

考虑对称性冲击的时变信息风险测量^①

熊春连^{1,2}, 王春峰¹, 房振明³

(1. 天津大学管理与经济学部, 天津 300072; 2. 天津城建大学理学院, 天津 300384;
3. 渤海证券有限股份公司, 天津 300381)

摘要: 准确测量股票信息风险对资产定价、风险管理和市场业绩的衡量均具有重要的作用. 针对 Duarte 和 Young 模型中参数假设为常数这一缺陷, 采用已有金融文献使用交易量建模消息状态概率和对称性订单流冲击概率的方法, 允许 Duarte 和 Young 模型中的消息发生概率和对称性订单流冲击概率均时变, 从而提出了时变信息风险测量模型. 该模型可以测量日间和日内两个时间窗口的 APIN 和 PSOS. 进一步通过选取我国股市交易活跃股票的交易数据, 实证比较了所建构的时变信息风险测量模型与 Duarte 和 Young 模型对数据的拟合效果及其对价差的解释能力. 最后, 实证研究了我国股市的周内信息风险的特征.

关键词: 对称性冲击概率; 时变信息风险; 知情交易概率; 时变信息状态

中图分类号: F830 文献标识码: A 文章编号: 1007-9807(2015)06-0070-14

0 引 言

信息风险或者信息不对称问题是市场微观结构研究领域的主要问题之一. 当市场存在信息不对称时, 相对于其他投资者, 部分投资者因拥有关于资产真正价值的私有信息进行知情交易而获利. 因此, 信息风险意味着投资者由于对某种资产信息的不对称而遭受损失, 它是对投资者面临的信息不对称程度的衡量, 通常用知情交易概率测量信息风险的大小. 准确测量股票信息风险对资产定价、风险管理和市场业绩的衡量都具有重要的作用.

自 Easley 等^[1-3]开创性地提出利用买卖交易笔数直接测量信息风险的知情交易概率 (probability of informed trading, PIN) 模型 (以下简称 EKOP_PIN 模型) 以来, 许多金融文献利用该模型测量了股票市场的信息不对称程度^[3-11]. 随着 EKOP_PIN 模型的广泛应用和研究的深入, 同时

研究者对 EKOP_PIN 模型测量信息不对称程度的适用性产生了质疑^[12-16]. Aktas 等^[12]发现公司兼并公告前的 PIN 较兼并公告后的低, 这显然与兼并公告前有信息泄露的事实相矛盾. 他们认为这种异象可能跟 EKOP_PIN 模型假设消息发生概率为常数有关. Easley 等^[14]也指出现实市场中消息发生概率应该是时变的, 而不是常数. 与 EKOP_PIN 模型利用每日的买卖交易笔数估计知情交易概率不同, Tay 等^[15]则使用不规则交易间隔数据, 利用非对称自回归条件久期模型估计知情交易概率. 他们的模型通过允许消息发生概率时变, 扩展了 EKOP_PIN 模型的框架.

Duarte 和 Young^[16]从模型对数据拟合程度的角度出发, 发现 EKOP_PIN 模型中隐含的买卖订单流之间的负相关关系与实际数据中买卖订单流正相关关系并不相符. 他们认为 EKOP_PIN 模型假设中的交易动机过于简单. 在 EKOP_PIN 模型中交易者的交易动机分为非知情交易者由于流动

① 收稿日期: 2013-01-25; 修订日期: 2013-12-15.

基金项目: 国家自然科学基金资助项目(71271146); 教育部长江学者和创新团队发展计划资助项目(IRT1028).

作者简介: 熊春连(1976—), 女, 江西九江人, 博士生. Email: xiongchunlian@163.com

性需求引起的交易动机和知情交易者拥有私有信息而引起的交易动机两种. Duarte 和 Young 认为, 投资者对公共信息解读的差异可能引起买卖订单流的同时增加, 从而实际买卖订单流呈现正相关关系. 因此, 他们把对称性买卖订单流冲击(即买卖订单流同时增加)交易动机引入 EKOP_PIN 模型, 提出了一个修订的知情交易概率测量模型, 该模型测量的知情交易概率是调整的知情交易概率(adjusted PIN, APIN). 同时, 该模型还测量了交易来自对称性订单流冲击的概率, 即 PSOS(probability of symmetric order-flow shock), 该概率是交易来自买卖订单流冲击的无条件概率. Duarte 和 Young 发现, 在公共信息发布日交易量大幅增加, 而其他交易日交易量很低的公司通常 PSOS 较高. 因此, 他们认为 PSOS 是一个有效的非流动性代理指标. DY_APIIN-PSOS 模型(Duarte 和 Young 的模型简称) 虽然提高了模型对数据的拟合程度, 但该模型的消息状态概率和对称性订单流冲击概率均假设为常数.

由上述可见, 虽然 EKOP_PIN 模型是测量信息不对称程度的经典模型, 但由于该模型假设过于简单和理想化, 导致其适用性受到质疑. 于是, 研究者们不断在其理论框架下完善该模型, 使其更符合市场实际. 一方面, 研究者将对称性买卖订单流冲击交易动机引入 EKOP_PIN 模型, 另一方面, 研究者允许 EKOP_PIN 模型中的消息发生概率时变. 但已有的文献在扩展 EKOP_PIN 模型时, 并没有将上述两方面同时考虑进去. 本文则试图在同时考虑上述两方面的情况下扩展 EKOP_PIN 模型, 即提出考虑对称性冲击和消息状态时变的时变信息风险测量模型.

针对 DY_APIIN-PSOS 模型的消息状态概率和对称性订单流冲击概率均假设为常数的缺陷, 将已有金融文献使用交易量建模消息状态概率和对称性订单流冲击概率的方法引入 Duarte 和 Young 模型中, 使得 Duarte 和 Young 模型中的消息发生概率和对称性订单流冲击概率均时变, 从而扩展了 Duarte 和 Young 模型, 即提出了考虑对称性冲击和消息状态时变的时变信息风险测量模型(记作 NV_APIIN-PSOS 模型). NV_APIIN-PSOS 模型考虑了买卖订单流的对称性冲击和允许消息发生概率是时变的, 因此, 从引入对称性买卖订单流冲击

交易动机和允许消息发生概率时变两方面扩展了经典的 EKOP_PIN 模型. NV_APIIN-PSOS 模型不仅提出了测量时变知情交易概率的方法而且还提出了测量时变对称性订单流冲击交易概率(即交易来自对称性订单流冲击的无条件概率)的方法. 该模型可以测量日间和日内两种时间窗口的知情交易概率和对称性订单流冲击交易概率, 这是一个很大的突破, 使得 PIN 的应用领域更为深入和广泛; 此外, 本文实证测量了我国股票市场的时变知情交易概率和时变对称性订单流冲击交易概率. 在模型的估计过程中采用一种新方法彻底解决了参数估计过程中由于日内交易笔数过大或买卖订单笔数过于不平衡引起数值溢出而导致参数估计失败的问题, 而这个估计失败问题是以往 PIN 模型族的参数估计过程中常见的问题.

