} + P_{min,t}}$ ,  $P_{max,t}$  为第  $t$  天的最高价,  $P_{min,t}$  为第  $t$  天的最低价。

(4) 换手率 ( $TR_t$ )。股票转手买卖的频率, 是交易量的替代指标,  $TR_t = \frac{VOL_t}{AMOUNT_t} \times 100\%$ ,  $TR_t$  为第  $t$  天的换手率,  $VOL_t$  为第  $t$  天的成交量,  $AMOUNT_t$  为第  $t$  天的流通股数。

(5) 订单不平衡程度 ( $StB_t$ )。第  $t$  天的订单不平衡程度  $StB_t = \frac{2(S_t - B_t)}{S_t + B_t}$ ,  $S_t$  为第  $t$  天的内盘交易量,  $B_t$  为第  $t$  天的外盘交易量。该比值越高, 说明市场中看空的人越多。

根据前文分析, 模型的机制个数应取3, 即模型应在无消息、好消息和坏消息3种信息状态间转换。通过比较各模型的似然函数值和信息准则值, 选定MSIAH(3)-VARX(1)的模型形式, 得到表1所示估计结果。对比表1中不同机制下指标系数间的关系可以看出, 在机制I的收益模型和价格模型中, 换手率、日内波幅和订单不平衡程度3个指标当期的系数都显著为负, 即这种信息状态下, 股票的成交量越大、波动越剧烈、买卖订单越不平衡, 股票的价格就越低, 收益就越少。显然, 这是股票市场处于坏消息状态下的表现。同理, 机制II中换手率、日内波幅和订单不平衡程度3个指标与同期股票收益率和股票价格正相关, 应是股票市场好消息状态的反映。与前两个机制不同, 在机制III中, 虽然换手率对同期价格和收益仍有显著的正影响, 但日内波幅和订单不平衡程度的影响并不显著。即在此状态下, 高换手率会带来高收益, 但此时订单的增加是买单和卖单同时增加的结果, 高交易强度并没有带来股价的大幅波动, 这也是市场流动性好的表现。也就是说, 在这种机制状态下, 交易强度的增加是由于流动性需求的推动, 而不是信息交易的需要。因此可以判定, 机制III反映了无信息的市场状态。因此, 机制I、机制II和机制III分别表示的是坏消息、好消息和无信息的状态, 可将所求得的信息状态概率代入信息交易概率模型中做进一步分析。

#### 5.2 信息交易概率模型的估计及替代指标的计算

在实证分析时需要考虑样本所处的实际市场环境。传统的EKOP模型是基于报价驱动市场提出的, 模型分析的主要是做市商的信息获取过程。但中国的交易所市场并不是报价驱动的, 而是指令驱动的。指令驱动市场并不存在做市商, 此时的股价也不再是投资者与做市商博弈的结果, 而是市场中多、空双方势力的比拼。其实, 报价驱动市场最终的走势也

表1 机制转换模型估计结果  
Table 1 Estimation Output of the MSVAR Model

	机制 I				机制 II				机制 III			
	收益模型		价格模型		收益模型		价格模型		收益模型		价格模型	
	系数	t 值	系数	t 值	系数	t 值	系数	t 值	系数	t 值	系数	t 值
常数项	1.445	2.737	0.218	4.620	-0.341	-0.566	-0.086	-2.009	0.346	0.208	0.065	0.435
$r_{-1}$	-0.054	-1.685	0.040	14.257	-0.018	-0.437	0.043	13.921	-0.276	-2.861	0.033	3.705
$p_{-1}$	0.002	0.027	0.986	185.195	-0.061	-0.932	1.005	209.760	-0.214	-1.353	0.991	70.104
$StB$	-0.025	-7.940	-0.002	-5.888	0.015	3.165	0.001	1.759	0.012	0.923	0.001	0.532
$StB_{-1}$	0.002	0.809	0.000	1.279	0.004	1.241	0.001	2.688	0.005	0.512	-0.001	-0.834
$v$	-0.653	-10.180	-0.019	-3.360	0.701	10.287	0.032	6.972	0.050	0.266	-0.024	-1.421
$v_{-1}$	-0.098	-1.738	-0.023	-4.420	-0.110	-1.403	-0.005	-0.880	0.530	2.807	0.042	2.468
$TR$	-1.521	-3.953	-0.152	-4.555	0.717	2.649	0.062	3.564	2.685	4.406	0.255	4.664
$TR_{-1}$	1.065	3.442	0.118	4.394	-0.859	-3.171	-0.070	-3.905	-2.189	-3.346	-0.214	-3.673
$SE$	0.694		0.063		0.705		0.046		1.277		0.119	

