

doi:10.3969/j.issn.1672-626x.2014.06.008

# 如何衡量市场中的知情交易:理论与文献综述

尹康<sup>1,2</sup>,程志芬<sup>3</sup>

(1. 湖北经济学院 经济学系,湖北 武汉 430205; 2.上海财经大学 统计与管理学院,上海 200433;  
3. 武汉大学 珞珈学院,湖北 武汉 430064)

**摘要:**知情交易在市场微观结构的信息模型中扮演着关键角色,特别是在理性预期模型(REE)引入到市场微观结构以后,知情交易更是理解市场流动性、市场波动性以及价格发现等诸多市场运行特征的钥匙。本文对知情交易的基本特征和意义进行了介绍,对学术界衡量市场知情交易的各类方法进行了梳理总结,并对各方法特点作了简要评价。

**关键词:**知情交易;流动性交易;知情交易概率

**中图分类号:**F830.9 **文献标识码:**A **文章编号:**1672-626X(2014)06-0053-07

## 一、引言

信息在证券市场中所扮演的角色一直是市场参与各方所共同关心的话题。从传统意义上讲,金融市场是为经济社会提供投融资的场所,但在市场实际运行过程中,场外的一点风吹草动都会让市场波澜迭起。市场参与各方对交易标的资产的信息掌握是不尽相同的,或有管理层内部渠道,或基于市场传言,或立足于投资者自身的技术分析结论,不管是何种途径,市场上总有一部分投资者掌握有关资产价格变动更多的信息,而另外一类投资者(即所谓的散户)只能基于市场上的公开信息进行投资决策,显然在这两类投资者中存在信息不对称。

信息不对称会给市场带来许多不良影响,最为人们熟悉的便是逆向选择问题。Akerlof<sup>[1]</sup>在对二手车市场的分析中首次提出,市场交易者内部的信息不对称可能会导致市场交易无法有效进行。Bagehot<sup>[2]</sup>首次把证券市场的交易者分为知情交易者和非知情交易者,并从信息不对称的角度考虑了信息成本对价差的影响。当然,对于证券市场而言,信息不对称还涉及市场运行是否公正的问题。因此,合理地度量市场中信息不对称程度,无论是对投资者的理性决策还是对监管部门的有效监管,都是有益的。

在整个市场运行过程中,知情交易就是联接场内外信息的桥梁。<sup>[3]</sup>知情交易者通过自身的交易行为向市场内的其他投资者传递其独占的私有信息,这一过程充满着矛盾。从知情交易者自身角度来讲,其没有向市场内其他投资者传递信息的利他动机,不仅如此,他在交易过程中还要极力隐藏自己的身份,防止信息泄露,其目的就在于利用私有信息获取更大的交易利润。但是市场上的其他投资者会有一个动态学习的过程,相比噪音交易者,知情交易者的交易单往往具有明显的方向性和规律性,特别是内幕消息通常会有一个存续期限,在私有信息即将曝光成为市场的共有信息之前,知情交易者存在交易的紧迫感,他不得不在交易利润和隐藏信息这两者之间做出取舍。

正如 Kyle<sup>[4]</sup>在连续拍卖交易模型中所发现的那样,知情交易者通常在信息产生的初期,交易比较谨慎,这一时期,其交易模式和其他非知情交易者的交易模式没有太大区别,因而难以识别。而在信息存续周期的中期,知情交易者的交易行为将变得较有攻击性,其交易模式和其他非知情交易者相比会表现出明显的差异。随着市场传言逐渐增多,市场预期逐渐形成共识,证券的交易价格也逐渐逼近

收稿日期:2014-8-20

作者简介:尹康(1981-),男,湖北洪湖人,湖北经济学院讲师,上海财经大学博士研究生,主要从事金融计量学与市场微观结构研究。

其真实价值,此时知情交易者继续入场交易所获利润会逐渐下降,因此其交易意愿下降,在市场中也难以有效识别。

要把知情交易从众多流动性交易订单中分离出来并不是一件容易的事情。理论上,我们可以从机构或内幕交易者的订单与个人投资者的订单对市场的冲击来进行识别,但在实际中却缺乏这方面的数据,因此,我们必须依赖一些反映知情交易强度的代理变量来进行度量。这些度量的结果反映的是市场在一段时间内知情交易的活跃程度,或者在一定数量的订单流中,由知情交易者发起的订单所占的比例。大体上,知情交易概率的测度经历了两个阶段:间接测度阶段和直接测度阶段。在市场微观结构的早期文献中,知情交易通常是信息不对称的同义词,为了将信息不对称的程度以量化形式表现出来,研究者使用了一些指标,或称之为代理变量。其中间接测量的指标有三类,分别是价差、逆向成本 HS 以及未预期的价格冲击 HFV; 直接测量的指标统一记为 PIN, 也被称为知情交易概率。这些度量方法一般都是从某种角度提炼价格变化方向、订单流的规模以及订单不平衡所释放的信息。