1 消息时变的知情交易概率模型的构建与计算方法

1.1 EKOP_PIN 模型和 DY_APIIN-PSOS 模型

1.1.1 EKOP_PIN 模型

在 EKOP_PIN 模型中, B_d 和 S_d 分别记作是 d 日的买单总数和卖单总数, 并假定它们是服从独立泊松分布的随机变量. 在不同的消息状态下, B_d 和 S_d 到达强度不同. θ_E 记作每日消息发生的概率, θ_B 记作发生消息的条件下, 发生坏消息的条件概率. 因此, 每日的消息状态有 3 种(即发生坏消息、好消息及不发生消息这 3 种状态), 其概率分别是: $\pi_B = \theta_E \theta_B$, $\pi_C = \theta_E (1 - \theta_B)$, $\pi_N = (1 - \theta_E)$. 在无消息发生日, B_d 和 S_d 的泊松到达率分别为 λ_1 和 λ_{-1} ; 在发生坏消息日, 卖单的泊松到达率增加一个常数 δ , 而买单的泊松到达率保持不变; 在好消息发生日, 买单的泊松到达率增加一个常数 δ , 而卖单的泊松到达率保持不变. EKOP_PIN 模型中的 PIN 计算的是来自知情交易者的交易到达强度与总的交易者的交易到达强度的相对比值, 即

$$PIN = \frac{\theta_E \delta}{\lambda_1 + \lambda_{-1} + \theta_E \delta} \quad (1)$$

从方程(1)可知 PIN 与知情交易者和非知情交易者的交易到达率以及消息发生概率有关.

1.1.2 DY_APIN-PSOS 模型

DY_APIN-PSOS 模型扩展了 EKOP_PIN 模型,他们不仅允许知情买单的到达率(δ₁)和知情卖单的到达率(δ₋₁)不同,而且允许某些交易日(甚至无消息发生的交易日)的买卖订单流同时增加,即在模型中引入对称性订单流冲击动机.在DY_APIN-PSOS 模型中,发生坏消息且发生共同订单流冲击的交易日、发生好消息且发生共同订单流冲击的交易日以及无消息但发生共同订单流冲击的交易日

APIN = (θ_E [(1 - θ_B) δ₁ + θ_B δ₋₁]) / (λ₁ + λ₋₁ + θ_E [(1 - θ_B) δ₁ + θ_B δ₋₁] + θ_C (Δ₁ + Δ₋₁)) (2)

定义对称性订单流冲击交易概率(即交易来自对称性冲击的无条件概率) PSOS 为

PSOS = (θ_C (Δ₁ + Δ₋₁)) / (λ₁ + λ₋₁ + θ_E [(1 - θ_B) δ₁ + θ_B δ₋₁] + θ_C (Δ₁ + Δ₋₁)) (3)

其中当 δ₁ = δ₋₁ 且 θ_C = 0 时, APIN 即为 PIN.

1.2 NV_APIN-PSOS 模型的建构

1.2.1 消息状态概率及对称性订单流冲击概率的建模

本文在DY_APIN-PSOS 模型的基础上提出消息时变的知情交易概率模型,该模型与DY_APIN-PSOS 模型的主要区别是假定 θ_E、θ_B 及 θ_C 不再是固定的常数,而是时变的. Tay 等^[15] 使用不规则交易间隔数据,利用非对称自回归条件久期模型估计知情交易概率.他们的模型利用交易量建模消息状态概率,从而允许消息状态概率是时变的. Preve 等^[17] 则将对称性买卖订单流冲击交易动机引入 Tay 等^[15] 的模型中,并利用交易量建模了时变的对称性订单流冲击发生概率.借鉴 Tay 等^[15] 对时变消息状态概率的建模方法和 Preve 等^[17] 对时变对称性订单流冲击发生概率的建模方法,将其引入DY_APIN-PSOS 的模型中,即

对 θ_E 和 θ_B 的建模采用 Tay 等^[15] 的建模方法,对 θ_C 的建模则采用 Preve 等^[17] 的方法.下面对 θ_E、θ_B 及 θ_C 的建模进行介绍.

关于交易量和信息之间的正相关关系的实证结果^[18-20] 启发 Duarte 和 Young 不再假设 θ_E 及 θ_B 为固定的常数,而是利用交易量建模消息状态概率,从而允许消息状态概率是时变的.具体地,他们记 V^B 及 V^S 分别为日平均买量和日平均卖量, V^B_d 及 V^S_d 则为 d 日的主动性买量和主动性卖量.则假设 d 日发生消息的概率为

θ_{Ed} = 1 - 1 / (1 + exp{γ₁ + γ₂ [ln(V^B_d + V^S_d) - ln(V^B + V^S)]}) (4)

其中 γ₂ > 0, γ₁ 可正可负.

假设 d 日发生消息的条件下,发生坏消息的概率为

θ_{Bd} = 1 - 1 / (1 + exp{γ₃ [ln(V^S_d - V^S) - γ₄ ln(V^B_d - V^B)]}) (5)

其中 γ₃ > 0, γ₄ > 0. 因此 d 发生坏消息、好消息及不发生消息的概率分别为 π_{Bd} = θ_{Ed}θ_{Bd}, π_{Gd} = θ_{Ed}(1 - θ_{Bd}), π_{Nd} = 1 - θ_{Ed}.

Preve 等^[17] 扩展了 Tay 等^[15] 的模型,并建模了 d 日对称性订单流冲击发生的概率,即假设为

θ_{Cd} = 1 - 1 / (1 + exp[γ₅ (lnV^B_d - lnV^B) + (lnV^S_d - lnV^S)]) (6)

其中 γ₅ ≥ 0, (u)₊ = {u, 若 u > 0; 0, 若 u ≤ 0}. 因此 d 日发生坏消息且发生对称性订单流冲击、发生好消息且发生对称性订单流冲击及不发生消息但发生对称性订单流冲击的概率分别为 π_{CB} = π_{Bd}θ_{Cd},

π_{CG} = π_{Gd}θ_{Cd}, π_{CN} = π_{Nd}θ_{Cd}.

1.2.2 NV_APIN-PSOS 模型对非对称信息及对称性订单流冲击交易概率的测量

与 DY_APIN-PSOS 模型类似, NV_APIN-PSOS 模型用调整的知情交易概率作为非对称信息的测量,从而得到 d 日的 APIN 为

$$APIN_d = \frac{\theta_{Ed} [(1 - \theta_{Bd}) \delta_1 + \theta_{Bd} \delta_{-1}]}{\lambda_1 + \lambda_{-1} + \theta_{Ed} [(1 - \theta_{Bd}) \delta_1 + \theta_{Bd} \delta_{-1}] + \theta_{Cd} (\Delta_1 + \Delta_{-1})} \quad (7)$$

与此同时, d 日的对称性订单流冲击交易概率 PSOS 为

$$PSOS_d = \frac{\theta_{Cd} (\Delta_1 + \Delta_{-1})}{\lambda_1 + \lambda_{-1} + \theta_{Ed} [(1 - \theta_{Bd}) \delta_1 + \theta_{Bd} \delta_{-1}] + \theta_{Cd} (\Delta_1 + \Delta_{-1})} \quad (8)$$

式(7)、(8)中的分子和分母中参数如果是日内时段的数值,则式(7)、(8)可以用来计算日内 APIN 和 PSOS.