注： $r_{-1}$ 为滞后一阶的日收益率， $p_{-1}$ 为滞后一阶的日均价， $StB_{-1}$ 为滞后一阶的订单不平衡程度， $v_{-1}$ 为滞后一阶的日内波幅， $TR_{-1}$ 为滞后一阶的换手率。依据统计学惯例，对于大样本数据( $n > 30$ )，以 $t > 1.700$ 为标准确定显著性。

是由市场中多、空双方势力决定的，做市商只不过充当了分离股票空头与多头的中介，他的存在有利于提高股票市场的流动性，降低噪声交易对市场的冲击，但却不会最终决定股票的实际价值和市场的整体走势。因此，针对指令驱动市场依然可以采用信息交易概率模型对不同信息状态下不同类型的投资者的交易状态进行分析，但由于此时的“做市商”不再是无形的市场主体，而是由所有投资者组成的市场集合，在模型实证中也需要做出一些调整，主要是对买、卖订单笔数估计方法的调整。

传统的信息交易概率模型主要考察买、卖交易的笔数，而没有考虑每笔交易的数量。Jones等<sup>[32]</sup>认为，报价驱动市场的交易次数几乎可以解释波动现象的全部，而每笔交易的大小甚至总交易量对波动的解释力都相对不足。因为报价驱动市场中的做市商面对的是独立订单，所以会对直观的订单笔数变化更敏感。但在指令驱动市场，买、卖订单的区分很难再简单地以笔数来衡量，市场整体所关心的也不再是一个个单独的订单，而是等待以不同价格成交的不同数量的委托买、卖订单的集合。因此，本研究选取的是内、外盘的总交易量指标，同时为了避免不同股票流通盘大小的差异，本研究以各股票单位时间内的平均交易量为标准对交易量进行转换，将交易量化为等价的交易笔数。实证结果显示，当所取单位交易量标准小幅变动时，不会影响模型的估计结果。

以2010年1月600000的日交易数据为例，除去元旦假期和周末，2010年1月共有20个有效的日交易数

据。将内、外盘交易量以及流通股指标和外生信息状况概率代入似然函数，可以得到表2所示的估计结果。

表2 改进的信息交易概率模型的估计结果  
Table 2 Estimation Output of the Improved PIN Model

指标	参数	估计值
交易活跃度	$\rho$	1.917%
信息交易概率	$\eta$	16.130%
主动买入比例	$\theta_B$	28.001%
主动卖出比例	$\theta_S$	29.108%
不交易比例	$1 - \theta_B - \theta_S$	42.891%
似然函数值	$\log\text{-likelihood}$	1 577.831

表2的估计结果显示了在交易数据背后的市场信息结构和交易者特征，该股票每天大概有1.917%发生交易，被信息交易者关注而发生的交易比例有近16.130%，而剩下的近84%的交易是由非信息交易者提出的，非信息交易者买入、卖出和不交易的比例分别为28.001%、29.108%和42.891%。