## 二、知情交易的间接度量方法

### (一) 买卖询价差

在早期的文献中,价差只被视为做市商存货以及交易成本的市场体现。<sup>[5]</sup> 但 Copeland 和 Galai<sup>[6]</sup>、Glosten 和 Harris<sup>[7]</sup> 以及 Amihud 和 Mendelson<sup>[8]</sup> 等人的研究表明,价差同时反映了做市商对不知情交易者的盈余和对知情交易者的损失之间的一个平衡。这也说明,除了存货成本和交易成本外,价差中还包含有市场交易过程中的信息成本。鉴于交易成本由市场制度外生给定,一般无太大变化,存货成本则是均值为常数的随机变量,因此,价差的系统性变化主要来自于信息成本。从这一角度看,价差可被视为市场知情交易或信息不对称的一个函数。在这些模型中,金融中间商所面临的是一个标准的“柠檬市场”。鉴于知情交易者和流动性交易者不容易区分,中间商在买价和卖价之间打入一个楔子,这个价差使得代理商或特约交易商能从流动性交易者那里获取利润,用以补偿与知情交易者交易所承担的损失,这也变相给知情交易者提供了流动性。因此,价差的大小也可以作为市场知情交易比例的一个代理变量。

### (二) 价差中的逆向选择成本

价差的大小不仅包含了信息不对称的因素,同时也包括了做市佣金以及订单处置的成本,为了测量价差中纯粹由信息成本所产生的那一部分,我们利用 Huang 和 Stoll<sup>[9]</sup> 所提出的模型估计价差中的信息成本部分,该模型实际上是 Glosten 和 Harris<sup>[7]</sup> 所提出的交易指示模型的一般化版本。Huang 和 Stoll<sup>[9]</sup> 推导了一个相对简单的模型,借助该模型可以对价差中信息成分做一步分解(One-step Decomposition)。定义一个示性变量  $Q_t$ , 用来反映最近一次交易价格变动的方向,取值为  $\{-1, 0, 1\}$ ,  $P_t$  为交易价格,  $M_t$  为中间价,因此:

$$Q_t = \begin{cases} -1, & P_t < M_t; \\ 0, & P_t = M_t; \\ 1, & P_t > M_t; \end{cases}$$

$S$  为询价差,  $\varepsilon_t$  反映随机的公共信息冲击,对应的回归模型为:

$$\Delta P_t = \frac{S}{2} \Delta Q_t + \alpha \frac{S}{2} Q_{t-1} + \varepsilon_t \quad (1)$$

$\alpha$  测度了  $\frac{S}{2}$  中信息成本的比例,余下的  $(1-\alpha)$  为订单处置成本和做市佣金。以往有些文献认为  $\alpha$  中应包含存货成本,而后续的研究表明,在该模型的设定下,  $\alpha$  中存货成本的比例接近于零。如果不是以比例而是以实际货币单位来度量逆向选择成本,那么一个常用的表示是  $\alpha S$ 。

### (三) HFV

该测量方法来源于 Foster 和 Viswanathan<sup>[10]</sup> 以及 Hasbrouck<sup>[11]</sup> 两篇论文中所提出的模型,鉴于订单的提交经常表现出序列相关性,因此,部分订单流是可预期的。Hasbrouck<sup>[11]</sup> 的模型则把未预期到的订单流作为解释变量来说明信息性交易的变化。令  $V_t$  为  $t$  时刻成交量,  $\Delta P_t$  为价格变化,  $Q_t$  为示性变量,定义如前所述,建立如下模型:

$$V_t = \theta_0 + \sum_{i=1}^I \beta_i \Delta P_{t-i} + \sum_{i=1}^I \gamma_i V_{t-i} + \tau_i \quad (2)$$

$$\Delta P_t = \phi_0 + \phi_1 \Delta Q_t + HFV \times \tau_t + \varepsilon_t \quad (3)$$