1.2.3 NV_APIN-PSOS 模型的参数估计

为了估计模型参数,首先求出 D 个交易日 NV_APIN-PSOS 模型的似然函数,然后极大化模型的对数似然函数. NV_APIN-PSOS 模型 d 日的似然函数是

$$\begin{aligned} L_d &= L_d(\theta_{Ed}, \theta_{Bd}, \theta_{Cd}, \lambda_1, \lambda_{-1}, \delta_1, \delta_{-1}, \Delta_1, \Delta_{-1} | B_d, S_d) \\ &= (1 - \theta_{Ed}) \theta_{Cd} \frac{(\lambda_1 + \Delta_1)^{B_d}}{B_d!} e^{-(\lambda_1 + \Delta_1)} \frac{(\lambda_{-1} + \Delta_{-1})^{S_d}}{S_d!} e^{-(\lambda_{-1} + \Delta_{-1})} + (1 - \theta_{Ed}) (1 - \theta_{Cd}) \frac{\lambda_1^{B_d}}{B_d!} e^{-\lambda_1} \frac{\lambda_{-1}^{S_d}}{S_d!} e^{-\lambda_{-1}} + \\ &\quad \theta_{Ed} \theta_{Bd} \theta_{Cd} \frac{(\lambda_1 + \Delta_1)^{B_d}}{B_d!} e^{-(\lambda_1 + \Delta_1)} \frac{(\lambda_{-1} + \delta_{-1} + \Delta_{-1})^{S_d}}{S_d!} e^{-(\lambda_{-1} + \delta_{-1} + \Delta_{-1})} + \\ &\quad \theta_{Ed} \theta_{Bd} (1 - \theta_{Cd}) \frac{\lambda_1^{B_d}}{B_d!} e^{-\lambda_1} \frac{(\lambda_{-1} + \delta_{-1})^{S_d}}{S_d!} e^{-(\lambda_{-1} + \delta_{-1})} + \\ &\quad \theta_{Ed} (1 - \theta_{Bd}) \theta_{Cd} \frac{(\lambda_1 + \delta_1 + \Delta_1)^{B_d}}{B_d!} e^{-(\lambda_1 + \delta_1 + \Delta_1)} \frac{(\lambda_{-1} + \Delta_{-1})^{S_d}}{S_d!} e^{-(\lambda_{-1} + \Delta_{-1})} + \\ &\quad \theta_{Ed} (1 - \theta_{Bd}) (1 - \theta_{Cd}) \frac{(\lambda_1 + \delta_1)^{B_d}}{B_d!} e^{-(\lambda_1 + \delta_1)} \frac{\lambda_{-1}^{S_d}}{S_d!} e^{-\lambda_{-1}} \end{aligned} \quad (9)$$

则 D 个交易日的对数似然函数是

$$L = \sum_{d=1}^D \ln L_d(\theta_{Ed}, \theta_{Bd}, \theta_{Cd}, \lambda_1, \lambda_{-1}, \delta_1, \delta_{-1}, \Delta_1, \Delta_{-1} | B_d, S_d) \quad (10)$$

极大化似然函数(10)即得模型的参数估计,但当交易笔数较大或买卖交易笔数及其不平衡时,极大化似然函数(10)估计模型参数会由于数据向上或向下溢出的问题导致参数估计失败. 为了避免这个问题,采用以下方法对似然函数(9)做如

下处理:

第一步,提取式(9)中的公因子 $\frac{\lambda_1^{B_d}}{B_d!} e^{-\lambda_1} \frac{\lambda_{-1}^{S_d}}{S_d!} \times e^{-\lambda_{-1}}$ 并取自然对数,则有

$$\begin{aligned} \ln(L_d) &= B_d \ln(\lambda_1) + S_d \ln(\lambda_{-1}) - (\lambda_1 + \lambda_{-1}) - (\ln(B_d! S_d!)) + \\ &\quad \left\{ \begin{aligned} &(1 - \theta_{Ed}) \theta_{Cd} \exp[-(\Delta_1 + \Delta_{-1}) - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \Delta_1}\right) - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \Delta_{-1}}\right)] + \\ &(1 - \theta_{Ed}) (1 - \theta_{Cd}) + \\ &\theta_{Ed} \theta_{Bd} \theta_{Cd} \exp[-(\Delta_1 + \Delta_{-1} + \delta_{-1}) - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \Delta_1}\right) - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \Delta_{-1} + \delta_{-1}}\right)] + \\ &\theta_{Ed} \theta_{Bd} (1 - \theta_{Cd}) \exp[-\delta_{-1} - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \delta_{-1}}\right)] + \\ &\theta_{Ed} (1 - \theta_{Bd}) \theta_{Cd} \exp[-(\Delta_1 + \Delta_{-1} + \delta_1) - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \Delta_1 + \delta_1}\right) - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \Delta_{-1}}\right)] + \\ &\theta_{Ed} (1 - \theta_{Bd}) (1 - \theta_{Cd}) \exp[-\delta_1 - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \delta_1}\right)] \end{aligned} \right\} \quad (11) \end{aligned}$$

第二步 由于 $\ln(B_d! S_d!)$ 是常数 不影响参数的极大似然估计结果 仅影响似然值的观测值 因此进行

极大似然估计时可以去掉对数似然函数中该项 第三步 继续对似然函数进行处理: 令

$$\begin{aligned}
 x_1 &= -(\Delta_1 + \Delta_{-1}) - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \Delta_1}\right) - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \Delta_{-1}}\right); \\
 x_2 &= -(\Delta_1 + \Delta_{-1} + \delta_{-1}) - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \Delta_1}\right) - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \Delta_{-1} + \delta_{-1}}\right) \\
 x_3 &= -\delta_{-1} - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \delta_{-1}}\right) \\
 x_4 &= -(\Delta_1 + \Delta_{-1} + \delta_1) - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \Delta_1 + \delta_1}\right) - S_d \ln\left(\frac{\lambda_{-1}}{\lambda_{-1} + \Delta_{-1}}\right) \\
 x_5 &= -\delta_1 - B_d \ln\left(\frac{\lambda_1}{\lambda_1 + \delta_1}\right)
 \end{aligned}$$

$$\begin{aligned}
 X_{\max} &= \max\{x_1, x_2, x_3, x_4, x_5, 0\}; X_1 = x_1 - X_{\max}; X_2 = x_2 - X_{\max}; \\
 X_3 &= x_3 - X_{\max}; X_4 = x_4 - X_{\max}; X_5 = x_5 - X_{\max};
 \end{aligned}$$

第四步 在极大化似然函数(10)时 对数似然函数 $\ln(L_d)$ 取值为

$$B_d \ln(\lambda_1) + S_d \ln(\lambda_{-1}) - (\lambda_1 + \lambda_{-1}) + X_{\max} + \ln \left\{ \begin{aligned} &(1 - \theta_{Ed}) \theta_{Cd} \exp(X_1) + \\ &(1 - \theta_{Ed}) (1 - \theta_{Cd}) \exp(-X_{\max}) + \\ &\theta_{Ed} \theta_{Bd} \theta_{Cd} \exp(X_2) + \\ &\theta_{Ed} \theta_{Bd} (1 - \theta_{Cd}) \exp(X_3) + \\ &\theta_{Ed} (1 - \theta_{Bd}) \theta_{Cd} \exp(X_4) + \\ &\theta_{Ed} (1 - \theta_{Bd}) (1 - \theta_{Cd}) \exp(X_5) \end{aligned} \right\} \quad (12)$$

