

基于修正的 PIN 模型的股票信息风险测度研究^①

郑振龙¹，杨伟²

(1. 厦门大学金融系，福建 厦门，361005；2. 兴业证券，上海，200135)

摘要：对股票信息风险进行准确的测度无论对资产定价、风险管理还是市场绩效的衡量都有着重要意义。Easley, Kiefer, O'Hara and Paperman (1996)[1]最早提出了直接测度信息风险的经典 PIN 模型，此后该模型成为测度信息风险的炙手可热的模型。然而，PIN 模型隐含的买卖指令之间的负相关性与实际数据中买卖指令之间的正相关性并不相符。本文在 Easley, Kiefer, O'Hara and Paperman (1996)[1]提出的经典的 PIN 模型基础上，通过增加交易动机，提出了修正的 PIN 模型。本文基于中国股票的逐笔交易数据，利用修正的 PIN 模型对我国股票具有的信息风险进行的实证研究表明，修正的 PIN 模型隐含的买卖指令之间的相关性和买卖指令的方差能够更好地与实际数据相匹配。经典的 PIN 模型由于忽视了市场指令流冲击事件发生时引起的交易动机，倾向于高估股票具有的信息风险。

关键词：信息风险；测度；实证研究

中图分类号：F830

文献标识码：A

文章编号：479

0 引言

信息风险或信息不对称问题一直都是各国股票市场微观结构中备受关注的问题之一。当市场上存在信息风险或信息不对称的情形时——即一部分交易者相对于另外的交易者拥有与资产真实价值相关的私有信息时，这部分知情交易者利用其掌握的私有信息进行知情交易便可以从中获利^②。信息风险即是指投资者由于在某种资产上的信息不对称而遭受损失的可能性，它是对投资者面临的信息不对称程度的衡量，其大小通常用知情交易概率来衡量^③。对股票信息风险进行准确的测度无论对资产定价、风险管理还是市场绩效的衡量都有着重要意义。

由于知情交易不能从市场上直接观测到，早期的文献主要从间接的角度、采用一些知情交易的替代变量来衡量市场中的信息风险或信息不对称程度，如 Bagehot (1971)^[2]和 Jaffe and

^①基金项目：国家自然科学基金面上项目：非完美信息下基于观点偏差调整的资产定价（70971114）；教育部“国际金融危机应对研究”应急项目：金融市场的信息功能与金融危机预警（2009JYJR051）；福建省自然科学基金：卖空交易对证券市场的影响研究（2009J01316）。

^②从 Bagehot (1971)^[2]开始，Copeland and Galai (1983)^[3]、Glosten and Milgrom (1985)^[4]以及 Easley and O'Hara (1987)^[5]等提出证券市场的参与者可以区分为知情交易者（informed traders）和非知情交易者（uninformed traders）两类。所谓知情交易者指的是拥有与资产真实价值相关的私有信息的交易者；与此相对应，市场上除了知情交易者之外的交易者便是非知情交易者，这一类的交易者在交易时不是没有依据与资产真实价值相关的私有信息就是所依据的信息为公开信息或是谣言。

^③ Easley, Hvidkjaer and O'Hara (2002)^[12]最早用知情交易概率来衡量股票具有的信息风险。

Winkler (1976)^[6]等将买卖价差作为信息不对称的简单测度。但是这些方法都不能明确地测度交易者在市场中面临的信息不对称状况,而且这些结果都没有标准化,让人无法对不同市场上的信息风险或信息不对称的严重程度进行直接比较。Easley, Kiefer, O'Hara and Paperman (1996)^[1]最早提出了直接测度信息风险的 PIN 模型,此后, Easley, Kiefer and O'Hara (1997a)^[7]、Easley, Kiefer and O'Hara (1997b)^[8]、Easley, O'Hara and Paperman (1998)^[9]、Brockman and Chung (2000)^[10]、Easley, O'Hara and Saar (2001)^[11]、Easley, Hvidkjaer and O'Hara (2002)^[12]、Easley and O'Hara (2004)^[13]、Vega (2006)^[14]、Boehmer, Grammig and Theissen (2007)^[15]等大量学者以该模型为基础,对各个证券市场的信息风险进行了估计和应用。

郑振龙和杨伟 (2009a)^[16]对信息风险测度的国内外相关研究文献进行了较为全面的回顾和评述。虽然Easley, Kiefer, O'Hara and Paperman (1996)^[1]提出的PIN模型目前已成为对证券具有的信息风险进行测度的炙手可热的模型,但是也有不同学者对其提出了一些质疑(如 Venter and De Jongh (2006)^[17]和Duarte and Young (2008)^[18]等),认为该模型并不能准确测度股票信息风险。郑振龙和杨伟 (2009b)^[19]的研究发现PIN模型隐含的买卖指令之间的负相关性与实际数据中买卖指令之间的正相关性并不相符;此外, PIN模型中隐含的买卖指令的方差与实际数据中买卖指令相对较大的方差也不能很好地匹配。这证实了Venter and De Jongh (2006)^[17]和Duarte and Young (2008)^[18]的结论。

因此,我们有必要对Easley, Kiefer, O'Hara and Paperman (1996)^[1]的PIN模型进行修正,以便使得修正后的模型能够更加准确地对信息风险进行测度。本文的主要贡献之一是对Easley, Kiefer, O'Hara and Paperman (1996)^[1]提出的测度信息风险的经典PIN模型进行了修正。本文通过在交易动机中增加由于市场指令流冲击引起的交易,提出了修正的PIN模型。本文基于中国股票的逐笔交易数据,利用修正的PIN模型对我国股票具有的信息风险进行的实证研究表明,修正的PIN模型隐含的买单和卖单数量之间的相关性为正,这能够较好地匹配实际数据中买单和卖单数量之间的正相关关系。此外,修正的PIN模型也能够较好地匹配实际数据中买单和卖单数量较大的方差。

本文的另一个贡献是选取换手率作为交易活跃程度的衡量指标,利用修正的 PIN 模型对交易活跃程度不同的股票具有的信息风险进行了实证研究,发现股票具有的信息风险大小同交易活跃程度之间呈负相关关系,换手率越高的股票具有的信息风险越低。而经典的 PIN 模型由于忽视了市场指令流冲击事件发生时引起的交易动机,倾向于高估股票具有的信息风险。

本文的剩余部分安排如下:第一部分对 Easley, Kiefer, O'Hara and Paperman (1996)^[1]提出的测度信息风险的经典 PIN 模型进行了修正,提出了修正的 PIN 模型;第二部分对修正的 PIN 模型隐含的买卖指令的均值、方差以及买卖指令之间的相关性进行了推导;第三部分基于中国股票的逐笔交易数据,利用修正的 PIN 模型对交易活跃程度不同的股票具有的信息风险进行了测度;最后是结论。

1 修正的 PIN 模型

1.1 对 PIN 模型修正的思路

Easley, Kiefer, O'Hara and Paperman (1996)^[1]的 PIN 模型假定交易只有两种可能的动机,即根据公司特有的私人信息进行的知情交易和外生的流动性交易。与私人信息相关的交易发生在知情交易者收到私人信息的日子,如果私人信息是好消息,这会导致更大数量的买入指令,如果私人信息是坏消息则导致更大数量的卖出指令。结果是,大量的买入和卖出指令在不同日到达,这造成了买入和卖出指令之间的负相关关系(郑振龙和杨伟 (2009b)^[19])。