方程(3)中的  $\tau_t$ , 规范讲,应为第一个方程中的估计残差, HFV 作为其系数,表示未预期的订单流对价格的冲击,其含义与 HS 中  $\alpha$  类似。

## 三、知情交易的直接度量方法

### (一) EKOP 模型

Easley 等<sup>[12]</sup> 在 Glosten 和 Milgrom<sup>[13]</sup> 序贯交易模

型的基础上,对事件树的结构做了部分修改,提出了著名的 EKOP 模型。虽然这一模型提出的初衷仍是为了分析价差的变化特征,但在该模型框架下所给出的知情交易概率的计算公式却是这篇论文最吸引后续研究者之处。正是从这篇论文开始,在市场微观结构的文献中出现了一系列的相关研究,均是考虑对知情交易概率的直接测算。EKOP 模型的基本设定如下:

1. 市场由三类参与者组成:流动性交易者(非知情交易者)、知情交易者以及风险中性的做市商。

2. 信息事件如果有,则在交易初期发生。消息树有三个分叉:无消息、好消息和坏消息。无消息发生时,市场预期没有任何改变;好消息发生,证券价值变为  $\bar{V}$ ;坏消息发生,证券价值变为  $\underline{V}$ ,证券期初价值为  $V_0$ ,显然有  $\bar{V} > V_0 > \underline{V}$ 。

3. 流动性交易者随机交易,且进场后必须交易,或者买入,或者卖出。知情交易者则视消息情况而定,若无消息,则不交易;如果是好消息,则买入;如果是坏消息,则卖出。

4. 交易过程的统计特征。假定在交易日前,有信息事件发生的概率为  $\alpha$ ,则无信息事件的概率为  $1-\alpha$ 。在给定信息事件发生的条件下,发生坏消息的概率为  $\delta$ ,发生好消息的概率为  $1-\delta$ 。知情交易者无论买入或是卖出,其提交的订单流均服从强度为  $\mu$  的泊松分布。非知情交易者的买单和卖单均服从强度为  $\varepsilon$  的泊松分布。

在上述设定下,知情交易的概率被定义为:

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} \quad (4)$$

接下来的工作便是利用经验数据,估计其中参数  $\alpha$ 、 $\mu$  以及  $\delta$ 。假定在一个时长为  $T$  的交易时段内有  $B$  次买入和  $S$  次卖出,可以给出其似然函数

$$L(B, S | \alpha, \delta, \mu, \varepsilon) = (1-\alpha)f_1(B, S) + \alpha\delta f_2(B, S) + \alpha(1-\delta)f_3(B, S) \quad (5)$$

其中

$$f_1(B, S) = e^{-\varepsilon T} \frac{(\varepsilon T)^B}{B!} e^{-\varepsilon T} \frac{(\varepsilon T)^S}{S!} \quad (6)$$

$$f_2(B, S) = e^{-\varepsilon T} \frac{(\varepsilon T)^B}{B!} e^{-(\varepsilon+\mu)T} \frac{[(\varepsilon+\mu)T]^S}{S!} \quad (7)$$

$$f_3(B, S) = e^{-(\varepsilon+\mu)T} \frac{[(\varepsilon+\mu)T]^B}{B!} e^{-\varepsilon T} \frac{(\varepsilon T)^S}{S!} \quad (8)$$

如果有  $n$  个相互独立的交易时段,则其似然函数为

$$L(\alpha, \delta, \mu, \varepsilon | B, S) = \prod_{i=1}^n L(B_i, S_i) \quad (9)$$

通过对式(9)进行最大化,给出各个参数的极大似然估计,<sup>①</sup>并根据估计结果求解出 PIN。

## (二) Nyholm 模型

Glosten 和 Milgrom<sup>[13]</sup>所提出的模型也被称为交易—指示 (Trade-indicator) 模型,该模型最早由 Glosten<sup>[14]</sup>、Glosten 和 Harris<sup>[7]</sup>等人提出,它假定做市商以一种成本平滑的方式进行运作,通过不断调整价差来平衡交易过程中的收益与损失。交易—指示模型也可视作一种混合分布模型,正如 Easley 和 O'hara<sup>[15]</sup>、Easley 等<sup>[12]</sup>所设定的那样,由“自然”决定交易日的事件发生类型(好消息、坏消息和无消息),各类事件的发生概率不同,对应的交易结构也不尽相同。