在以往 PIN 模型族估计中由于似然函数中有指数项,当指数过大或者过小时,常因产生数据溢出,导致估计失败.虽然许多研究者都试图解决这一问题,但当交易笔数的非常大或者买卖交易及其不平衡时,仍无法彻底解决^[14,16,21].本文对似然函数做上述处理后可以彻底解决极大似然估计失败的问题.式(12)中的 $B_d \ln(\lambda_1) + S_d \ln(\lambda_{-1}) - (\lambda_1 + \lambda_{-1}) + X_{\max}$ 不可能会出现数据溢出的情形,而式(12)中最后一项中的 $\exp(X_1)$ 、 $\exp(X_2)$ 、 $\exp(X_3)$ 、 $\exp(X_4)$ 、 $\exp(X_5)$ 及 $\exp(-X_{\max})$ 中的 X_1 、 X_2 、 X_3 、 X_4 、 X_5 及 $-X_{\max}$ 均小于等于0,当这些指数为很小的负数时会出现数据向下溢出的情形,计算机显示指数运算结果为0(笔者计算机所采用的 matlab 软件出现向下溢出的临界值是 -708,即 $\exp(-708)$ 的值显示为0),但 X_1 、 X_2 、 X_3 、 X_4 、 X_5 及 $-X_{\max}$ 中至少有一项为0,因此指数运算结果至少有一个值为1.从而整个

$$\ln \left\{ \begin{aligned} &(1 - \theta_{Ed}) \theta_{Cd} \exp(X_1) + \\ &(1 - \theta_{Ed}) (1 - \theta_{Cd}) \exp(-X_{\max}) + \\ &\theta_{Ed} \theta_{Bd} \theta_{Cd} \exp(X_2) + \\ &\theta_{Ed} \theta_{Bd} (1 - \theta_{Cd}) \exp(X_3) + \\ &\theta_{Ed} (1 - \theta_{Bd}) \theta_{Cd} \exp(X_4) + \\ &\theta_{Ed} (1 - \theta_{Bd}) (1 - \theta_{Cd}) \exp(X_5) \end{aligned} \right\}$$

不会因为溢出而导致极大似然估计失败,而此时所计算的似然值是真正似然值的高度近似.可见,采用此方法彻底解决了参数估计过程中由于数据溢出导致不能估计的问题.虽然这会损失计算的精度,但相对参数估计失败而言,这种损失显然是值得的.

2 实证分析

2.1 样本选择与数据描述

选取交易活跃的5只股票:中国联通(股票

代码为 600050)、紫金矿业(股票代码为 601899)、国电电力(股票代码为 600795)、万科 A(股票代码为 000002)和中国石化(股票代码为 600028)。5 只股票的样本期是 2009 年 1 月 5 日至 2010 年 6 月 30 日。本文数据来源于 CS-

MAR 高频交易数据库中的分笔交易数据。处理高频数据时,选取交易日的正常交易时间,即 9:30-11:30 和 13:00-15:00 内的连续竞价数据,剔除了开盘集合竞价交易数据和收盘后的交易数据。

表 1 5 只样本股票数据的描述性统计特征

Table 1 Descriptive statistics of 5 sample stocks

	中国联通	紫金矿业	国电电力	万科 A	中国石化
买单比例(%)	50.65	51.52	49.81	51.87	50.11
卖单比例(%)	49.35	48.48	50.19	48.13	49.89
平均日交易笔数	2 643.45	2 587.41	2 342.00	2 740.89	2 554.82
买笔数的均值	1 338.86	1 332.96	1 166.51	1 421.71	1 280.23
卖笔数的均值	1 304.59	1 254.44	1 175.49	1 319.18	1 274.58
平均对数交易量	10.09	9.95	8.84	9.87	9.23
平均交易量	70 445.97	57 743.72	20 291.19	52 514.43	31 572.43
观测总数	948 999	926 292	786 913	983 979	919 734

从表 1 可知: 样本股票的买卖单比例比较接近,并未出现文献中卖比明显高于买比的情形^[16]。中国联通的平均日交易量最大,国电电力的平均日交易量最小。万科 A 的平均日交易笔数最大,国电电力的日平均交易笔数最小。

2.2 各种模型的极大似然估计结果

EKOP_PIN 模型、DY_APIN-PSOS 模型及 NV_APIN-PSOS 模型的极大似然估计均采用遗传算法计算。每只股票随机选取初始点,并优化 1 000 次,然后取这 1 000 次中似然值最大时的参数估计为本文的估计结果。

表 2 EKOP_PIN 模型的估计结果

Table 2 Results of the EKOP_PIN model

	中国联通	紫金矿业	国电电力	万科 A	中国石化
λ_1	1 228.084 *** (0.499)	609.292 *** (0.491)	985.976 *** (0.458)	1 259.569 *** (1.034)	1 145.148 *** (0.499)
λ_{-1}	1 119.700 *** (0.503)	1 246.046 *** (0.439)	1 102.285 *** (0.437)	1 247.200 *** (1.028)	1 165.574 *** (0.514)
δ	342.622 *** (0.772)	752.175 *** (0.653)	364.510 *** (0.817)	261.287 *** (0.906)	307.433 *** (0.798)
θ_E	0.903 *** (0.014)	0.981 *** (0.008)	0.689 *** (0.030)	0.897 *** (0.018)	0.800 *** (0.023)
θ_B	0.634 *** (0.028)	0.007 ** (0.003)	0.283 *** (0.034)	0.309 *** (0.026)	0.501 *** (0.030)
π_N	0.097	0.019	0.311	0.103	0.200
π_C	0.330	0.974	0.494	0.620	0.399
π_B	0.573	0.007	0.195	0.278	0.401
PIN	0.116	0.285	0.107	0.086	0.096
AIC	26 727.288	24 605.810	32 275.051	13 274.859	30 879.687
SBC	26 730.063	24 608.580	32 277.683	13 277.635	30 882.469

注: *, **, *** 依次表示在 10%、5%、1% 显著性水平下显著。括号中数值为标准误差。

表3 DY_APIN-PSOS 模型的估计结果
Table 3 Results of the DY_APIN-PSOS model

	中国联通	紫金矿业	国电电力	万科 A	中国石化
θ_E	0.567 *** (0.049)	0.770 *** (0.034)	0.796 *** (0.851)	0.240 *** (1.765)	0.786 *** (1.076)
θ_B	0.500 *** (0.038)	0.089 ** (0.039)	0.561 *** (0.978)	0.375 *** (1.661)	0.350 *** (1.073)
θ_C	0.339 *** (0.046)	0.898 *** (0.023)	0.802 *** (1.285)	0.480 *** (2.790)	0.796 *** (1.399)
λ_1	1 141.172 *** (0.832)	574.577 *** (1.058)	697.035 *** (1.359)	1 178.880 *** (3.429)	984.898 *** (1.791)
λ_{-1}	1 114.401 *** (0.780)	979.881 *** (1.246)	779.256 *** (1.024)	1 169.476 *** (2.329)	967.879 *** (1.277)
δ_1	379.821 *** (1.655)	365.473 *** (1.326)	325.985 *** (1.158)	450.102 *** (2.302)	338.308 *** (1.391)
δ_{-1}	460.270 *** (2.071)	1 001.175 *** (4.971)	256.001 *** (0.024)	439.362 *** (0.037)	261.495 *** (0.023)
Δ_1	207.559 *** (1.642)	561.744 *** (1.411)	422.483 *** (0.031)	229.477 *** (0.045)	170.595 *** (0.029)
Δ_{-1}	215.935 *** (1.792)	310.171 *** (1.422)	339.914 *** (0.023)	171.898 *** (0.050)	269.103 *** (0.026)
π_N	0.433	0.230	0.204	0.760	0.214
π_B	0.284	0.068	0.446	0.090	0.275
π_C	0.284	0.701	0.350	0.150	0.511
APIN	0.090	0.122	0.099	0.040	0.096
PSOS	0.054	0.294	0.264	0.073	0.137
AIC	23 681.649	20 976.947	25 064.008	15 949.552	22 107.532
SBC	23 686.645	20 981.932	25 068.745	15 954.547	22 112.539

注: *, **, *** 依次表示在 10% 5% 1% 显著性水平下显著. 括号中数值为标准误差.