而 Hasbrouck and Seppi (2001)^[20]证明了指令流非平衡在不同的股票之间是相关的,这意味着存在驱动指令流非平衡的共同因素。而且, Chordia, Roll and Subrahmanyam (2002)^[21]

的研究表明整个市场的交易指令流非平衡对同时期的指数收益具有显著影响。整个市场的指令流非平衡不太可能是由于信息不对称引起的，因此可能存在其他造成非平衡的因素。

因此，在修正的PIN模型中，我们增加驱动整个市场范围内的交易增加的动机，即市场指令流冲击，以使得每天的买卖指令增加，从而使得模型隐含的买入指令和卖出指令之间的相关性与实际数据中买入指令和卖出指令之间的相关性一致（为正）并产生更接近于数据中观测到的买卖指令的方差。

市场指令流冲击的存在至少有两种原因。市场指令流冲击存在的一种可能的原因是使投资者产生分歧的公开信息事件的发生。这种对公开信息分歧使得买单和卖单的数量都增加。在文献中有关对公开信息事件的分歧是交易量的重要来源具有相当长的历史，这一历史可以追溯到Bachelier (1900)^[22]。正如Kandel and Pearson (1995)^[23]所发现的，观察到同样信息的投资者对于如何解释这些信息会产生分歧——如果投资者对于应该被用来解释这些信息的方式存在不同看法的话。大量的研究提供了对这一假说的实证证据，包括Harris and Raviv (1993)^[24]、Kandel and Pearson (1995)^[23]、Bessembinder, Chan and Seguin (1996)^[25]和Bamber, Barron and Stober (1999)^[26]等。市场指令流冲击存在的另外一种可能的原因是交易者仅仅在某个交易日合作交易以降低交易成本，具体论述可见Admati and Pfleiderer (1988)^[27]。

尽管市场指令流冲击的存在超过一种可能的原因，本文的目的并不是要在这些不同的原因之间进行区分。我们的目的是为了证明在经典的PIN模型中加入这一市场指令流冲击引起的交易动机，可以提高模型对实际数据的解释力度，这对于更加准确地测度信息风险具有重要意义。

1.2 修正的PIN模型的基本假设

为了将Easley, Kiefer, O'Hara and Paperman (1996)^[1]的模型应用到指令驱动交易机制下的我国股票市场上，本文假定存在一个隐形的做市商，知情交易者和非知情交易者通过提交限价委托单与隐形做市商进行交易。同Easley, Kiefer, O'Hara and Paperman (1996)^[1]一样，本文假设交易者在有限的天数内进行交易，但是一天之内的交易是连续的，假定一天之内的交易存在基于公司特有的私人信息进行的知情交易和由于市场指令流冲击引起的交易以及外生的流动性交易（即非知情交易）。假定由于市场指令流冲击引起的交易者、非知情交易者和知情交易者的到达由相互独立的泊松过程决定^①。来自于前一天的每个资产的所有信息价值在交易开始之前被揭示。

因此，本文修正后的PIN模型与原来经典的PIN模型存在以下区别。首先，本文对公司特有的知情事件发生时买者和卖者采用不同的知情交易到达率。本文用 μ^b 表示公司特有的知情事件发生时买者的知情交易到达率，用 μ^s 表示公司特有的知情事件发生时卖者的知情交易到达率。这样做的原因在于在实际数据中，买单数量的方差通常总是比卖单数量的方差更大，对模型做这一改动之后可以使新的模型更加能够与实际数据的特征相一致。其次，本文在模型中加入了驱动整个市场交易增加的因素，即市场指令流冲击。加入这一因素后可以使得在每个交易日中当市场指令流冲击事件发生时买单和卖单的数量都增加，从而使得模型隐含的买单和卖单数量之间的相关性为正，并且买单和卖单数量的变动也增加，以便更好地与实际数据的特征相匹配。在修正的PIN模型中，我们用 λ 表示市场指令流冲击事件发生的可能性。同样，本文也对市场指令流冲击事件发生时买者和卖者采用不同的交易到达率。本文用 γ_b 表示市场指令流冲击事件发生时买者的交易到达率，用 γ_s 表示市场指令流冲击事件

^① Easley, Kiefer and O'Hara (1997b)^[8]检验了不同交易日之间知情事件的独立性，发现他们不能拒绝不同交易日之间的知情事件相互独立的假设。因此，我们在修正的PIN模型中直接做出这一假定。

发生时卖者的交易到达率。

图1显示了在修正的PIN模型中股票在一个交易日的交易过程。如图1所示，在第一个节点即每个交易日开盘之前，自然决定该交易日市场指令流冲击事件是否发生。市场指令流冲击事件发生的概率为 λ ，市场指令流冲击事件没有发生的概率为 $1-\lambda$ 。自然接下来决定与资产真实价值相关的公司特有的知情事件是否发生，公司特有的知情事件发生的概率为 α ，公司特有的知情事件没有发生的概率则为 $1-\alpha$ 。如果知情事件发生，自然再将之区分为坏消息和好消息，其中，坏消息发生的概率为 δ ，好消息发生的概率为 $1-\delta$ 。给定已区分的事件之后，买卖单到达率都服从泊松分布。