和所有分析市场信息不对称的模型类似,交易—指示模型也假定市场上的交易者分为两类:知情交易者和非知情交易者。在其他研究知情交易的文献中,通常假定做市商无论在交易发生前或是发生后都无法区分订单的提交者属于哪一类别。但在 Nyholm 模型中放松了这一假定,允许做市商在每次交易发生后,能识别交易者分属哪一类。在模型中,非知情交易者出于外生动机(流动性需求)进行交易,他们掌握的信息是市场公开信息,记为  $\Omega_t^i$ ;类似的,记  $\Omega_t^i$  为知情交易者的信息集,他们或通过专业的分析手段,或是抢先获得新的内幕消息,使得其信息集优于市场的公开信息。一般有  $\Omega_t^i = \{\Omega_t^i, G_t\}$ ,  $G_t$  表示额外信息集。对于做市商而言,在其制定新的市场买卖价之前,其信息集  $\Omega_{t+}^i = \{\Omega_t^i, Y_{t+}\}$ ,  $Y_{t+}$  表示做市商所知道的关于订单流的内部信息,包括买卖方向和价位等信息,一旦新的买卖价格公布以后,该信息集就变成了  $\Omega_{t+}^i$ 。显然  $G_t \cup Y_{t+}$ , 但是做市商通常会利用  $Y_{t+}$  去推断  $G_t$ , 以减少当其制定的价格偏离股票真实价值所遭受的损失。

记  $\tilde{V}_t$  为股票的基础价值,  $M_t$  为中间价,  $S_t$  为价差,知情交易者仅在下列情形之一发生时才进行交易:

$$\tilde{V}_t > M_t + \frac{S_t}{2} \text{ 或 } \tilde{V}_t < M_t - \frac{S_t}{2}$$

因此做市商面临的潜在损失为:

$$\lambda_t = (M_t + \frac{S_t}{2} - \tilde{V}_t) P\{\tilde{V}_t > M_t + \frac{S_t}{2} | \Omega_t^i, \Omega_{t+}^i\} - (M_t - \frac{S_t}{2} - \tilde{V}_t) P\{\tilde{V}_t < M_t - \frac{S_t}{2} | \Omega_t^i, \Omega_{t+}^i\} \quad (10)$$

(10)式中条件概率也即知情交易者进行交易的概率。价差的大小用于平衡与知情交易者交易的损失以及与非知情交易者的收益,做市商要能做到盈亏平衡,其制定的买卖价差应遵循“无悔”原则。做市商清楚知情交易者的动机,并且需要最小化预期损失,所以必须尽快更新其中间价,相应决策准则如下:

$$\Delta M_{t+1} = \begin{cases} > \frac{S_t}{2}, P\{\tilde{V}_t > M_t + \frac{S_t}{2} | \Omega_t^1, \Omega_{t+1}^s\} > 0.5 \\ < -\frac{S_t}{2}, P\{\tilde{V}_t < M_t - \frac{S_t}{2} | \Omega_t^1, \Omega_{t+1}^s\} > 0.5 \end{cases} \quad (11)$$

虽然中间价的调整幅度还有待实证,但是面临知情交易者所发起的交易,其调整方向必须满足上式中的原则。对非知情交易者发起的交易,其调整原则如下:

$$\Delta M_{t+1} = \begin{cases} < \frac{S_t}{2}, P\{\tilde{V}_t > M_t + \frac{S_t}{2} | \Omega_t^1, \Omega_{t+1}^s\} < 0.5 \\ > -\frac{S_t}{2}, P\{\tilde{V}_t < M_t - \frac{S_t}{2} | \Omega_t^1, \Omega_{t+1}^s\} < 0.5 \end{cases} \quad (12)$$

根据上述价格调整机制, Nyholm 引入了一个便于估计的价格调整模型:

$$\Delta M_t = (\alpha_0 + \alpha_1 I_{t-1}) \frac{S_t - 1}{2} Q_{t-1} + \varepsilon_t \quad (13)$$