表2和表3分别是EKOP_PIN模型和DY_APIN-PSOS模型的估计结果,该结果表明, EKOP_PIN模型和DY_APIN-PSOS模型的所有估计系数在5%的显著性水平上均显著. EKOP_PIN模型估计的5只股票的PIN值均较DY_APIN-PSOS模型估计的APIN值要大.这与Duarte和Young的结果一致.他们认为PIN值较APIN值要大的原因是PIN估计的不仅仅是知情交易概率而且还估计了对称性订单流冲击交易概率.

EKOP_PIN模型的估计结果表明:不发生消息的概率较低,而发生消息的概率较高,尤其是发生好消息的概率.除中国联通外,其余股票发生好消息的概率均大于发生坏消息的概率.而DY_APIN-PSOS模型的估计结果表明:除个别股票外,

该模型估计的发生好消息及发生坏消息的概率值均较EKOP_PIN模型估计的值有所减小.这可能是因为EKOP_PIN模型估计的发生好消息及发生坏消息的概率包含了对称性订单流冲击的概率.而DY_APIN-PSOS模型估计的对称性订单流冲击概率 θ_C 较大,介于0.339-0.898,且交易来自对称性订单流冲击的概率PSOS也较大,介于0.054-0.294. DY_APIN-PSOS模型估计的不发生消息的概率较EKOP_PIN模型估计的值总体有所上升,这可能是EKOP_PIN模型中发生消息的概率包含了对称性订单流冲击概率.以上结果分别表明,测量我国股票市场的信息风险也应考虑买卖订单流的对称性冲击.

表4是NV_APIN-PSOS模型的估计结果.该

结果表明：在 5% 的显著性水平下，所有模型参数除中国联通和紫金矿业的 γ_2 及紫金矿业的 γ_4 外，其余均显著。 γ_2 不显著说明消息发生的概率并不是时变的，但发生好消息和坏消息的概率则是时变的。 γ_4

不显著说明好消息和坏消息发生的概率仅受 γ_3 的影响。所有股票估计的 $\gamma_2, \gamma_3, \gamma_4, \gamma_5$ 均大于 0，这与前面的理论相符。但所有股票估计的 γ_1 大于零，这意味着平均每日发生消息的概率大于 0.5。

表 4 NV_APIN-PSOS 模型的估计结果

Table 4 Results of the NV_APIN-PSOS model

	中国联通	紫金矿业	国电电力	万科 A	中国石化
λ_1	1 202.660 *** (0.719)	524.250 *** (0.779)	934.068 *** (0.490)	1 261.417 *** (1.574)	1 081.780 *** (0.722)
λ_{-1}	1 188.386 *** (0.694)	1 200.221 *** (0.520)	1 056.260 *** (0.546)	1 243.517 *** (1.243)	1 159.901 *** (0.654)
δ_1	313.734 *** (1.627)	820.600 *** (0.960)	360.380 *** (1.107)	276.871 *** (2.145)	318.396 *** (1.369)
δ_{-1}	311.433 *** (1.563)	768.014 *** (3.522)	243.506 *** (1.170)	256.730 *** (2.691)	277.528 *** (1.385)
Δ_1	259.761 *** (2.914)	581.887 *** (2.945)	354.713 *** (1.810)	16.353 *** (2.387)	272.007 *** (2.650)
Δ_{-1}	86.412 *** (6.449)	6.915 *** (1.626)	28.964 *** (1.712)	190.881 *** (7.980)	38.824 *** (2.044)
γ_1	1.122 *** (0.177)	3.617 *** (0.404)	1.884 *** (0.212)	1.540 *** (0.198)	1.738 *** (0.176)
γ_2	0.147 (0.217)	0.439 (0.282)	3.156 *** (0.378)	0.678 ** (0.263)	1.178 *** (0.277)
γ_3	7.195 *** (2.006)	2.458 ** (1.012)	5.195 *** (1.330)	7.101 *** (1.443)	6.580 *** (1.805)
γ_4	5.334 *** (1.864)	1.282 (0.914)	3.088 *** (1.013)	6.143 *** (1.315)	4.032 *** (1.428)
γ_5	0.742 *** (0.287)	0.426 ** (0.197)	7.184 *** (0.830)	1.306 *** (0.366)	4.410 *** (0.419)
AIC	19 664.720	24 057.699	21 125.807	12 146.920	18 567.519
SBC	19 670.826	24 063.792	21 131.597	12 153.026	18 573.638

注：*，**，*** 依次表示在 10%、5%、1% 显著性水平下显著。括号中数值为标准误差。

表 5 是 NV_APIN-PSOS 模型估计的主要变量描述性统计结果。该结果表明，大体而言（除个别情况外），发生消息的概率 $1 - \pi_{Nd}$ 的均值比 EKOP_PIN 模型估计的 π_N 值要小，但比 DY_APIN-PSOS 模型估计的 π_N 值要大，而 π_{Cd} 和 π_{Bd} 也具有相同的结果。该模型估计的调整的知情交易概率 $APIN_d$ 的均值大体介于 EKOP_PIN 模型估计的 PIN 值和 DY_APIN-PSOS 模型估计的 APIN 值

之间。该模型估计的对称性订单流冲击概率 θ_{Cd} 的均值较 DY_APIN-PSOS 模型的估计值 θ_C 有较大幅度的下降，其值介于 0.040 - 0.189，而估计交易来自对称性订单流冲击的概率 $PSOS_d$ 的均值较 DY_APIN-PSOS 模型的估计值 $PSOS$ 也下降很多，其值介于 0.006 - 0.028。以上结果产生的原因可能是：由于 NV_APIN-PSOS 模型允许消息状态概率及对称性订单流冲击概率时变，模型变得更精确了。

表 5 NV_APIN-PSOS 模型估计的主要变量描述性统计结果

Table 5 Descriptive statistics of main estimated variables by the NV_APIN-PSOS model

	中国联通	紫金矿业	国电电力	万科 A	中国石化
π_{Nd}					
最小值	0.134	0.013	0.001	0.102	0.038
最大值	0.391	0.161	1.000	0.408	0.909
均值	0.195	0.032	0.281	0.199	0.185
标准差	0.030	0.014	0.266	0.065	0.095
π_{Gd}					
最小值	0.027	0.000	0.000	0.011	0.017
最大值	0.791	0.855	0.783	0.852	0.776
均值	0.415	0.539	0.352	0.410	0.428
标准差	0.193	0.179	0.181	0.176	0.183
π_{Bd}					
最小值	0.170	0.130	0.000	0.130	0.083
最大值	0.841	0.954	0.920	0.883	0.938
均值	0.472	0.448	0.447	0.474	0.474
标准差	0.161	0.178	0.210	0.153	0.187
θ_{Cd}					
最小值	0.000	0.000	0.000	0.000	0.000
最大值	0.674	0.526	1.000	0.668	0.999
均值	0.053	0.040	0.189	0.082	0.150
标准差	0.117	0.094	0.327	0.153	0.285
$APIN_d$					
最小值	0.084	0.111	0.000	0.072	0.023
最大值	0.112	0.319	0.185	0.095	0.118
均值	0.105	0.309	0.104	0.085	0.105
标准差	0.003	0.014	0.033	0.005	0.007
$PSOS_d$					
最小值	0.000	0.000	0.000	0.000	0.000
最大值	0.085	0.132	0.146	0.048	0.110
均值	0.007	0.011	0.028	0.006	0.017
标准差	0.015	0.026	0.049	0.011	0.032