## 6 信息交易概率指标的估计结果及统计特性分析

运用同样的方法，本研究对每一支股票的信息



状态和信息交易概率参数进行估计,并通过统计箱图和描述统计分析直观地展示了不同时段、不同样本信息交易概率的波动范围和离散程度,具体见图4、图5以及表3和表4。

从图4和图5的截面和时序统计箱图可以看出,各样本、各时段的信息交易概率的平均水平都在15%~35%之间,即在通常情况下,市场中各股票发生的交易中大概有15%~35%的交易量是由信息交易者提出的,这样的信息交易比例应该比较适中。从取值范围看,各样本、各时段的信息交易概率的变化幅度及其取值分布情况的差异都比较大。信息交易概率超过50%的情况并不少见,个别样本某些时段的信息交易概率甚至接近80%。这说明在某些情况下,信息交易者会在市场中占主导地位,他们的交易会市场整体的走势。因此,不能因为市场大多数情况下的正常就忽视其中潜藏的信息风险危机。

为了更全面、准确地把握市场中信息交易概率的水平及其变化情况,本研究进一步对估计结果进行时序和截面统计。均值和方差分别显示了信息交易概率的平均水平和取值的离散状况。偏度和峰度主要反映信息交易概率取极端值的概率,峰度越高,分布曲线的尾部也就越厚,信息交易概率取极端值的概率就越大;偏度越高(低),信息交易概率就更有可能取到较大(小)的极端值。

各样本的截面统计结果按均值排序可以得到表3所示的结果。从表3可以发现,信息交易概率水平相近的个股,其所属的类别也具有一定的相似性。信息交易概率水平较低的股票(如安信信托、贵州茅台、中兴通讯、中储股份等),其标的资产主要是与金融、物流和资源等与稀有垄断资源相关的公司,这些公司较强的竞争优势和行业壁垒,加上所处行业较高的市场集中度,使利用信息对这些股票进行炒