对建模者而言,  $I_t$  为不可观测变量, 但做市商知道其取值, 当  $I_{t-1}=0$  时, 表示在  $t-1$  时刻发生的交易为非知情交易,  $I_{t-1}=1$  时, 交易为知情交易。  $S$  表示价差,  $Q$  表示交易指示变量,  $Q_t=1$  时, 交易在买方叫价 (Ask Price) 执行,  $Q_t=-1$ , 交易在买方出价 (Bid price) 执行;  $\varepsilon_t$  表示公共信息对价格的冲击。  $\alpha_0$  反映非知情交易对价格的冲击,  $\alpha_1$  反映知情交易对价格的冲击。从前面设定可知, 非知情交易对价格冲击幅度一般小于  $S/2$ , 故  $\alpha_0 < 1$ ; 而知情交易只在基础价值与买卖价格偏离较大时才进场交易, 故对价格的冲击一般大于  $S/2$ , 所以  $\alpha_1 > 1$ ,  $\alpha_0$  和  $\alpha_1$  的经济含义则表示交易次数对价格更新的影响。

在参数的估计过程中,  $I_t$  是不可观测的, 因此借用制度转换 (Regime-switching) 的思路, 考虑  $I_t$  动态调整过程为一阶马尔科夫链, 其状态转移矩阵记为

$$P = \begin{pmatrix} P_{00} & 1 - P_{11} \\ 1 - P_{00} & P_{11} \end{pmatrix} \quad (14)$$

记  $\theta = (\alpha_0, \alpha_1, \delta, P_{00}, P_{11})$ ,  $X_t = \{S_t, \dots, S_t; Q_t, \dots, Q_t; M_t, \dots, M_t\}$ ,  $\varepsilon_t \sim N(0, \sigma^2)$ , 定义

$$\eta_t = \begin{pmatrix} f(\Delta M_t | I_{t-1}=0, X_{t-1}; \theta) \\ f(\Delta M_t | I_{t-1}=1, X_{t-1}; \theta) \end{pmatrix} \quad (15)$$

$$\hat{\xi}_{dt-1} = \begin{pmatrix} P\{I_{t-1}=0 | X_{t-1}, I_{t-1}; \theta\} \\ P\{I_{t-1}=1 | X_{t-1}, I_{t-1}; \theta\} \end{pmatrix} \quad (16)$$

则模型估计的对数似然函数为  $L(\theta) = \sum_{t=1}^T \log L(\theta)_t$ , 其中  $\xi_{dt-1} = P \cdot \hat{\xi}_{dt-1}$ ,  $L(\theta) = \eta_t \cdot \hat{\xi}_{dt-1}$ 。根据极大似然的估计结果, 可以给出  $I_t$  的平滑估计  $\hat{I}_t$ , 因此对应的知情交易为

$$PIN_T = \frac{\sum_{t=1}^T \hat{I}_t}{T} \quad (17)$$

### (三) VPIN 模型

Easley 等<sup>[16]</sup>基于订单的不平衡特征, 提出了一种新的知情交易概率的估计方法, 即交易量同步的知情交易概率, 简称 VPIN (Volume-Synchronized Probability of Informed Trading)。这种新的估计方法通常也被视作利用高频数据的 PIN 估计 VPIN 的估计不需要借助极大似然的方法估计未知参数 ( $\alpha, \delta, \varepsilon, \mu$ ), 因此简化了知情交易概率的求解过程。此外, 即使在一个交易日内, VPIN 也可以根据新信息抵达市场的情况不断更新其取值。

在市场交易结构的设定上, VPIN 方法基本沿用 EKOP 模型的设定。这两种方法最关键的差别在于对交易过程的刻度处理上, 和传统的金融时序模型类似, EKOP 模型对交易过程的记录是以自然时间为刻度, 但在 VPIN 的计算过程中, 则采用了一种新的等交易量刻度 (Volume Clock), 这种记录方法在 Easley 等<sup>[17]</sup>考察交易时间间隔所包含信息量时曾使用过。具体来说, 选取某一交易量  $V$  作为分隔标准, 在任意交易日内, 从交易开始, 当交易量达到  $V$  时, 记该刻度  $\tau=1$ , 依次  $\tau=2, \dots, n$ 。在任一刻度内, 交易量均为  $V$ , 这也被称为一个“交易篮子” (Volume Bucket)。在一些交易比较活跃的日子,  $n$  的取值会比较大; 反之则较小。在实际的标记过程中,  $V$  的取值并无统一标准,<sup>②</sup>具体视标的资产的交易规模而定。