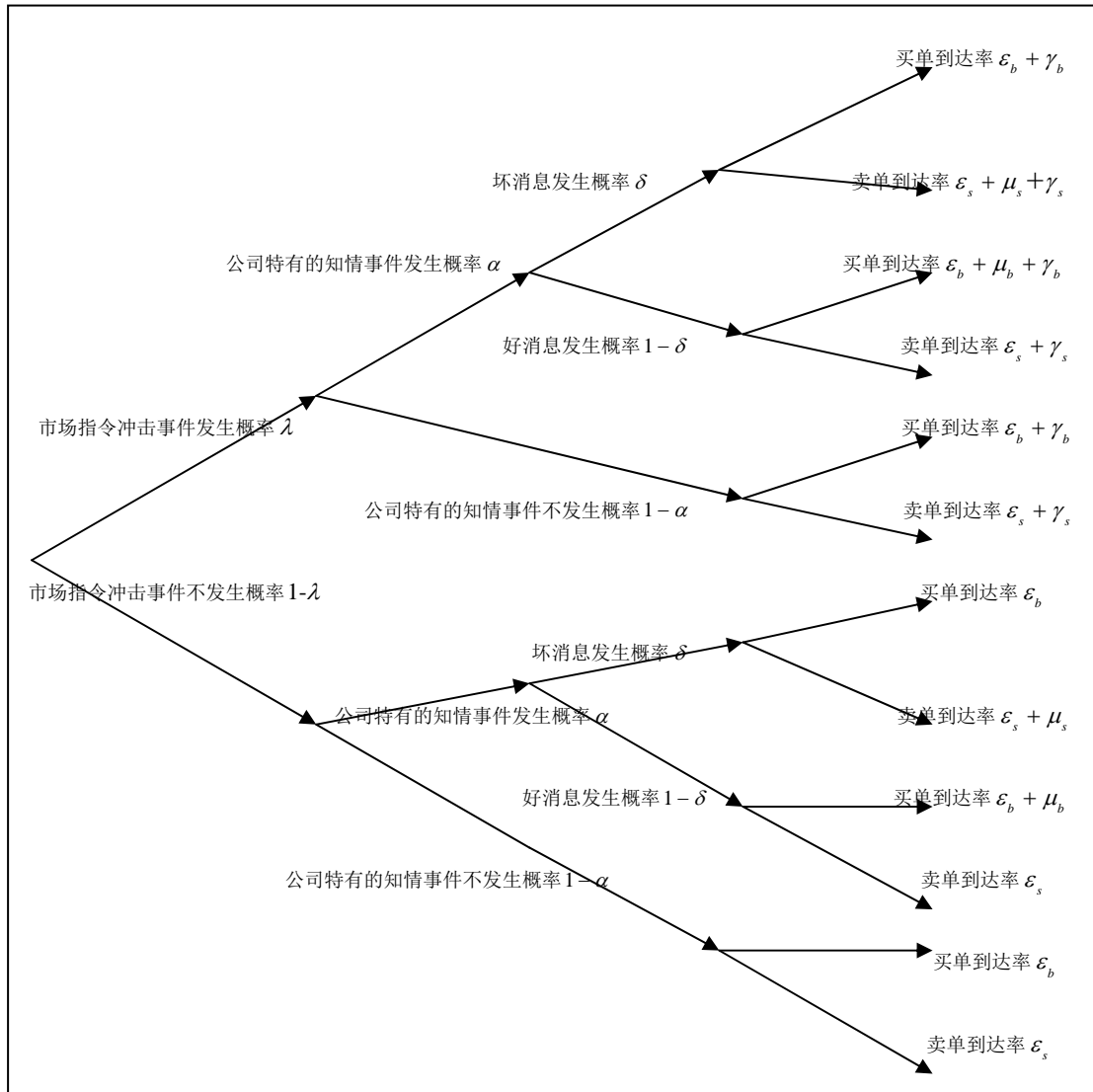


图 1 修正的 PIN 模型中股票在一个交易日的交易过程

当市场指令流冲击事件发生、且公司特有的坏消息发生时，买单到达率为 $\epsilon_b + \gamma_b$ ，卖单到达率为 $\epsilon_s + \mu_s + \gamma_s$ ；当市场指令流冲击事件发生、且公司特有的好消息发生时，买单到达率为 $\epsilon_b + \mu_b + \gamma_b$ ，卖单到达率为 $\epsilon_s + \gamma_s$ ；当市场指令流冲击事件发生、且公司特有的

知情事件没有发生时，买单到达率为 $\varepsilon_b + \gamma_b$ ，卖单到达率为 $\varepsilon_s + \gamma_s$ ；当市场指令流冲击事件没有发生、且公司特有的坏消息发生时，买单到达率为 ε_b ，卖单到达率为 $\varepsilon_s + \mu_s$ ；当市场指令流冲击事件没有发生、且公司特有的好消息发生时，买单到达率为 $\varepsilon_b + \mu_b$ ，卖单到达率为 ε_s ；当市场指令流冲击事件没有发生、且公司特有的知情事件也没有发生时，买单到达率为 ε_b ，卖单到达率为 ε_s 。

1.3 修正的 PIN 模型对信息风险的测度

类似地，信息风险可以由修正的 PIN 模型中知情交易者发出的交易指令占总交易指令的概率即新的知情交易概率（New Probability of Informed Trading, NPIN）来测度。在修正的 PIN 模型中，新的基于公司特有的私人信息的知情交易概率为：

$$NPIN = \frac{\alpha \times (\delta \times \mu_s + (1 - \delta) \times \mu_b)}{\alpha \times (\delta \times \mu_s + (1 - \delta) \times \mu_b) + \lambda \times (\gamma_b + \gamma_s) + \varepsilon_b + \varepsilon_s} \quad (1)$$

同时，我们也可以利用修正的 PIN 模型估计出市场指令流冲击事件发生时引起的交易概率（Probability of Market Order Flow Shock, PMOS）。在修正的 PIN 模型中，由于市场指令流冲击事件发生时引起的交易概率为：

$$PMOS = \frac{\lambda \times (\gamma_b + \gamma_s)}{\alpha \times (\delta \times \mu_s + (1 - \delta) \times \mu_b) + \lambda \times (\gamma_b + \gamma_s) + \varepsilon_b + \varepsilon_s} \quad (2)$$

与估计经典的 PIN 模型相比，对修正的 PIN 模型的参数估计需要估计出参数向量 $\theta = (\lambda, \alpha, \delta, \mu_b, \mu_s, \gamma_b, \gamma_s, \varepsilon_b, \varepsilon_s)$ ，在估计参数时只需要知道在一个交易日内买单的总数量和卖单的总数量。当不知道该交易日属于何种消息时，可以根据每一个分支发生的概率进行加权平均得到混合的似然函数。修正的 PIN 模型的混合似然函数可以写为：

$$\begin{aligned} L(\theta|B, S) = & \lambda \alpha \delta e^{-(\varepsilon_b + \gamma_b)} \frac{(\varepsilon_b + \gamma_b)^B}{B!} e^{-(\varepsilon_s + \mu_s + \gamma_s)} \frac{(\varepsilon_s + \mu_s + \gamma_s)^S}{S!} \\ & + \lambda \alpha (1 - \delta) e^{-(\varepsilon_b + \mu_b + \gamma_b)} \frac{(\varepsilon_b + \mu_b + \gamma_b)^B}{B!} e^{-(\varepsilon_s + \gamma_s)} \frac{(\varepsilon_s + \gamma_s)^S}{S!} \\ & + \lambda (1 - \alpha) e^{-(\varepsilon_b + \gamma_b)} \frac{(\varepsilon_b + \gamma_b)^B}{B!} e^{-(\varepsilon_s + \gamma_s)} \frac{(\varepsilon_s + \gamma_s)^S}{S!} \\ & + (1 - \lambda) \alpha \delta e^{-\varepsilon_b} \frac{(\varepsilon_b)^B}{B!} e^{-(\varepsilon_s + \mu_s)} \frac{(\varepsilon_s + \mu_s)^S}{S!} \\ & + (1 - \lambda) \alpha (1 - \delta) e^{-(\varepsilon_b + \mu_b)} \frac{(\varepsilon_b + \mu_b)^B}{B!} e^{-\varepsilon_s} \frac{(\varepsilon_s)^S}{S!} \\ & + (1 - \lambda) (1 - \alpha) e^{-\varepsilon_b} \frac{(\varepsilon_b)^B}{B!} e^{-\varepsilon_s} \frac{(\varepsilon_s)^S}{S!} \end{aligned} \quad (3)$$

由于交易日之间彼此独立， I 个观察日之内的买卖单数量 $M = (B_i, S_i)_{i=1}^I$ 的似然函数为

每天的似然函数的乘积:

$$L(M|\theta) = \prod_{i=1}^I L(\theta|B_i, S_i) \quad (4)$$

为了从任意的数据集 M 中估计出参数向量 θ , 只需要最大化方程 (4) 中的似然函数。

2 修正的 PIN 模型隐含的买单和卖单的均值、方差及相关性

本文接下来对修正的 PIN 模型隐含的买单和卖单的均值、方差及相关性进行推导。

修正的 PIN 模型中买单数量的期望值等于图 1 中每个分支的概率乘以该分支上买单到达率的加总:

$$\begin{aligned} E(B) &= \lambda\alpha\delta(\varepsilon_b + \gamma_b) + \lambda\alpha(1-\delta)(\varepsilon_b + \mu_b + \gamma_b) + \lambda(1-\alpha)(\varepsilon_b + \gamma_b) \\ &\quad + (1-\lambda)\alpha\delta\varepsilon_b + (1-\lambda)\alpha(1-\delta)(\varepsilon_b + \mu_b) + (1-\lambda)(1-\alpha)\varepsilon_b \\ &= \varepsilon_b + \alpha(1-\delta)\mu_b + \lambda\gamma_b \end{aligned} \quad (5)$$