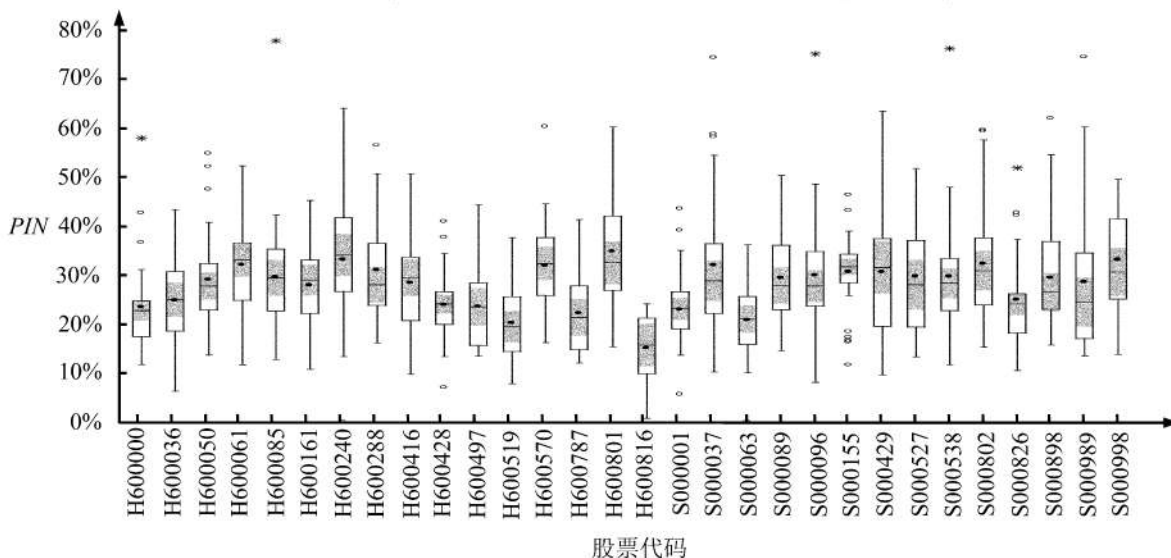


图4 各股票信息交易概率箱图  
Figure 4 Boxplots of Each Stock's PIN

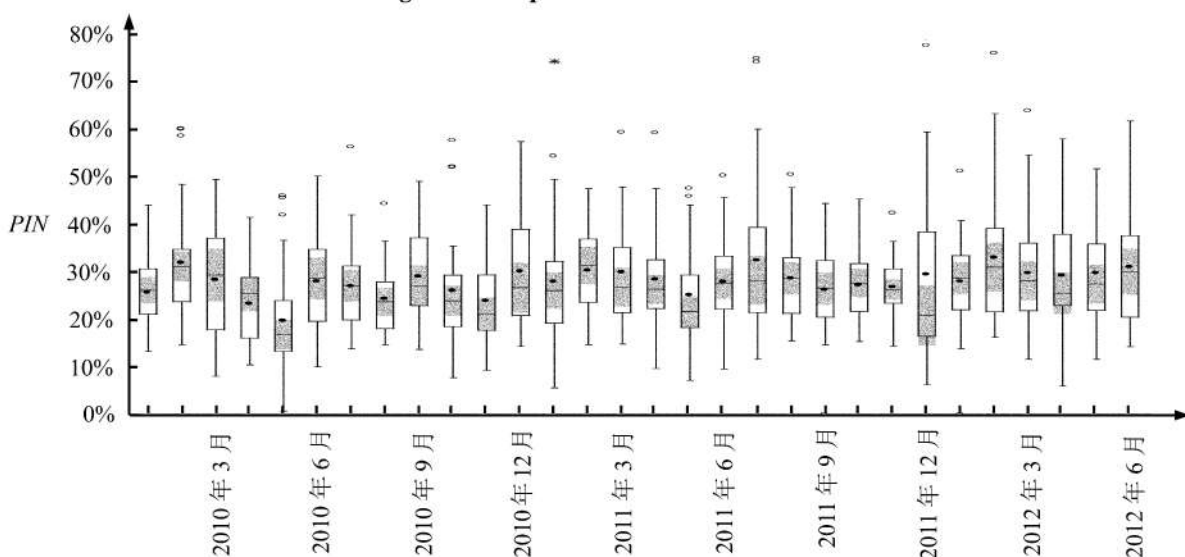


图5 各月份信息交易概率箱图  
Figure 5 Boxplots of Each Month's PIN

表3 各股票信息交易概率统计表(按均值升序排列)  
Table 3 Statistical Table of Each Stock's PIN (Ascending Means)

股票名称	股票代码	均值	方差	偏度	峰度	所属版块	所属类别
安信信托	H600816	0.161	0.013	0.810	1.261	券商信托	金融相关
贵州茅台	H600519	0.201	0.005	0.444	-0.126	酿酒行业	稀有垄断
中兴通讯	S000063	0.207	0.006	-0.344	1.126	通讯行业	通讯行业
中储股份	H600787	0.221	0.006	0.620	-0.089	交运物流	物流运输
平安银行	S000001	0.227	0.006	0.598	1.087	银行业	金融相关
驰宏锌锗	H600497	0.229	0.006	0.721	0.261	有色金属	稀有垄断
中远航运	H600428	0.238	0.004	0.242	1.625	港口水运	物流运输
浦发银行	H600000	0.246	0.018	2.859	9.445	银行业	金融相关
招商银行	H600036	0.247	0.007	0.086	-0.214	银行业	金融相关
桑德环境	S000826	0.249	0.009	1.098	1.400	公用事业	公用事业
天坛生物	H600161	0.278	0.008	0.063	-0.388	医药行业	医药行业
湘电股份	H600416	0.283	0.008	0.252	0.227	输配电气	输配电气
九芝堂	S000989	0.285	0.021	1.519	2.304	医药行业	医药行业
中国联通	H600050	0.289	0.010	0.970	1.017	通讯行业	通讯行业
深圳机场	S000089	0.293	0.007	0.678	0.190	民航机场业	民航机场业
鞍钢股份	S000898	0.293	0.013	1.150	1.220	钢铁行业	钢铁行业
同仁堂	H600085	0.295	0.015	2.038	7.799	医药行业	医药行业
云南白药	S000538	0.296	0.014	2.211	7.966	医药行业	医药行业
美的电器	S000527	0.296	0.012	0.353	-0.951	家电行业	家电行业
广聚能源	S000096	0.298	0.018	1.287	3.324	石油行业	能源化工
*ST川化	S000155	0.305	0.006	-0.637	0.392	化工行业	能源化工
粤高速 A	S000429	0.308	0.015	0.527	0.316	高速公路业	高速公路业
大恒科技	H600288	0.309	0.012	0.752	-0.339	电子信息业	高新科技
恒生电子	H600570	0.318	0.010	0.450	0.990	电子信息业	高新科技
深南电 A	S000037	0.319	0.023	1.082	0.852	电力行业	能源化工
中纺投资	H600061	0.320	0.010	0.063	-0.430	化纤行业	能源化工
隆平高科	S000998	0.320	0.021	0.888	3.174	农牧饲渔业	高新科技
华业地产	H600240	0.331	0.017	-0.165	0.807	房地产业	房地产相关
北京旅游	S000802	0.334	0.018	1.193	1.187	旅游酒店业	旅游业
华新水泥	H600801	0.347	0.013	0.706	0.092	水泥建材业	房地产相关