2.3 日间 APIN、PSOS 的估计结果

由于篇幅所限,这里仅列出中国联通的估计结果.图 1 是 NV_APIN-PSOS 模型估计的中国联通日间状态概率,横坐标是股票样本期的交易日.由图 1 可知,在整个样本期间,好消息发生概率和坏消息发生概率的波动均较大,且两者的均值都在 0.5 左右,而不发生消息概率较稳定,均值在 0.2 左右.估计的每日各种情形下的对称性订单流冲击概率(π_{CG} 、 π_{CN} 、 π_{CB}) 在样本期出现聚集性

且比较稀疏,在大约前 2/3 样本期,非 0 的概率值较多,而大约后 1/3 样本期,概率值大多为 0.

图 2 是 NV_APIN-PSOS 模型估计的中国联通日间 APIN 和 PSOS,横坐标是股票样本期的交易日.由图 2 可知,APIN 的值在整个样本期较稳定,大约在 0.1 左右,大约前 2/3 样本期的 PIN 值较后 1/3 样本期的值略大.PSOS 的值在整个样本期有较多为 0,且大约前 2/3 样本期,PSOS 的非 0 的值较多,而大约后 1/3 样本期,值大多为 0.

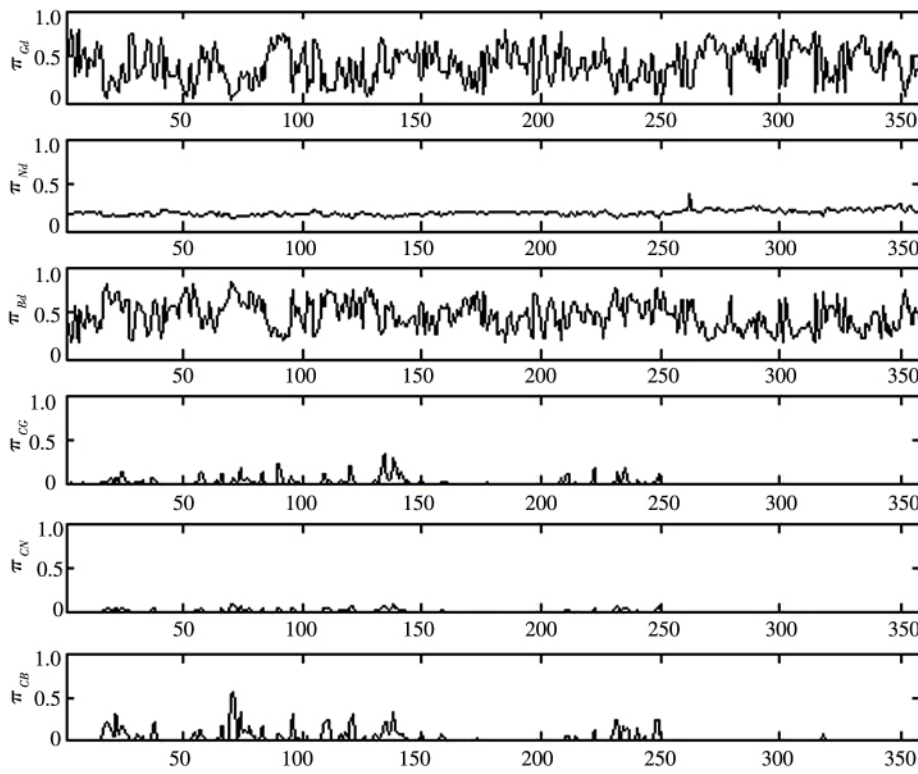


图 1 中国联通日间状态概率
Fig. 1 Daily probability of states of China Unicom

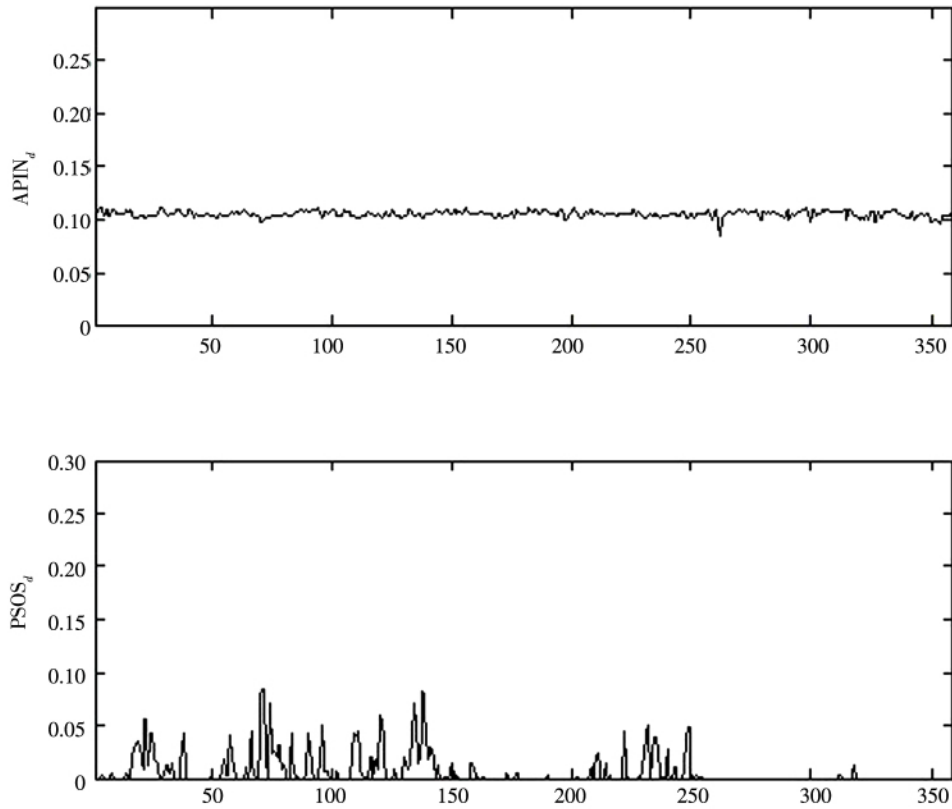


图 2 中国联通日间 APIN_d 及 PSOS_d
Fig. 2 Daily APIN_d and PSOS_d of China Unicom

2.4 模型的选择

DY_APIN-PSOS 模型(当 $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = 0$) 是 NV_APIN-PSOS 模型的嵌套模型,但此时似然比检验原假设的伽方统计量的极限分布的标准正则条件并不满足,因此,采用 AIC 准则 (Akaike's Information Criterion) 和 SBC (Schwartz's Bayesian Criterion) 准则选择模型. 其中

