

买卖价差与限价指令簿信息： 基于时变 MRR 模型的实证研究^{*}

郑振龙 戴 嵩

〔摘 要〕本文的主要目的是研究限价指令簿信息与买卖价差之间的关系。本文引入限价指令簿信息指标作为模型参数的状态变量,提出了时变 MRR 模型,并基于此模型对中国 A 股市场买卖价差进行了实证研究。本文实证表明,限价指令簿中所体现出的净卖出(买入)压力对原 MRR 模型中的流动性成本参数具有显著影响,且这种影响在买单与卖单中是非对称的;限价指令簿中的订单总量,则可以反映出交易流数据中无法反映的信息不对称程度,对原 MRR 模型中的信息不对称成本参数具有显著影响。另外,通过时变 MRR 模型估计出的隐含价差的日内走势与真实绝对价差及真实相对价差走势吻合,这说明模型可以较好地反映我国 A 股市场买卖价差的性质。

关键词: 限价指令簿信息 时变 MRR 模型 买卖价差

JEL 分类: G11 G12 G14

一、引言

(一)研究背景

股票的价格变动可以分为高频变动(日内变动)与低频变动(日间变动)两类。其中低频变动由于易于观测且数据可得性较强,所以较早进入了金融学家的视野,并且至今仍是学术研究的热点。而对于股票日内价格变化的研究,起步则相对较晚。进入 1980 年代以后,随着高频数据和较高性能计算机的普及,使得研究日内股价变动成为了可能。

导致日内股价变化的因素除了公共信息外,还包括各种市场摩擦。这些摩擦使得股票成交价格并不等于其内在价值——即引起了买卖价差的形成。因此,早期的日内股价变动研究内容大多着眼于买卖价差的构成因素。其中 Demsetz(1968)的研究最早从流动性方面进行建模,来解释买卖价差的形成原因。而 Smidt(1971)、Garman(1976)、Amihud and Mendelson(1980)、O'Hara and Oldfield(1986)、Madhavan and Smidt(1993)等人则在此领域进行了更为深入的研究,并指出流动性因素(做市商的存货成本)是导致买卖价差的因素之一;与此同时,Glosten and Milgrom(1985)、Easley and O'Hara(1987)、Kyle(1985)等学者将交易者区分为知情交易者与非知情交易者,从信息角度对买卖价差进行建模,并指出信息不对称也是导致买卖价差的重要因素之一。此类研究虽然较好地解释了买卖价差形成的原因及其日内变化规律,但都具有一个共同的局限,即模型都是在传统的做市商市场(dealer market)框架下建立的,并没有考虑限价指令簿信息对价格的影响。

近年来,指令驱动市场因其高效、透明、易监管等特点,在全球范围流行开来。许多世界主要股

^{*} 郑振龙,厦门大学闽江学者,教授,博士生导师;戴嵩,华泰联合证券研究所。本文得到以下基金项目的资助: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(项目号:70971114)”;国家自然科学基金青年项目“投资者风险偏好:度量与应用(项目号:71101121)”;教育部人文社科一般项目(07JA790077);教育部留学回国人员科研启动基金“人民币即期与远期汇率关系及外汇市场协同稳定机制研究(教外司留[2008]890)”。

票交易所,如东京股票交易所(Tokyo Stock Exchange)、多伦多股票交易所(Toronto Stock Exchange)、香港联交所以及我国的上海、深圳证券交易所都采取了这一架构。与传统做市商市场不同的是,指令驱动市场并没有固定的做市商,所有交易者均通过递交限价指令(limit order)申报自己希望买入或卖出证券的价格与数量,并由交易系统自动撮合交易。在这类市场中,所有交易者自动为交易对手提供流动性,变相承担起做市商的角色。由于指令驱动市场的限价指令簿(limit order book)中含有大量偏离当前最优报价的订单,很多学者由此认为这些订单具有一定的信息含量,并且会影响投资者的投资策略。所以近年来,金融学术界涌现出大量研究限价指令簿的文献,旨在研究限价指令簿中所含有的信息及指令驱动市场中投资者的行为模式与价格形成机制。其中比较著名的研究成果包括 Glosten(1994)、Rock(1996)、Seppe(1997)、Foucault(1999)等文章。但这类文章的不足之处在于,理论模型大都通过最优化或重复博弈等理论推导市场均衡解,过于复杂,无法直接指导实证检验。