表4 各月份信息交易概率统计表

Table 4 Statistical Table of Each Month's PIN

月份	均值	方差	偏度	峰度
1	0.268	0.013	0.318	1.332
2	0.322	0.016	0.919	1.146
3	0.294	0.014	0.515	-0.016
4	0.267	0.011	0.326	0.340
5	0.245	0.012	0.698	0.422
6	0.290	0.012	0.436	0.037
7	0.297	0.017	1.160	1.816
8	0.263	0.007	0.771	0.677
9	0.274	0.008	-0.170	0.896
10	0.273	0.013	1.606	3.984
11	0.253	0.006	0.332	-0.098
12	0.298	0.024	1.056	0.546

作的可能性相对较小。因此,这些样本信息交易概率的均值和方差都较小,即这些股票的信息交易概率相对较低,且波动幅度不大。相对而言,信息交易概率水平较高的股票(如深南电A、隆平高科、北京旅游和华新水泥等)主要属于能源化工、高新科技、房地产相关和旅游行业,这些行业的特点是市场参与者较多,发展较快,话题性较强,容易被炒作而成为信息交易者关注的对象,所以这些股票自然拥有较高的信息交易概率。同时,这些股票较大的方差也反映了其信息交易概率分布较高的离散程度,这意味着信息交易者会在不同时期给予这些股票不同程度的关注,这是投资者对这些股票较强的投机倾向的反映,这些股票应成为信息不对称风险监控和管理中重点关注的对象。信息交易概率中等的股票(如九芝堂、中国联通、深圳机场和美的电器等)大多为医药、家电、公用事业等与民生相关的行业,这些股票的标的资产大多没有太大的垄断优势,话题性也相对较弱。从偏度和峰度看,市场中大多数样本都具有正偏和尖峰厚尾的分布,即使有个别偏度或峰度小于0的样本,其负偏和平峰的程度都非常低,更直接地说明市场中的信息交易概率风险不可忽视,即使大多数的情况下信息风险不会对市场造成太大影响,但信息不对称水平突然发生异常升高的状况并不是小概率事件,图4和图5中许多过大的异常值正是正偏和尖峰厚尾分布的直观反映。