类似地, 修正的 PIN 模型中卖单数量的期望值为:

$$\begin{aligned} E(S) &= \lambda\alpha\delta(\varepsilon_s + \mu_s + \gamma_s) + \lambda\alpha(1-\delta)(\varepsilon_s + \gamma_s) + \lambda(1-\alpha)(\varepsilon_s + \gamma_s) \\ &\quad + (1-\lambda)\alpha\delta(\varepsilon_s + \mu_s) + (1-\lambda)\alpha(1-\delta)\varepsilon_s + (1-\lambda)(1-\alpha)\varepsilon_s \\ &= \varepsilon_s + \alpha\delta\mu_s + \lambda\gamma_s \end{aligned} \quad (6)$$

买单数量的方差为 $\sigma^2(B) = E(B^2) - E^2(B)$, 同上面计算 $E(B)$ 的方式相同, 我们可

以计算出 $E(B^2)$, 利用上面计算出来的 $E(B)$ 的表达式, 我们可以得到:

$$\begin{aligned} \sigma^2(B) &= (\varepsilon_b)^2 + \alpha(1-\delta)(\mu_b)^2 + \lambda(\gamma_b)^2 + 2\alpha(1-\delta)\varepsilon_b\mu_b + 2\lambda\varepsilon_b\gamma_b \\ &\quad + 2\alpha(1-\delta)\lambda\mu_b\gamma_b - (\varepsilon_b + \alpha(1-\delta)\mu_b + \lambda\gamma_b)^2 \\ &= \alpha(1-\delta)(1-\alpha(1-\delta))(\mu_b)^2 + \lambda(1-\lambda)(\gamma_b)^2 \end{aligned} \quad (7)$$

卖单数量的方差为:

$$\begin{aligned} \sigma^2(S) &= (\varepsilon_s)^2 + \alpha\delta(\mu_s)^2 + \lambda(\gamma_s)^2 + 2\alpha\delta\varepsilon_s\mu_s + 2\lambda\varepsilon_s\gamma_s + 2\lambda\alpha\delta\mu_s\gamma_s \\ &\quad - (\varepsilon_s + \alpha\delta\mu_s + \lambda\gamma_s)^2 \\ &= \alpha\delta(1-\alpha\delta)(\mu_s)^2 + \lambda(1-\lambda)(\gamma_s)^2 \end{aligned} \quad (8)$$

买单和卖单数量之间的协方差为 $COV(B, S) = E(B \times S) - E(B) \times E(S)$, 以同样的

方式计算 $E(B \times S)$ 并利用上面计算出来的 $E(B)$ 和 $E(S)$ 的表达式, 我们可以得到:

$$\begin{aligned} COV(B, S) &= E(B \times S) - E(B) \times E(S) \\ &= \varepsilon_b\varepsilon_s + \alpha\delta\varepsilon_b\mu_s + \lambda\varepsilon_b\gamma_s + \alpha(1-\delta)\mu_b\varepsilon_s + \lambda\alpha(1-\delta)\mu_b\gamma_s \\ &\quad + \lambda\gamma_b\varepsilon_s + \lambda\alpha\delta\gamma_b\mu_s + \lambda\gamma_b\gamma_s \\ &\quad - (\varepsilon_b + \alpha(1-\delta)\mu_b + \lambda\gamma_b)(\varepsilon_s + \alpha\delta\mu_s + \lambda\gamma_s) \\ &= \lambda(1-\lambda)\gamma_b\gamma_s - \alpha^2\delta(1-\delta)\mu_b\mu_s \end{aligned} \quad (9)$$

从公式(9)中可以看出,通过增加市场指令流冲击引起的交易动机,可以使得修正的PIN模型隐含的买单和卖单数量之间的协方差为正,从而使得修正的PIN模型隐含的买单和卖单数量之间的相关性与实际数据中买单和卖单数量之间的正相关性一致。

3 中国股票的信息风险测度实证研究

3.1 数据及描述性统计

为了估计修正的PIN模型的参数,我们需要每只股票每天由买方和卖方发起的交易数量的数据。本文使用来自于万得金融高频数据库的分笔数据。万得金融高频数据库的分笔数据包含每个交易日内的每分钟内的每笔成交报价数据。分笔数据中包含证券代码、日期、时间、最新价、本笔成交量、本笔成交金额、买卖方向等共计三十六个字段的数据。由于该数据中包含了买卖方向,因此我们就不需要像国外学者一样在研究中对于没有提供买卖方向的交易数据通常采用Lee and Ready (1991)^[28]的买卖判断准则^①。由于数据的可得性,我们只有2007年一年内共计241个交易日的日内实时交易数据^②。我们选择的样本为上证180指数的成份股^③。

为了研究交易活跃程度不同的股票其信息风险大小之间的区别^④,本文按样本期内日均换手率的高低将上证180指数的成份股股票分为5组,为了避免新股上市首日的高换手率对分组的影响,本文剔除了上证180指数的成份股在2007年中上市的股票,这样样本中只有157只股票。

鉴于相邻两组之间的差别不是很大,为了更好地比较交易活跃程度不同的股票其信息风险大小之间的区别,本文选择换手率最低、换手率居中和换手率最高的三组股票作为我们最后的研究样本,每组股票平均在30只左右。本文剔除了发生在开盘之前的所有交易数据以避免包括发生在开盘期间的竞价交易。

表1列出了换手率低、换手率中和换手率高的三组股票在2007年内的241个交易日中平均每日换手率的描述性统计结果。

表1 各组股票在样本期内的日均换手率描述统计 (单位: %)

组别	均值	中位数	最大值	最小值	标准差	偏度	峰度	J-B	P 值
换手率低	1.80	1.98	2.53	0.77	0.60	-0.39	1.63	2.60	0.27
换手率中	3.77	3.69	4.17	3.28	0.28	0.04	1.62	1.98	0.37
换手率高	5.71	5.69	6.69	5.07	0.42	0.54	2.66	1.35	0.51

^① PIN模型的估计需要判断一笔交易是由买方发起还是由卖方发起,而大部分交易数据都不提供交易的买卖方向,对此国外学者通常采用Lee and Ready (1991)[28]的买卖判断准则,即将一笔交易的成交价与买卖报价的中点相比较,高于中点的这笔交易为卖,低于中点的交易为买,如果成交价与买卖报价的中点相同,则与上一笔成交价相比较,以此来判断出一笔交易的买卖方向。当然,这种判断准则不一定完全准确。Boehmer, Grammig and Theissen (2007)[15]的研究表明,如果将买卖方向错误分类的话,会导致下偏的PIN估计值。

^② Easley, Kiefer and O' Hara (1997a)[7]的研究表明只需要60个交易日的数据就可以充分获得PIN模型的可靠参数估计值。

^③ 基于其客观科学的选择方法,通常认为上证180指数的成份股具有很好的代表性,能够较好地代表上海股票市场的概貌和运行状况。

^④ 大部分学者都根据股票交易量的高低来衡量股票交易活跃程度(如Easley, Kiefer, O' Hara and Paperman (1996)[1]和Brockman and Chung (2000)[10]等),但是本文认为股票交易量的高低还受到股票自身规模的影响,而换手率能够更加准确地反映股票交易的活跃程度。