$$AIC = -2\hat{L} + 2k \text{ 和 } SBC = -2\hat{L} + \ln(D)k$$

式中 \hat{L} 极大似然估计的极大似然值, k 是模型的参数个数, D 是样本观测数. AIC 和 SBC 的值越小,则模型拟合越好. 表 3 中 DY_APIN-PSOS 模型和表 2 中 EKOP_PIN 模型对应的 AIC 和 SBC 值表明 5 只样本股票的 DY_APIN-PSOS 模型对应的 AIC 和 SBC 的值均较 EKOP_PIN 模型对应的 AIC 和 SBC 值小. 这说明 DY_APIN-PSOS 模型较 EKOP_PIN 模型对实际数据的拟合效果要好; 对比表 4 中 NV_APIN-PSOS 模型的 AIC 和 SBC 的值与表 3 中 DY_APIN-PSOS 模型的 AIC 和 SBC 值可知,紫金矿业的 NV_APIN-PSOS 模型的 AIC 和 SBC 的值则较 DY_APIN-PSOS 模型对应的 AIC 和 SBC 值大,这说明 DY_APIN-PSOS 模型比 NV_APIN-PSOS 模型对实际数据的拟合效果要好. 而表 4 中 NV_APIN-PSOS 模型参数估计的结果也发现紫金矿业的 γ_2, γ_4 参数并不显著. 除紫金矿业外,其他 4 只样本股票的 NV_APIN-PSOS 模型的 AIC 和 SBC 的值均较 DY_APIN-PSOS 模型对应的 AIC 和 SBC 值小,这说明对这些样本股票而言 NV_APIN-PSOS 模型较 DY_APIN-PSOS 模型对实际数据的拟合效果要好. 因此,总体而言, NV_APIN-PSOS 模型较 DY_APIN-PSOS 模型对实际数据的拟合效果要好.

2.5 稳健性检验

为了进一步检验 NV_APIN-PSOS 模型计算信息风险的有效性和解释能力,这里对 NV_APIN-PSOS 模型和 DY_APIN-PSOS 模型做进一步的比

较. 借鉴 Kalev 等^[20],李广川等^[22]比较两个信息风险模型的思路,建立一个面板回归模型,用于检验 NV_APIN-PSOS 模型和 DY_APIN-PSOS 模型对不同价差的解释能力,从而比较上述两个模型对不同价差的解释能力. 式(13)给出了面板回归模型.

$$Spread_{i,t} = \beta_0 + \beta_1 VAPIN1_{i,t} + \beta_2 VAPIN2_{i,t} + \beta_3 VOL_{i,t} + \eta_{i,t} \quad (13)$$

其中 $Spread$ 是价差变量,在实证分析中分别采用开盘后 10 min 平均价差,收盘前 10 min 平均价差和日均价差. $VAPIN1$ 是 DY_APIN-PSOS 模型估计出的日频知情交易概率(由于 DY_APIN-PSOS 模型并不能直接估计出日度知情交易概率,这里采用 DY_APIN-PSOS 模型估计出每个季度的知情交易概率,然后把估计出的知情交易概率赋值给该季度的每个交易日)与股价的乘积, $VAPIN2$ 是 NV_APIN-PSOS 模型估计出的日频知情交易概率与股价的乘积,股票价格为日收盘价, VOL 为日均交易额的对数值. 在进行面板回归时,采用 Husman 检验确定采用固定效应模型还是随机效应模型.

这里选取上证 50 样本股在 2009-01-01 至 2009-12-31 的高频交易数据作为样本,同样数据来源于国泰安高频分笔数据. 剔除了 5 只在样本期调出上证 50 的股票,并剔除了一只交易日数不足的股票,剩余 44 只股票作为分析样本. 剔除了连续竞价期外的所有成交数据.

根据已有研究,知情交易概率对价差有正向影响,而交易量对价差有负向影响. 因此,这里期望 $VAPIN1$ 和 $VAPIN2$ 有正的系数值,而 VOL 有负的系数值. 若 NV_APIN-PSOS 模型优于 DY_APIN-PSOS 模型,则期望 $VAPIN2$ 在回归方程中的系数估计值为正且显著,而 $VAPIN1$ 在回归方程中的系数估计值的显著性较低或不显著. 表 6 给出了式(13)的估计结果,这里仅列出重要系数的估计结果.

表 6 面板回归方程的估计结果

Table 6 Results of panel regression

回归变量	开盘后 10 min 平均价差		收盘前 10 min 平均价差		日均价差	
VAPIN1	-0.000 4	(-2.05)	-0.001 5	(-12.96)	-0.001 2	(-14.26)
VAPIN2	-0.000 6	(-1.16)	0.000 4	(7.9)	0.000 1	(3.03)
VOL	0.001 2	(1.35)	0.001 5	(2.92)	0.000 4	(0.8)

注:不同价差栏分别代表因变量为相应价差时的估计结果,任一价差栏包括两列,第 1 列为与行变量相对应的系数估计值,第 2 列括号内的数值为相应系数的 t 统计量或 z 统计量(固定效应面板回归是 t 统计量,随机效应面板回归是 z 统计量). 一般情形下,绝对值大于 2 意味着系数显著.

从表 6 的结果可知, $VAPIN1$ 变量的系数对所有价差的解释中均为负值, 且显著, 这与预期不符; $VAPIN2$ 变量的系数对开盘后 10 min 平均价差的解释为负值, 这与预期不符, 但这个系数并不显著, 而 $VAPIN2$ 变量的系数对收盘前 10 min 平均价差和日均价差的解释均为正值, 且显著, 这与预期的符号一致. VOL 变量的系数对收盘前 10 min 平均价差的解释中为正值, 且显著, 这与预期不符, VOL 变量的系数对开盘后 10 min 平均价差和日均价差中的解释仍为正, 但不显著. 以上结果表明, $NV_APIN-PSOS$ 模型估计的信息风险相对于 $DY_APIN-PSOS$ 模型估计的信息风险具有更好的有效性和解释力.

2.6 信息风险的周内特征

在这部分检验信息风险的周内特征. 已有关于价差的研究表明, 价差呈现明显的周内特征, 即周一的买卖报价价差较大, 一些研究者把这个特征解释为信息风险引起的^[21]. 由于逆向选择的存在, 买卖价差的周内特征反应出信息风险的特征. 周末不交易积累的周末信息, 可能导致周一的信息风险水平高于其它交易日. 那么是否信息风险也具有价差类似的周内模式呢? 图 3 是 $NV_APIN-PSOS$ 模型计算出的上证 50 在 2009-01-01

至 2009-12-31 的知情交易概率的周内模式, 也即信息风险经横截面和时间序列平均后得出的周内模式. 由图 3 可知, 周一的信息风险高于其它交易日的信息风险, 信息风险自周一开始逐渐降低, 直至周五, 信息风险略有上升.

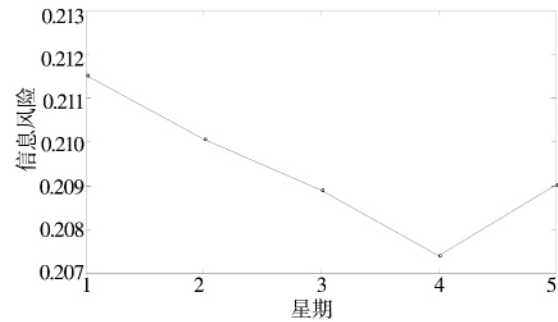


图 3 信息风险的周内模式

Fig. 3 Weekly pattern of the information risk

表 7 是信息风险周内特征的 t 检验, 由表 7 可知, 周一的信息风险的平均百分比变化最大, 其值为 1.13. 周二至周四的信息风险的平均百分比变化均小于 0, 而周五的信息风险的平均百分比变化为 0.97. 由于对这些变化序列进行 t 检验的概率值 P 均大于 0.1, 因此信息风险的周内特征并不显著, 即信息风险没有显著的周内模式.