接下来需要解决的问题是如何区分“交易篮子”中买单和卖单的数量。如果能获得逐笔交易数据 (Tick Data), 那么对交易单方向的划分将非常简单明了。对于分笔数据, 公开平台所获得的数据已无法准确表示交易单的发起方。Lee 和 Ready<sup>[18]</sup>曾以 Level-2 的数据为基础, 提出过一种辨识交易方向

的算法,但在实际研究过程中,Level-2 的数据并不容易获得,而且数据规模庞大不易处理。Easley 等<sup>[19]</sup>所提出的 BVC(Bulk Volume Classification)算法则具有更好的可操作性,该方法不仅适用于分笔数据,也适用于每分钟记录的高频数据。对于某一“交易篮子” $\tau$ ,令

$$V_{\tau}^B = \sum_{i=(\tau-1)+1}^{i(\tau)} V_i \cdot Z\left(\frac{P_i - P_{i-1}}{\delta_{\Delta P}}\right) \quad (18)$$

$$V_{\tau}^S = \sum_{i=(\tau-1)+1}^{i(\tau)} V_i \cdot [1 - Z\left(\frac{P_i - P_{i-1}}{\delta_{\Delta P}}\right)] = V - V_{\tau}^B \quad (19)$$

$t(\tau)$ 是第 $\tau$ 个“交易篮子”最后一笔交易发生的时刻, $Z(\cdot)$ 是标准正态分布的累积分布函数, $\delta_{\Delta P}$ 是基于时间轴的价格变动标准差。从公式中可以看出,如果某一个“篮子”中价格从始至终都没有变化,则该“篮子”中买卖单相等;如果价格在上升,则买单比例大于卖单;如果价格在下降,则卖单比例大于买单。当然如同 Lee 和 Ready<sup>[18]</sup>算法类似,该分类算法也不可能完全准确区分每一笔交易的方向。

最后回到要求解的知情交易概率 VPIN。由 EKOP 模型的结果可知, $PIN = \alpha\mu / (\alpha\mu + 2\varepsilon)$ ,根据 Easley 等<sup>[17]</sup>的研究结论可知:

$$E(V_{\tau}^B + V_{\tau}^S) = \alpha\mu + 2\varepsilon \quad (20)$$

$$E(V_{\tau}^B - V_{\tau}^S) \approx \alpha\mu \quad (21)$$

因此,VPIN 的一个近似计算结果为:

$$VPIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} \approx \frac{E(V_{\tau}^B - V_{\tau}^S)}{E(V_{\tau}^B + V_{\tau}^S)} = \frac{\sum_{\tau=1}^n |V_{\tau}^B - V_{\tau}^S|}{nV} \quad (22)$$

#### 四、简要评述

从发展的角度看,对市场中知情交易的度量方法可以分为两类:间接度量和直接度量。在早期的市场微观结构文献中,知情交易的度量主要采用间接手段,而其中的媒介便是价差。价差常被用于度量不同的市场特征,最早用于反映做市商的存货状况,后来也被用于描述市场的流动性,Stoll<sup>[20]</sup>建议把价差分解成三个部分:订单处置成本(Order Processing Cost)、存货成本(Inventory Holding Cost)以及逆向选择成本(Adverse Selection Cost)。前两者也被一起称为交易成本,而逆向选择成本则是做市商为市场提供流动性时,用于补偿与知情交易者进行交易所发生的损失。因此当知情交易越活跃时,逆向成本越高,价差也越大。显然,如果直接用价差来度量知情交易的程度,其结果势必受交易成本的影响。后续的 HS 模型以及 HFV 模型本质上都是对价差进行

分解,把价差中反映知情交易的那部分逆向成本分离出来,只不过两者所用的模型不同。HS 模型是建立在交易—指示模型基础之上的,注重交易行为的初始动机;HFV 模型则借鉴了时间序列中的向量自回归(VAR)模型,更多强调交易动机的表现。总的说来,基于市场买卖价差所构建的直接度量方法,缺乏对交易过程的结构化描述,无法刻画知情交易的动态变化特征,价差分解的准确性也令人怀疑,结果难有说服力。因此,用逆向成本来反映知情交易,本质上只是找到了一个近似的替代变量,而不是对问题本身的直接度量。