从表 1 中可以看出，三组股票在样本期内的日均换手率上存在较大的差异^①，换手率最低的一组股票的平均换手率为 1.80%，换手率居中的一组股票的平均换手率为 3.77%，换手率最高的一组股票的平均换手率为 5.71%。

表 2 列出了换手率最低、换手率居中和换手率最高的三组股票在样本期内的每日平均买单数量、卖单数量和样本期内的买单方差、卖单方差以及样本期内的买单和卖单数量之间的相关性在第 95 分位、第 75 分位、第 50 分位、第 25 分位和第 5 分位的统计数。

表 2 各组股票的买单和卖单数量、方差及其相关性的分位数统计

组别	类别	第 95 分位	第 75 分位	中位数	第 25 分位	第 5 分位
换手率低	买单数量	1232	1022	798	703	488
	卖单数量	1123	996	785	714	506
	买单方差	71699	60254	48957	38769	33857
	卖单方差	60106	47796	37870	28473	23036
	买卖单数量 相关性	0.69	0.60	0.50	0.43	0.30
	换手率中	买单数量	1264	1171	1088	935
卖单数量		1137	1096	993	902	751
买单方差		78223	60464	52148	47824	41658
卖单方差		63870	52347	43476	35982	27940
买卖单数量 相关性		0.75	0.68	0.60	0.54	0.43
换手率高		买单数量	1221	1131	1082	967
	卖单数量	1100	1052	998	913	850
	买单方差	77211	62894	58943	51057	45115
	卖单方差	59195	53550	45943	40349	34579
	买卖单数量 相关性	0.82	0.75	0.70	0.62	0.43

从表 2 中可以看出，换手率最低的一组股票的买单和卖单数量之间的相关性为从第 5 分位的 0.30 到第 95 分位的 0.69；换手率居中的一组股票的买单和卖单数量之间的相关性为从第 5 分位的 0.43 到第 95 分位的 0.75；换手率最高的一组股票的买单和卖单数量之间的相关性为从第 5 分位的 0.43 到第 95 分位的 0.82。可见，换手率不同的各组股票的买单和卖单数量之间的相关性都为正。

3.2 修正的 PIN 模型的估计结果

最大化公式 (4) 中的联合似然函数可以得到修正的 PIN 模型的参数估计值。然而对于该最大化问题也没有解析解，只能采用数值最大化技术来得到该最大化问题的解。本文采用同 Easley, Hvidkjaer and O'Hara (2005)^[29]类似的办法，通过对联合似然函数进行因式分解来使得数值最大化过程变得更加容易。通过因式分解后，联合似然函数的形式为：

^① 对各组股票换手率之间的差异进行的 Kruskal-Wallis 非参数检验结果在 1% 的置信度水平下拒绝了原假设，这表明各组股票换手率之间的差异在 1% 的置信度水平下是显著的。

$$\ln(L(B_i, S_i)) = \ln \left(\begin{aligned} & \lambda \alpha \delta e^{-\gamma_b - \mu_s - \gamma_s} (x_{b\mu})^{B_i - M_i} (x_{b\gamma})^{-M_i} (x_{b\mu\gamma})^{B_i - M_i} (x_{s\mu})^{S_i - M_i} (x_{s\gamma})^{S_i - M_i} (x_{s\mu\gamma})^{-M_i} \\ & + \lambda \alpha (1 - \delta) e^{-\mu_b - \gamma_b - \gamma_s} (x_{b\mu})^{B_i - M_i} (x_{b\gamma})^{B_i - M_i} (x_{b\mu\gamma})^{-M_i} (x_{s\mu})^{S_i - M_i} (x_{s\gamma})^{-M_i} (x_{s\mu\gamma})^{S_i - M_i} \\ & + \lambda (1 - \alpha) e^{-\gamma_b - \gamma_s} (x_{b\mu})^{B_i - M_i} (x_{b\gamma})^{-M_i} (x_{b\mu\gamma})^{B_i - M_i} (x_{s\mu})^{S_i - M_i} (x_{s\gamma})^{-M_i} (x_{s\mu\gamma})^{S_i - M_i} \\ & + (1 - \lambda) \alpha \delta e^{-\mu_s} (x_{b\mu})^{B_i - M_i} (x_{b\gamma})^{B_i - M_i} (x_{b\mu\gamma})^{B_i - M_i} (x_{s\mu})^{-M_i} (x_{s\gamma})^{S_i - M_i} (x_{s\mu\gamma})^{S_i - M_i} \\ & + (1 - \lambda) \alpha (1 - \delta) e^{-\mu_b} (x_{b\mu})^{-M_i} (x_{b\gamma})^{B_i - M_i} (x_{b\mu\gamma})^{B_i - M_i} (x_{s\mu})^{S_i - M_i} (x_{s\gamma})^{S_i - M_i} (x_{s\mu\gamma})^{S_i - M_i} \\ & + (1 - \lambda) (1 - \alpha) (x_{b\mu})^{B_i - M_i} (x_{b\gamma})^{B_i - M_i} (x_{b\mu\gamma})^{B_i - M_i} (x_{s\mu})^{S_i - M_i} (x_{s\gamma})^{S_i - M_i} (x_{s\mu\gamma})^{S_i - M_i} \end{aligned} \right)$$

$$+ \left\{ \begin{aligned} & -\varepsilon_b - \varepsilon_s + M_i \left[\ln(x_{b\mu}) + \ln(x_{b\gamma}) + \ln(x_{b\mu\gamma}) + \ln(x_{s\mu}) + \ln(x_{s\gamma}) + \ln(x_{s\mu\gamma}) \right] \\ & + B_i \left[\ln(\varepsilon_b + \mu_b) + \ln(\varepsilon_b + \gamma_b) + \ln(\varepsilon_b + \mu_b + \gamma_b) \right] - 2B_i \ln(\varepsilon_b) \\ & + S_i \left[\ln(\varepsilon_s + \mu_s) + \ln(\varepsilon_s + \gamma_s) + \ln(\varepsilon_s + \mu_s + \gamma_s) \right] - 2S_i \ln(\varepsilon_s) \end{aligned} \right\}$$

其中， $M_i = (\min(B_i, S_i) + \max(B_i, S_i)) / 2$ ， $x_{b\mu} = \frac{\varepsilon_b}{\varepsilon_b + \mu_b}$ ， $x_{b\gamma} = \frac{\varepsilon_b}{\varepsilon_b + \gamma_b}$ ，

$$x_{b\mu\gamma} = \frac{\varepsilon_b}{\varepsilon_b + \mu_b + \gamma_b}，x_{s\mu} = \frac{\varepsilon_s}{\varepsilon_s + \mu_s}，x_{s\gamma} = \frac{\varepsilon_s}{\varepsilon_s + \gamma_s}，x_{s\mu\gamma} = \frac{\varepsilon_s}{\varepsilon_s + \mu_s + \gamma_s}。$$

为了比较修正的 PIN 模型与经典的 PIN 模型之间对数据的拟合能力的高低，本文首先进行似然比检验^① (Likelihood Ratio Test)。修正的 PIN 模型与经典的 PIN 模型相比，增加了 λ 、 μ_s 、 γ_b 和 γ_s 这四个参数，因此，对于每只股票，检验的原假设为：

$$H_0 : \lambda = \mu_s = \gamma_b = \gamma_s = 0$$

在原假设下，没有市场范围内的指令流冲击事件发生，基于公司特有的私人信息进行交易的买单和卖单的到达率相同。我们用 $\log(L_0)$ 表示经典的 PIN 模型的对数似然函数值，