表 7 信息风险周内特征的 t 检验

Table 7 T test of the weekly pattern of the information risk

	周一	周二	周三	周四	周五
信息风险平均变化 (%)	1.13	-0.39	-0.38	-0.22	0.97
t 检验的概率值 P	0.16	0.61	0.68	0.79	0.31

注: 信息风险百分比变化指的是当日信息风险与前一个交易日的信息风险的差与当日信息风险的百分比值, 周一的信息风险百分比变化则是指的周一的信息风险与周五的信息风险的差与周一的信息风险的百分比值.

3 结束语

本文扩展了 $DY_APIN-PSOS$ 模型, 采用已有金融文献使用交易量建模消息状态概率和对称性订单流冲击概率的方法, 允许上述两种概率是时变的, 构建了 $NV_APIN-PSOS$ 模型, 这放松了 $DY_APIN-PSOS$ 模型中这些状态变量为常数的假设. 实证研究结果表明: 本文提出的 $NV_APIN-PSOS$

模型较 $DY_APIN-PSOS$ 模型更好地拟合了我国股票市场的数据以及对价差具有更好的解释力. 本文构建的模型在允许上述状态时变的基础上, 提出了一种估计日间和日内两个时间窗口的 $APIN$ 和 $PSOS$ 方法, 可以应用于多个方面, 如计算事件(如盈余公告和公司兼并重组)前后的知情交易概率水平; 研究非流动性和非对称信息在资产定价中的作用, 这些都是后续研究的一部分.

参考文献:

- [1] Easley D, Kiefer N M, O'Hara M. Cream-skimming or profit-sharing? The curious role of purchased order flow [J]. *The Journal of Finance*, 1996, 51(3): 811–833.
- [2] Easley D, Kiefer N M, O'Hara M. The information content of the trading process [J]. *Journal of Empirical Finance*, 1997, 4(2): 159–186.
- [3] Easley D, Hvidkjaer S, O'Hara M. Is information risk a determinant of asset returns? [J]. *The Journal of Finance*, 2002, 57(5): 2185–2221.
- [4] Easley D, O'Hara M, Paperman J. Financial analysts and information-based trade [J]. *Journal of Financial Markets*, 1998, 1(2): 175–201.
- [5] Brockman P, Chung D Y. Informed and uninformed trading in an electronic order-driven environment [J]. *Financial Review*, 2005, 35(2): 125–146.
- [6] Easley D, O'Hara M, Saar G. How stock splits affect trading: A microstructure approach [J]. *Journal of Financial and Quantitative Analysis*, 2001, 36(1): 25–52.
- [7] Easley D, O'Hara M. Information and the cost of capital [J]. *The Journal of Finance*, 2005, 59(4): 1553–1583.
- [8] Vega C. Stock price reaction to public and private information [J]. *Journal of Financial Economics*, 2006, 82(1): 103–133.
- [9] Boehmer E, Grammig J, Theissen E. Estimating the probability of informed trading: Does trade misclassification matter? [J]. *Journal of Financial Markets*, 2007, 10(1): 26–47.
- [10] 才静涵, 夏乐. 卖空制度、流动性与信息不对称问题研究——香港市场的个案 [J]. *管理科学学报*, 2011, 14(02): 71–85.
- Cai Jinghan, Xia Le. Short-sale constrains liquidity and information asymmetry: Evidence from Hong Kong stock market [J]. *Journal of Management Sciences in China*, 2011, 14(02): 71–85. (in Chinese)
- [11] 屈文洲, 谢雅璐, 叶玉妹. 信息不对称、融资约束与投资—现金流敏感性——基于市场微观结构理论的实证研究 [J]. *经济研究*, 2011, 6: 105–117.
- Qu Wenzhou, Xie Yalu, Ye Yumei. Information asymmetry and investment-cash flow sensitivity: An empirical research based on market microstructure theory [J]. *Economic Research Journal*, 2011, 6: 105–117. (in Chinese)
- [12] Aktas N, De Bodt E, Declerck F, et al. The PIN anomaly around M&A announcements [J]. *Journal of Financial Markets*, 2007, 10(2): 169–191.
- [13] Yan Y, Zhang S. An improved estimation method and empirical properties of the probability of informed trading [J]. *Journal of Banking & Finance*, 2012, 36(2): 454–467.
- [14] Easley D, Engle R F, O'Hara M, et al. Time-varying arrival rates of informed and uninformed trades [J]. *Journal of Financial Econometrics*, 2008, 6(2): 171–207.
- [15] Tay A, Ting C, Tse Y K, et al. Using high-frequency transaction data to estimate the probability of informed trading [J]. *Journal of Financial Econometrics*, 2009, 7(3): 288–311.
- [16] Duarte J, Young L. Why is PIN priced? [J]. *Journal of Financial Economics*, 2009, 91(2): 119–138.
- [17] Preve D, Tse Y K. Estimation of time-varying adjusted probability of informed trading and probability of symmetric order-flow shock [J]. *Journal of Applied Econometrics*, 2013, 28: 1138–1152.
- [18] Berry T D, Howe K M. Public information arrival [J]. *The Journal of Finance*, 1994, 49(4): 1331–1346.
- [19] Mitchell M L, Mulherin J H. The impact of public information on the stock market [J]. *The Journal of Finance*, 1994, 49(3): 923–950.
- [20] Kalev P S, Liu W M, Pham P K, et al. Public information arrival and volatility of intraday stock returns [J]. *Journal of Banking & Finance*, 2004, 28(6): 1441–1467.
- [21] Easley D, Hvidkjaer S, O'Hara M. Factoring information into returns [J]. *Journal of Financial and Quantitative Analysis*,

2010, 45(2): 293.

[22] 李广川, 刘善存, 邱苑华. 连续指令驱动市场的信息交易概率: 一种新的方法[J]. 管理科学学报, 2010, 13(10): 8-20.

Li Guangchuan, Liu Shancun, Qiu Wanhua. The probability of informed trading in continuous order driven market: A new method[J]. Journal of Management Sciences in China, 2010, 13(10): 8-20. (in Chinese)

Measure of time-varying information risk with symmetric order-flow shock

XIONG Chun-lian^{1 2}, *WANG Chun-feng*¹, *FANG Zhen-ming*³

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

2. Tianjin Chengjian University, Department of Basic Science, Tianjin 300384, China;

3. Bohai Securities Co., Ltd, Tianjin 300381, China

Abstract: Accurate measure of the risk of stock information is of great significance for asset pricing, risk management, and the measure of market performance. To overcome the shortcoming of assuming constant probabilities of the states of news and symmetric order-flow shock of the model of Duarte and Young, we extend the methodology of Duarte and Young. We model the probabilities of the states of news and the probabilities of symmetric Order-Flow shock by using trade volumes. Our method allows both the probabilities of the states of news and the probabilities of symmetric Order-Flow shock to vary. Our APIN and PSOS estimates can be computed daily as well as over intraday intervals. Then, we select some actively traded stocks to do an empirical research and our model are compared with the model of Duarte and Young. Our empirical results show that our model provides a better fit for our data and has better explanatory power for spread. At last, we do an empirical research for the inter-day pattern of information risk.

Key words: probability of symmetric order-flow shock; time-varying risk of stock information; probability of informed trading; time-varying states of news