从时序角度看,图5各月份信息交易概率的数据统计箱图显示出一定的季节趋势,对信息交易概率

按月份统计可以得到表4所示的结果。统计结果显示,PIN数值最低的3个月份依次是5月、11月和8月,这3个时间分别与股票的年报、半年报和季报的业绩公布时间有关。一般来说,一季度季报要在4月底公布完毕,半年报是在7月起至8月底公布完毕,第三季度的季报需要在10月底公布完毕,而年度报告则需在第二年的1月中旬起至4月底公布完毕。因此,每年的5月、8月和11月应该是上市公司信息集中披露完毕之后未公开信息最少的时刻,这些月份较低的信息交易水平合乎市场实际。而PIN数值最高的2月则与春节较长的年假休市有关,因为在较长的不可交易时段会有大量的信息发生,这些私有信息的存在增加了信息交易者凭借自身的优势获利的机会,信息交易概率相对较高也是正常的。因此,为了维护市场的公平和稳定,相关监管者应该在业绩公告之前的时段和较长的休市时段更多关注信息风险。

## 7 结论

本研究在详细考察信息交易概率模型的建模过程和假设条件的基础上,通过重新梳理信息交易者和非信息交易者在不同信息状态下的订单提交状况,提出考虑异质期望、交易活跃度和外生信息状态概率的改进的信息交易概率模型。实证结果表明,所构建的模型可以较好地刻画市场的实际信息状况,中国证券市场的信息交易概率水平与标的资产的垄断优势和话题性相关,垄断优势越强的股票信息交易概率越低,话题性越强的股票信息交易概率越高。因此,在防范信息风险时,应重点关注话题性较强的版块,特别是股价比较容易受舆论影响波动的个股。从时间上看,信息交易概率的变化也会受公告时间的影响,公告信息发布前后的信息交易概率变化较大,公告信息发布之后的信息交易概率会大幅减小。因此,相关管理者应该加强公告信息发布机制的制度和监管,在公告信息发布之前的时段更多关注市场波动和信息动向。

本研究构建的模型是在经典的信息交易概率模型的基础上,针对此类信息模型对外生信息状态的依赖问题以及经典模型对信息交易者在无信息状态下的行为特点考虑不足的问题提出的。通过本研究模型的修正,可以排除市场信息水平波动对信息交易概率估计的干扰,解决现有的信息不对称指标在不同信息状态下可比性较低的问题,从而提高市场信息不对称水平测度的准确性。在今后的研究中可以将这一指标应用到资产定价和风险管理中,深入研究信息风险对资产价格行为的影响,并对信息风险的管理提供理论支持。

## 参考文献:

- [1] Glosten L R, Milgrom P R. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders[J]. Journal of Financial Economics,