$\log(L_1)$ 表示修正的 PIN 模型的对数似然函数值。似然比检验的统计量为：

$$LR = -2(\log(L_0) - \log(L_1)) \square \chi_\alpha^2(n)$$

其中， α 为置信水平， n 为自由度。对修正的 PIN 模型与经典的 PIN 模型的似然比检验结果^②见表 3。

表 3 修正的 PIN 模型与经典的 PIN 模型的似然比检验

^① 修正的 PIN 模型与经典的 PIN 模型属于嵌套模型 (Nested Models)，对于这种嵌套模型优劣的比较通常采用似然比检验 (Likelihood Ratio Test)，似然比检验的统计量渐进服从 χ^2 分布。

^② 因为修正的 PIN 模型与经典的 PIN 模型相比增加了 4 个参数，因此似然比检验的统计量服从的 χ^2 分布的自由度为 4。本文对所有三组股票中的每只股票都分别进行了似然比检验，每只股票的检验结果都拒绝了原假设。限于篇幅的限制，本文在这里只给出各组似然比检验统计量的平均值。

组别	检验统计量	临界值
换手率低	6120	13.28
换手率中	6584	13.28
换手率高	7200	13.28

从表 3 中可以看出，在换手率最低、换手率居中和换手率最高这三组股票中，似然比检验的统计量都远远大于对应的临界值，我们可以全部拒绝原假设 H_0 。从似然比检验的结果可以看出，修正的 PIN 模型在对数据的拟合能力上明显优于经典的 PIN 模型。

对修正的 PIN 模型的参数估计结果见表 4。表 4 列出了换手率最低、换手率居中和换手率最高的三组股票在样本期内的修正的 PIN 模型的各项参数的估计值和信息风险在第 95 分位、第 75 分位、第 50 分位、第 25 分位和第 5 分位的统计数。

表 4 修正的 PIN 模型中各组股票的参数估计值及信息风险的分位数统计

组别	参数	第 95 分位	第 75 分位	中位数	第 25 分位	第 5 分位
换手率低	α	0.68	0.65	0.64	0.58	0.44
	δ	0.91	0.79	0.72	0.68	0.62
	λ	0.71	0.67	0.65	0.61	0.41
	ε_b	899.16	798.56	625.52	541.99	338.85
	γ_b	376.46	351.92	310.60	295.04	257.98
	μ_b	665.02	640.52	581.45	340.72	209.96
	ε_s	936.55	785.27	627.49	559.16	332.54
	γ_s	275.40	182.15	153.89	133.20	48.42
	μ_s	308.01	237.18	205.15	190.75	159.75
	NPIN	0.15	0.12	0.10	0.08	0.07
PMOS	0.28	0.20	0.17	0.14	0.11	
换手率中	α	0.67	0.64	0.60	0.58	0.53
	δ	0.90	0.77	0.71	0.69	0.66
	λ	0.71	0.67	0.65	0.62	0.57
	ε_b	937.05	890.08	818.19	744.24	601.47
	γ_b	398.99	373.78	355.63	319.62	302.62
	μ_b	672.55	635.61	560.53	300.62	234.94
	ε_s	958.75	859.05	801.45	720.53	577.51

	γ_s	211.55	188.47	151.39	106.02	64.46
	μ_s	313.58	238.04	196.13	171.75	140.96
	NPIN	0.11	0.09	0.08	0.07	0.05
	PMOS	0.21	0.18	0.15	0.14	0.11
	α	0.69	0.65	0.63	0.59	0.50
	δ	0.84	0.74	0.71	0.67	0.64
	λ	0.73	0.67	0.65	0.63	0.50
	ε_b	928.77	867.36	810.04	747.02	670.47
	γ_b	388.06	374.60	362.53	338.83	309.69
换手率高	μ_b	678.02	315.82	287.47	259.80	215.81
	ε_s	916.45	806.70	757.04	708.22	631.12
	γ_s	214.33	189.83	153.31	121.54	44.58
	μ_s	287.43	248.46	216.40	173.62	155.76
	NPIN	0.12	0.08	0.07	0.06	0.06
	PMOS	0.20	0.18	0.16	0.14	0.12

从表 4 中可以看出，知情事件发生的概率 α 在换手率最低的一组股票中的中位数为 0.64，其变动范围为从第 5 分位的 0.44 到第 95 分位的 0.68；知情事件发生的概率 α 在换手率居中的一组股票中的中位数为 0.60，其变动范围为从第 5 分位的 0.53 到第 95 分位的 0.67；知情事件发生的概率 α 在换手率最高的一组股票中的中位数为 0.63，其变动范围为从第 5 分位的 0.50 到第 95 分位的 0.69。可以看出换手率低的股票发生知情事件的概率高于换手率居中和换手率最高的股票。市场指令流冲击事件发生的概率 λ 在换手率最低的一组股票中的中位数为 0.65，其变动范围为从第 5 分位的 0.41 到第 95 分位的 0.71；市场指令流冲击事件发生的概率 λ 在换手率居中的一组股票中的中位数为 0.65，其变动范围为从第 5 分位的 0.57 到第 95 分位的 0.71；市场指令流冲击事件发生的概率 λ 在换手率最高的一组股票中的中位数为 0.65，其变动范围为从第 5 分位的 0.50 到第 95 分位的 0.73。可以看出无论是换手率最低的股票还是换手率居中和换手率最高的股票发生市场指令流冲击事件的概率都是大致相同的。

为了比较各组股票由修正的 PIN 模型估计得到的参数的分布在统计上是否存在显著差异，本文同样对各组股票的参数之间的差异进行 Kruskal-Wallis 非参数检验和 Mann-Whitney 非参数检验^①，检验结果见表 5。

表 5 修正的 PIN 模型估计得到的参数分布的 Kruskal-Wallis 和 Mann-Whitney 非参数检验

^① Kruskal-Wallis 非参数检验的原假设为多个独立样本来自的多个总体分布无显著差异；Mann-Whitney 非参数检验又称为 Wilcoxon Rank Sum 非参数检验，其原假设为如果两组样本的总体分布相同，那么两组样本都来自于同一个总体分布。

A 栏：参数分布的 Kruskal-Wallis 非参数检验			
参数	P 值		
α	0.2986		
δ	0.6213		
λ	0.8416		
ε_b	0.0003		
γ_b	0.0015		
μ_b	0.0014		
ε_s	0.0074		
γ_s	0.9367		
μ_s	0.5891		

B 栏：参数分布的 Mann-Whitney 非参数检验			
参数	低对中 P 值	低对高 P 值	中对高 P 值
α	0.2890	0.8228	0.1218
δ	0.9612	0.4031	0.4081
λ	0.6191	0.6190	0.9845
ε_b	5.1446e-04	6.3805e-04	0.6980
γ_b	0.0043	0.0011	0.5157
μ_b	0.8310	9.0701e-04	0.0038
ε_s	0.0059	0.0199	0.1870
γ_s	0.7415	0.9227	0.8009
μ_s	0.3417	0.8009	0.4492