考虑到基于价差的间接度量方法的一些弊端,在过去十多年中,理论界逐步发展出几类直接对市场知情交易程度进行度量的方法,而其中最具代表性,亦是研究者使用最多的是 EKOP 模型所提出的知情交易概率(PIN)。该指标的构建基于序贯交易框架,有较为严谨的理论基础,其指标内涵明确,结果易于解释,计算上也无太大障碍。因此该方法一经提出,便受到众多市场微观结构领域的研究者推崇。有不少学者利用 EKOP 模型的方法对欧美证券市场的知情交易进行了度量,并分析了对应的市场结构特征。此外,也有一些国内学者利用 EKOP 模型所介绍的方法对中国证券市场的知情交易概率进行了估算,并分析了与知情交易有关的一些市场微观结构特征。杨之曙和姚松瑶<sup>[21]</sup>以 EKOP 模型为基础,估算上海证券市场的知情交易概率。他们选取上交所 704 只股票,交易记录从 2002 年 7 月 1 日至 2003 年 6 月 30 日。在对样本股按日平均成交量大小分组之后,计算了第一、三、五组的平均 PIN,其结果分别为 0.1227、0.1532、0.1636。这一结论也同时验证了成交越活跃的股票,其知情交易的概率越低。其他类似文献还可参见王春峰等<sup>[22]</sup>、周开国等<sup>[23]</sup>、韩立岩等<sup>[24]</sup>。也有国内学者尝试对 EKOP 模型进行改进,马丹等<sup>[25]</sup>放弃了 EKOP 模型中交易者的订单提交无相关性这一假定,考虑知情交易者和非知情交易者的订单提交都存在自相关性,并设定信息事件的发生服从一阶遍历的马尔科夫过程。经过这些改进后重新估计了 PIN,模拟结果验证了改进后的估计结果更接近真实情况。

对其他几类知情交易的直接度量方法,研究的人相对较少,影响力也不及 EKOP 模型。至于 Nyholm 的交易指示模型所提出的知情交易度量方

法,在国内有攀登和施东晖<sup>[26]</sup>曾借鉴这一模型的基本思想,并考虑了 Kyle<sup>[4]</sup>所提出的交易者的价格反应函数,得出一种新的 PIN 估计方法。最后是 Easley 等<sup>[19]</sup>新近提出的 VPIN 指标,这一指标可以看做是 EKOP 模型中 PIN 指标的变形版,但是在计算上更有优势。VPIN 指标其实在 Easley 等人更早时间的工作论文中就已经出现,但是这个方法出现没多久就遭到其他学者 Andersen 和 Bondarenko<sup>[27]</sup>的猛烈批判,Andersen 等人对 VPIN 方法中按“交易篮子”标记高频率序列而不是按传统时间刻度的做法以及“交易篮子”中买单和卖单的区别方法表示了质疑,从而对这一指标的有效性表示怀疑。尽管之后 Easley 等<sup>[28]</sup>专门对这些质疑做了回应,但目前关于 VPIN 方法的讨论仍在持续。

对于市场中知情交易(或信息不对称)的度量,理论界虽然提出不少的度量方法,但由于研究者不可能获得市场上交易账户数据,因此对其度量结果的有效性应持相对谨慎的态度。同时,知情交易尽管在理论上可能有一个明确的定义,但是从实际监管过程中去判定某笔交易为知情交易并不是一件容易的事情。诸多现实性障碍使得知情交易概率这一指标的应用受到极大限制,目前也尚未见到有交易所把知情交易概率做为市场监管的工具。关于这一指标的后续研究和推广还有许多工作要做。

#### 注释:

- ① 在最大化过程中,需要注意参数取值范围,Easley 的建议是对  $\alpha, \beta$  做 Logit 变换,对  $\mu, \varepsilon$  做对数变换。
- ② Easley 等在分析 S&P500 期货合约时,建议  $V$  的取值为日平均交易量的 1/50。