- 1985, 14(1):71-100.
- [2] Grossman S J, Stiglitz J E. On the impossibility of informationally efficient markets [J]. *The American Economic Review*, 1980, 70(3):393-408.
- [3] Easley D, Kiefer N M, O'Hara M, Paperman J B. Liquidity, information, and infrequently traded stocks [J]. *The Journal of Finance*, 1996, 51(4):1405-1436.
- [4] Easley D, Kiefer N M, O'Hara M. The information content of the trading process [J]. *Journal of Empirical Finance*, 1997, 4(2/3):159-186.
- [5] Easley D, López de Prado M M, O'Hara M. Flow toxicity and liquidity in a high frequency world [J]. *Review of Financial Studies*, 2012, 25(5):1457-1493.
- [6] Bagehot W. The only game in town [J]. *Financial Analysts Journal*, 1971, 27(2):12-14, 22.
- [7] Kyle A S. Continuous auctions and insider trading [J]. *Econometrica*, 1985, 53(6):1315-1335.
- [8] Copleand T E, Galai D. Information effects on the bid-ask spread [J]. *The Journal of Finance*, 1983, 38(5):1457-1469.
- [9] Glosten L R, Harris L E. Estimating the components of the bid/ask spread [J]. *Journal of Financial Economics*, 1988, 21(1):123-142.
- [10] Dufour A, Engle R F. Time and the price impact of a trade [J]. *The Journal of Finance*, 2000, 55(6):2467-2498.
- [11] 许敏, 刘善存. 基于VAR模型的知情交易者信息性交易概率研究 [J]. *北京航空航天大学学报: 社会科学版*, 2010, 23(5):69-71.  
Xu Min, Liu Shancun. The probability of informed trading based on VAR model [J]. *Journal of Beijing University of Aeronautics and Astronautics: Social Sciences Edition*, 2010, 23(5):69-71. (in Chinese)
- [12] Easley D, O'Hara M. Time and the process of security price adjustment [J]. *The Journal of Finance*, 1992, 47(2):577-605.
- [13] 郑振龙, 杨伟. 基于经典PIN模型的股票信息风险测度研究 [J]. *管理科学*, 2010, 23(6):91-99.  
Zheng Zhenlong, Yang Wei. Measuring information risk of stocks based on classical PIN model [J]. *Journal of Management Science*, 2010, 23(6):91-99. (in Chinese)
- [14] 许敏, 刘善存. 交易者市场到达率及影响因素研究 [J]. *管理科学学报*, 2010, 13(1):85-94.  
Xu Min, Liu Shancun. Arrival rate of traders and influencing factors [J]. *Journal of Management Sciences in China*, 2010, 13(1):85-94. (in Chinese)
- [15] 陈辉, 顾乃康, 朱雪嫣. 信息不对称、异质性交易与企业资本结构的决定 [J]. *山西财经大学学报*, 2012, 34(2):95-105.  
Chen Hui, Gu Naikang, Zhu Xueyan. Information asymmetry, heterogeneous trading and the determinants of capital structure [J]. *Journal of Shanxi Finance and Economics University*, 2012, 34(2):95-105. (in Chinese)
- [16] Aktas N, de Bodt E, Declerck F, Van Oppens H. The PIN anomaly around M&A announcements [J]. *Journal of Financial Markets*, 2007, 10(2):169-191.
- [17] 韩立岩, 郑君彦, 李东辉. 沪市知情交易概率 (PIN) 特征与风险定价能力 [J]. *中国管理科学*, 2008, 16(1):16-24.  
Han Liyan, Zheng Junyan, Li Donghui. The feature of probability of informed trading and risk pricing in Shanghai stock market [J]. *Chinese Journal of Management Science*, 2008, 16(1):16-24. (in Chinese)
- [18] Duarte J, Young L. Why is PIN priced? [J]. *Journal of Financial Economics*, 2009, 91(2):119-138.
- [19] 朱元琪, 刘善存. 交易信息含量与资产定价: 来自A股的经验证据 [J]. *系统工程*, 2010, 28(6):1-8.  
Zhu Yuanqi, Liu Shancun. Trading information content and asset pricing: Evidence from Chinese A shares [J]. *Systems Engineering*, 2010, 28(6):1-8. (in Chinese)
- [20] Andersen T G, Bondarenko O. VPIN and the flash crash [J]. *Journal of Financial Markets*, 2014, 17:1-46.
- [21] Easley D, López de Prado M M, O'Hara M. VPIN and the flash crash: A rejoinder [J]. *Journal of Financial Markets*, 2014, 17:47-52.
- [22] Andersen T G, Bondarenko O. Reflecting on the VPIN dispute [J]. *Journal of Financial Markets*, 2014, 17:53-64.
- [23] Easley D, O'Hara M. Price, trade size, and information in securities markets [J]. *Journal of Financial Economics*, 1987, 19(1):69-90.
- [24] Easley D, López de Prado M M, O'Hara M. The exchange of flow toxicity [J]. *The Journal of Trading*, 2011, 6(2):8-13.
- [25] Easley D, López de Prado M M, O'Hara M. The microstructure of the "flash crash": Flow toxicity, liquidity crashes, and the probability of informed trading [J]. *The Journal of Portfolio Management*, 2011, 37(2):118-128.
- [26] Hasbrouck J. Trading costs and returns for U. S. equities: Estimating effective costs from daily data [J]. *The Journal of Finance*, 2009, 64(3):1445-1477.
- [27] 张强, 刘善存, 邱苑华, 林千惠. 流动性特征对知情、非知情交易的影响研究 [J]. *管理科学学报*, 2013, 16(7):55-65.  
Zhang Qiang, Liu Shancun, Qiu Wanhua, Lin Qianhui. Informed trading, uninformed trading and liquidity