从表5中可以看出,对各组股票由修正的PIN模型估计得到的参数进行的Kruskal-Wallis非参数检验结果表明,知情事件发生的概率 α 、知情事件发生时为坏消息的概率 δ 和市场指令流冲击事件发生的概率 λ 在换手率不同的各组股票之间并不存在显著的差异。而包括非知情交易者的买单到达率 ε_b 和卖单到达率 ε_s 、知情交易者的买单到达率 μ_b 和市场指令流冲击事件发生时的买单到达率 γ_b 都在99%的置信水平下拒绝了换手率不同的各组股票之间的分布相同的原假设,这表明这些参数在各组股票之间存在显著差异;根据Mann-Whitney非

参数检验，知情事件发生的概率 α 、知情事件发生时为坏消息的概率 δ 和市场指令流冲击事件发生的概率 λ 在换手率不同的各组股票之间并不存在显著的差异，这与前面的 Kruskal-Wallis 非参数检验结果相同。换手率最低的一组股票的非知情交易者的买单到达率 ε_b 和卖单到达率 ε_s 、市场指令流冲击事件发生时的买单到达率 γ_b 与换手率居中和换手率最高的股票之间在 99% 的置信水平下存在显著差异，而非知情交易者的买单到达率 ε_b 和卖单到达率 ε_s 、市场指令流冲击事件发生时的买单到达率 γ_b 在换手率居中和换手率最高的股票之间并不存在显著差异。换手率最高的一组股票的知情交易者的买单到达率 μ_b 与换手率最低和换手率居中的股票之间在 99% 的置信水平下存在显著差异。而知情交易者的卖单到达率 μ_s 和市场指令流冲击事件发生时的卖单到达率 γ_s 在换手率不同的各组股票之间并不存在显著差异。

从表 4 中可以看出，由修正的 PIN 模型估计得到的换手率最低、换手率居中和换手率最高的三组股票的信息风险 (PIN) 的中位数分别为 0.10、0.08 和 0.07，可见信息风险的大小 (NPIN 值) 随着交易活跃程度 (换手率) 的增加而下降。市场指令流冲击事件发生引起的交易概率 PMOS 的值在换手率最低、换手率居中和换手率最高的三组股票中的中位数分别为 0.17、0.15 和 0.16，可见市场指令流冲击事件发生引起的交易概率 PMOS 在交易活跃程度不同的三组股票之间并不明显差别。在各组股票中，NPIN 值都低于对应的 PMOS 值，这意味着交易在更大程度上是由于市场指令流冲击事件发生引起的，在经典的 PIN 模型中仅仅把交易动机区分为知情交易和流动性交易是不合理的，当市场指令流冲击事件发生时会在更大程度上促使投资者进行交易。在换手率最低的一组股票中，市场指令流冲击事件发生时引起的交易概率是最高的，结合郑振龙和杨伟 (2009b)^[19] 基于相同的数据利用经典的 PIN 模型估计得到的结果^①，我们可以得出以下结论，那就是经典的 PIN 模型由于没有考虑到市场指令流冲击事件发生时会在很大程度上引起投资者进行交易，从而高估股票具有的信息风险。

为了进一步从统计上检验各组股票的信息风险之间的差异，本文对各组股票由经典的 PIN 模型估计得到的信息风险进行 Kruskal-Wallis 非参数检验和 Mann-Whitney 非参数检验，检验结果见表 6。

表 6 修正的 PIN 模型估计得到的信息风险 Kruskal-Wallis 和 Mann-Whitney 非参数检验

A 栏：参数分布的 Kruskal-Wallis 非参数检验			
名称	P 值		
NPIN	0.0006		
PMOS	0.6788		
B 栏：参数分布的 Mann-Whitney 非参数检验			
名称	低对中 P 值	低对高 P 值	中对高 P 值
NPIN	0.0065	4.0078e-04	0.1310

^① 郑振龙和杨伟 (2009b) [19] 基于相同的数据利用经典的 PIN 模型估计得到的结果显示，换手率最低、换手率居中和换手率最高的三组股票的信息风险 (PIN) 的均值分别为 0.1090、0.1237 和 0.1238，这在总体上高于本文利用修正的 PIN 模型估计得到的换手率不同的三组股票的信息风险 (NPIN)。

PMOS	0.4185	0.5196	0.8146
------	--------	--------	--------

从表 6 中可以看出, Kruskal-Wallis 非参数检验的结果表明可以在 99% 的置信水平下拒绝各组股票之间的信息风险相同的原假设, 这意味着换手率不同的各组股票之间的信息风险确实存在显著的差异。而在 10% 的置信水平下也不能拒绝各组股票之间由市场指令流冲击事件发生引起的交易概率不存在显著差异的原假设, 这与我们的直觉是相符的, 市场指令流冲击事件发生时会引起所有股票的交易增加。根据 Mann-Whitney 非参数检验, 由修正的 PIN 模型估计出来的换手率最低的一组股票具有的信息风险在 99% 的置信水平下大于换手率居中和换手率最高的股票具有的信息风险, 这意味着交易活跃程度低的股票具有的信息风险高于交易活跃程度高的股票具有的信息风险。然而, 换手率居中的股票和换手率最高的股票具有的信息风险之间并不存在显著的差异。而各组股票由于市场指令流冲击事件发生时引起的交易概率之间都不存在显著的差异。

3.3 修正的 PIN 模型隐含的买卖指令的均值、方差和相关性与实际数据的比较

郑振龙和杨伟 (2009b)^[19]的研究发现经典的 PIN 模型中隐含的买单和卖单之间的负相关性与实际数据揭示的正相关性不相符合, 并且经典的 PIN 模型中隐含的买单和卖单的方差也远远低于实际数据揭示的买单和卖单的方差。我们希望修正后的 PIN 模型中隐含的买单和卖单的方差以及它们之间的相关性能够更好地与实际数据揭示的特征相匹配。利用对修正的 PIN 模型估计出来的参数, 可以计算出修正的 PIN 模型中隐含的买单和卖单数量的均值、方差及它们之间的相关性, 结果见表 7。

表 7 修正的 PIN 模型中隐含的各组股票的买卖单均值、方差及它们之间的相关性在不同分位数下的统计

组别	类别	第 95 分位	第 75 分位	中位数	第 25 分位	第 5 分位
换手率低	买单数量	1273	1058	908	760	590
	卖单数量	1089	972	822	739	507
	买单方差	98441	74718	55449	40154	30604
	卖单方差	22720	18810	17710	15206	12103
	买卖单数量相关性	0.42	0.26	0.19	0.13	0.05
	换手率中	买单数量	1262	1198	1116	1075
卖单数量		1108	1054	976	886	765
买单方差		89917	79497	55421	39508	35745
卖单方差		20816	18429	15723	14336	11322
买卖单数量相关性		0.49	0.24	0.14	0.05	0.01
换手率高		买单数量	1253	1154	1129	1038
	卖单数量	1035	979	957	890	837
	买单方差	65259	49988	42572	39361	31557
	卖单方差	22161	19502	17155	15725	14278
	买卖单数量相关性	0.51	0.31	0.16	0.08	0.002

从表 7 中可以看出, 修正的 PIN 模型在换手率最低、换手率居中和换手率最高的三组股票中隐含的买单和卖单之间的相关性都为正数, 对比表 2 中实际数据的买单和卖单之间

的正相关性，可以看出，修正的 PIN 模型隐含的买单和卖单之间的这种正相关性与实际数据中买单和卖单之间的正相关性能够很好地相匹配。此外，从表 7 中还可以看出，各组股票在修正的 PIN 模型中隐含的买单和卖单的方差都明显大于经典的 PIN 模型中隐含的买单和卖单的方差，这与实际数据中买单和卖单较大的方差更加匹配。