#### 参考文献:

- [1] Akerlof, G. A. The Market for "Lemons": Quality Uncertainty and the Market Mechanism[J].The Quarterly Journal of Economics, 1970, (3): 88-500.
- [2] Bagehot, W. The Only Game in Town [J].Financial Analysts Journal, 1971, (2): 12-14.
- [3] O'hara, M. Market Microstructure Theory[M].MA: Blackwell Cambridge, 1995.
- [4] Kyle, A. S. Continuous Auctions and Insider Trading[J].Econometrica, 1985, (6): 1315-1335.
- [5] Garman, M. Market Microstructure[J].Journal of Financial Economics, 1976, (3): 257-275.
- [6] Copeland, T. E., D. Galai. Information Effects on the Bid-Ask Spread [J].The Journal of Finance, 1983, (5): 1457-1469.
- [7] Glosten, L. R., L. E. Harris. Estimating the Components of the Bid/Ask Spread[J].Journal of Financial Economics, 1988, (1): 123-142.
- [8] Amihud, Y., H. Mendelson. Asset Pricing and the Bid-Ask Spread [J].Journal of Financial Economics, 1986, (2): 223-249.
- [9] Huang, R. D., H. R. Stoll. The Components of the Bid-Ask Spread: A General Approach [J].Review of financial studies, 1997, (4): 995-1034.
- [10] Foster, F. D., S. Viswanathan. Variations in Trading Volume, Return Volatility, and Trading Costs; Evidence on Recent Price Formation Models [J].The Journal of Finance, 1993, (1): 187-211.
- [11] Hasbrouck, J. Measuring the Information Content of Stock Trades[J].The Journal of Finance, 1991, (1): 179-207.
- [12] Easley, D., N. M. Kiefer, M. O'hara et al. Liquidity, Information, and Infrequently Traded Stocks [J].The Journal of Finance, 1996, (4): 1405-1436.
- [13] Glosten, L. R., P. R. Milgrom. Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders [J].Journal of Financial Economics, 1985, (1): 71-100.
- [14] Glosten, L. R. Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices [J].The Journal of Finance, 1987, (5): 1293-1307.
- [15] Easley, D., M. O'hara. Price, Trade Size, and Information in Securities Markets [J].Journal of Financial Economics, 1987, (1): 69-90.
- [16] Easley, D., M. López de Prado, M. O'hara. The Microstructure of the "Flash Crash": Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading [J].The Journal of Portfolio Management, 2011, (2): 118-128.
- [17] Easley, D., R. F. Engle, M. O'hara et al. Time-Varying Arrival Rates of Informed and Uninformed Trades [J].Journal of Financial Econometrics, 2008, (2): 171-207.
- [18] Lee, C., M. J. Ready. Inferring Trade Direction from Intraday Data [J].The Journal of Finance, 1991, (2): 733-746.
- [19] Easley, D., M. López de Prado, M. O'hara. Bulk Classification of Trading Activity[R].Available at SSRN, 2012.

- [20] Stoll, H. R. Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests [R].The Journal of Finance, 1989, (1): 115-134.
- [21] 杨之曙,姚松瑶.沪市买卖价差和信息性交易实证研究[J].金融研究, 2004, (4): 45-56.
- [22] 王春峰,董向征,房振明.信息交易概率与中国股市价格行为关系的研究[J].系统工程, 2005, (2): 62-67.
- [23] 周开国,何兴强,柴俊.股票交易活跃性,流动性与基于信息的交易——对H股的微观结构分析 [J]. 财经问题研究, 2006, (8): 50-59.
- [24] 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, (1): 16-24.
- [25] 马丹,牛秀敏,王芳.中国证券市场知情交易概率的动态马尔科夫状态转移模型研究[J].数理统计与管理, 2012, (4): 10.
- [26] 攀登,施东晖.知情交易概率的测度模型及其影响因素分析[J].管理世界, 2006, (6): 18-26.
- [27] Andersen, T. G., O. Bondarenko. Vpin and the Flash Crash[J].Journal of Financial Markets, 2014, 17(1): 1-46.
- [28] Easley, D., M. Lopez de Prado, M. O'Hara. Vpin and the Flash Crash: A Rejoinder [J].Journal of Financial Markets, 2014, (1): 47-52.

(责任编辑:彭晶晶)

## How to Measure the Informed-trading of Market: Theory and Literature Review

YIN Kang, CHENG Zhi-fen

(1.School of Economics, Hubei University of Economics, Wuhan Hubei 430205, China; 2.School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China; 3. luojia College, Wuhan University, Wuhan Hubei 430064, China)

**Abstract:** Informed-trading plays a key role in the information model of market microstructure. Particularly, after the Rational Expectations Model (REE) is introduced to the market microstructure, informed-trading has been the key to understanding market liquidity, market volatility and price discovery. This paper introduces the basic characteristics and significance of informed-trading. Then, the paper summarizes all kinds of methods to measure the informed-trading of market and gives a brief evaluation of the characteristics of each method.

**Key words:** informed-trading; liquidity trading; the probability of informed-trading