- ty [J]. *Journal of Management Sciences in China*, 2013, 16(7):55-65. (in Chinese)
- [28] 王春峰, 于婧晗, 房振明. 基于公司年报披露的交易量与信息不对称研究 [J]. *管理科学*, 2006, 19(6):74-80.  
Wang Chunfeng, Yu Jinghan, Fang Zhenming. Empirical research on information asymmetry and trading volume based on the disclosure of corporate earning announcements [J]. *Journal of Management Science*, 2006, 19(6):74-80. (in Chinese)
- [29] Gagnon L, Karolyi G A. Information, trading volume, and international stock return comovements: Evidence from cross-listed stocks [J]. *Journal of Financial and Quantitative Analysis*, 2009, 44(4):953-986.
- [30] Banerjee S, Kremer I. Disagreement and learning: Dynamic patterns of trade [J]. *The Journal of Finance*, 2010, 65(4):1269-1302.
- [31] 林虎, 刘冲. 投资者分歧、异常交易量和股票横截面收益率预测: 基于中国股票市场的经验证据 [J]. *投资研究*, 2011, 30(10):42-55.  
Lin Hu, Liu Chong. Investor divergence, abnormal trading volume and cross-sectional stock returns: Evidence from Chinese stock market [J]. *Review of Investment Studies*, 2011, 30(10):42-55. (in Chinese)
- [32] Jones C M, Kaul G, Lipson M L. Transactions, volume, and volatility [J]. *The Review of Financial Studies*, 1994, 7(4):631-651.

## Study on Measuring Information Risk Based on an Improved PIN Model

Yang Baochen<sup>1</sup>, Guo Can<sup>1</sup>, Chang Jianyong<sup>1,2</sup>

1 College of Management and Economics, Tianjin University, Tianjin 300072, China

2 Capital Normal University, Beijing 100048, China

**Abstract:** Measuring information asymmetry accurately is the research basis of financial market's information asymmetry issue. Based on the modeling and assumption of the classical PIN model, this paper examines different order submission behavior of information traders and non-information traders under different information states. Herein, an improved PIN model, which takes the information state probability as the exogenous variable and considers the heterogeneous beliefs and trading activity, is built up. The exogenous information state probability is estimated by Markov-switching vector autoregressive (MSVAR) model, which takes various market trading information into consideration, so the accuracy of the model could be improved. The empirical results indicate that the advanced model is more accurate in measuring the information asymmetry of financial market. First, the probability of informed trading (PIN) is related to the monopolistic advantage and topicality of underlying asset. On one hand, strong monopolistic advantage always results in a lower PIN; on the other hand, high topicality is usually accompanied by high PIN. Furthermore, the PIN vibration is influenced by reports. The PIN is always in a high level before the information being opened to the public, and subsequently decreases significantly after being reported. Therefore, for preventing information risk, it is needed to focus on the stocks that are easy to be influenced by public opinion and to pay more attention on the market fluctuations and the trends of information before announcement of related information released.

**Keywords:** information risk; PIN model; trading activity; heterogeneous beliefs; information state

**Received Date:** July 17<sup>th</sup>, 2014    **Accepted Date:** October 7<sup>th</sup>, 2014

**Funded Project:** Supported by the National Natural Science Foundation of China (71171144, 71471129) and the Specialized Research Fund for the Doctoral Program of Higher Education of China (20130032110016)

**Biography:** Dr. Yang Baochen, a Hebei Tangshan native (1966 -), graduated from Tianjin University and is a Professor and Ph. D. Advisor in the College of Management and Economics at Tianjin University. His research interests include economic system modeling and analysis, quantitative analysis of the financial markets and fixed income securities management, etc. E-mail: bchyang@tju.edu.cn □