4 结论

Easley, Kiefer, O'Hara and Paperman (1996)^[1]提出的 PIN 模型已成为直接测度股票信息风险的经典模型，然而，郑振龙和杨伟（2009b）^[19]利用该模型对我国股票信息风险进行的实证研究表明，经典的 PIN 模型隐含的买单和卖单数量之间的负相关性与实际数据中买单和卖单数量之间的正相关性并不相符，而且经典的 PIN 模型隐含的买单和卖单数量的方差与实际数据中买卖指令相对较大的方差也不是很匹配。这意味着经典的 PIN 模型并不能对股票信息风险进行准确测度。这证实了 Venter and De Jongh (2006)^[17]和 Duarte and Young (2008)^[18]的结论。

本文在 Easley, Kiefer, O'Hara and Paperman (1996)^[1]的基础上，对经典的 PIN 模型进行了修正，提出了修正的 PIN 模型，与经典的 PIN 模型相比，修正的 PIN 模型在交易动机中增加了由于市场指令流冲击事件发生时引起的交易。本文利用修正的 PIN 模型对我国股票信息风险进行的实证研究表明，修正的 PIN 模型隐含的买单和卖单数量之间的相关性为正，这能够较好地匹配实际数据中买单和卖单数量之间的正相关关系。此外，修正的 PIN 模型也能够较好地匹配实际数据中买单和卖单数量较大的方差。

本文选取换手率作为交易活跃程度的衡量指标，本文利用修正的 PIN 模型对交易活跃程度不同的股票具有的信息风险进行的实证研究发现股票信息风险的大小同交易活跃程度之间呈负相关关系，换手率越高的股票其信息风险越低。这主要是因为经典的 PIN 模型忽视了市场指令流冲击事件发生时引起的交易动机，从而倾向于高估股票具有的信息风险。

参考文献:

- [1] Easley, D., N. M. Kiefer, M. O'Hara, and J. B. Paperman, 1996, Liquidity, information and infrequently traded stocks, *Journal of Finance* 51, 1405-1436.
- [2] Bagehot, W., 1971, The only game in town, *Financial Analysts Journal* 27, 12-14.
- [3] Copeland, T., and D. Galai, 1983, Information effects on the bid-ask spread, *Journal of Finance* 38, 1457-1469.
- [4] Glosten, L., and P. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71-100.
- [5] Easley, D., and M. O'Hara, 1987, Price, Trade Size, and Information in Security Markets, *Journal of Financial Economics* 19, 69-90.
- [6] Jaffe, J. F., and R. L. Winkler, 1976, Optimal speculation against an efficient market, *Journal of Finance* 31, 49-91.
- [7] Easley, D., N. M. Kiefer, and M. O'Hara, 1997a, One Day in the Life of a Very Common Stock, *Review of Financial Studies*, 3, 805-835.
- [8] Easley, D., N. M. Kiefer, and M. O'Hara, 1997b, The information content of the trading process, *Journal of Empirical Finance* 4, 159-186.
- [9] Easley, D., M. O'Hara, and J. Paperman, 1998, Financial analysts and information-based trade, *Journal of Financial Markets* 1, 175-201.
- [10] Brockman, P., and D. Y. Chung, 2000, Informed and Uninformed Trading in an Electronic,

- Order-Driven Environment, *Financial Review* 35, 125-146.
- [11] Easley, D., M. O'Hara, and G. Saar, 2001, How Stock Splits Affect Trading: A Microstructure Approach, *Journal of Financial and Quantitative Analysis* 36.
- [12] Easley, D., S. Hvidkjaer, and M. O'Hara, 2002, Is Information Risk a Determinant of Asset Returns?, *The Journal of Finance* 57, 2185-2221.
- [13] Easley, D., and M. O'Hara, 2004, Information and the Cost of Capital, *The Journal of Finance* 59, 1553-1583.
- [14] Vega, C., 2006, Stock price reaction to public and private information, *Journal of Financial Economics* 82, 103-133.
- [15] Boehmer, E., J. Grammig, and E. Theissen, 2007, Estimating the probability of informed trading—does trade misclassification matter?, *Journal of Financial Markets* 10, 26-47.
- [16] 郑振龙和杨伟, 2009a, 信息风险与资产定价研究述评, [J], *经济学动态*, forthcoming.
- [17] Venter, J. H., and D. C. J. De Jongh, 2006, Extending the EKOP model to estimate the probability of informed trading, *Studies in Economics and Econometrics* 30, 25-39.
- [18] Duarte, J. and L. Young, 2008, Why is PIN Priced?, *Journal of Financial Economics* forthcoming.
- [19] 郑振龙和杨伟, 2009b, PIN 模型能准确测度股票信息风险吗? 厦门大学工作论文.
- [20] Hasbrouck, J., and D. J. Seppi, 2001, Common factors in prices, order flows, and liquidity, *Journal of Financial Economics* 59, 383-411.
- [21] Chordia, T., R. Roll, and A. Subrahmanyam, 2002, Order imbalance, liquidity, and market returns, *Journal of Financial Economics* 65, 111-130.
- [22] Bachelier, L., 1900, Theorie de la Speculation (Thesis), *Annales Scientifiques de Ecole Normale Supérieure* 3, 21-86. (English Translation in Cootner, P. H., 1964, Random Character of Stock Market Prices, *MIT Press*, 17-78).
- [23] Kandel, E., and N. D. Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy*, 831-872.
- [24] Harris, M., and A. Raviv, 1993, Differences of opinion make a horse race, *Review of Financial Studies* 473-506.
- [25] Bessembinder, H., K. Chan, and P. J. Seguin, 1996, An empirical examination of information, differences of opinion, and trading activity, *Journal of Financial Economics* 40, 105-134.
- [26] Bamber, L. S., O. E. Barron, and T. L. Stober, 1999, Differential interpretations and trading volume, *Journal of Financial and Quantitative Analysis* 369-386.
- [27] Admati, A. R., and P. Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, *Review of Financial Studies* 3-40.
- [28] Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733-746.
- [29] Easley, D., N. M. Kiefer, and M. O'Hara, 2005, Factoring Information into Returns, *Working Paper, Cornell University*.

A Study on Measuring Information Risk of Stocks Based on the Corrected PIN Model

ZHENG Zhen-long¹, YANG Wei¹

(1. Department of Finance, Xiamen University, Xiamen, 361005, China)

Abstract: Measuring information risk of stocks accurately has important significance on asset pricing, risk management and measuring the market performance. Easley, Kiefer, O'Hara and Paperman (1996) first put forward the PIN model used to directly measure information risk of stocks, hereafter this model became one of the most popular models used to measure information risk. However, the negative correlation between buy and sell orders implied in the PIN model is not accordant with the positive correlation between buy and sell orders in real data. This paper put forward a corrected PIN model by adding trading motive based on the classical PIN model put forward by Easley, Kiefer, O'Hara and Paperman (1996). Based on the trading data tick by tick of Chinese stocks, this paper applied the corrected PIN model to measure information risk of stocks empirically, and found that the correlation between buy and sell orders and the variance of buy and sell orders implied in the corrected PIN model can better match the real data. The classical PIN model tended to overestimate information risk of stocks due to ignoring the trading motive aroused by market order flow shock.

Key Words: Information Risk; Measurement; Empirical Study