近年来,随着高频数据可得性的增加,我国也涌现出了一批研究市场微观结构的文献。其中杨之曙(2003)、屈文洲和吴世农(2002)、孙培源和施东辉(2002)以及穆启国和吴冲锋(2004)分别对我国大陆股市的买卖价差进行了分析,其结论表明我国大陆股市的价差由信息成本(逆向选择成本)及非信息成本构成,且在日内呈“L”型或不对称“U”型,这是我国学者在微观结构层面研究买卖价差最早的一批文献,但其使用的均为非时变系数模型。雷觉铭和曾勇(2006)则在 30 分钟的频度上对我国大陆股市买卖价差进行了归因分析,认为价差受到最优买卖档订单不平衡程度、成交量、成交笔数及波动率等因素影响。该研究将买卖价差成因与限价指令簿联系起来,但在估计价差是使用的仍然是常系数模型,并未将指令簿信息指标直接整合进理论模型。戴洁(2005)研究了我国股市日内价格走势与当期最优买档订单量价、最优卖档订单量价、报价差异及交易量的关系,对我国股市限价指令簿的特征进行了较为全面的实证性分析。陈瑜(2008)基于 Sandas(2001)及 Glosten and Harris(1988)的模型,研究了大陆股市限价指令簿量价关系及逆向选择成本,并指出与 Glosten-Harris 模型相比,Sandas(2001)的量价关系模型更适合刻画中国股市短期价格行为。该文侧重于研究逆向选择成本,并未对流动性成本进行详细分析。

(二) 本文研究目的和创新之处

虽然之前的学者已经对我国股票市场买卖价差进行了较多实证研究,但直接将限价指令簿信息加入经典模型并同时研究信息成本和流动性成本的文章还非常少。本文的目的便是通过在简单直观的做市商市场模型(MRR 模型)的基础上,将限价指令簿中的信息指标作为状态变量加入其中,使其成为时变模型,并从分笔数据的层面检验当期指令簿信息对日内买卖价差的成分是否有显著影响。

本文的创新点在于提出了含有限价指令簿信息的时变 MRR 模型,并指出净卖出压力对原 MRR 模型中流动性成本参数的影响在买单与卖单中是非对称的。

二、研究方法

(一) 基本研究思路

本章的主要目的是研究买卖价差成分与限价指令簿信息之间的关系。所采用的方法是在 Madhavan, Richardson and Roomans(1997)提出的结构化模型(后文统称 MRR 模型)的基础上,加入基于限价指令簿信息构建的状态变量,将原本的常参数模型转化为时变参数模型。作者认为这样做的意义在于保证模型易于估计的同时,将指令簿信息加入模型,使其更适合中国这样的指令驱动市场。

(二) MRR 模型概述

MRR 模型是 Madhavan、Richardson and Roomans(1997)三位学者提出的用来研究做市商市场日内股价形成机制的模型,该模型由 Glosten and Milgrom(1985)提出的微观结构模型改进而来。

三位作者在文章中指出,日内股价^①的构成因素包括做市商对股票潜在价值的判断及其对提供流动性所要求的补偿。其中做市商对股票当期潜在价值的判断又由三种因素构成:股票上一期的潜在价值、当期的公共信息冲击与当期订单流新息项的冲击。Glosten and Milgrom(1985)的研究指出,订单流的信息项会导致信念的改变,且两者呈正相关关系。由此可以得到做市商对当期股票潜在价值的判断为:

$$\mu_t = \mu_{t-1} + \theta(x_t - E[x_t | x_{t-1}]) + \varepsilon_t \quad (1)$$

其中 μ_t 为 t 时刻股票潜在价值。 x_t 为 t 时刻订单流的方向——当订单为买单时 $x_t=1$, 订单为卖单时 $x_t=-1$ 。 $x_t - E[x_t | x_{t-1}]$ 为本期订单流新息,用来刻画 t 时刻市场上的信息不对称程度。 $\theta > 0$, 表示做市商针对信息不对称所要求的补偿,即逆向选择成本。 ε_t 表示新的公共信息,是一个均值为零的白噪声(white noise)过程。

而真正的交易价格则等于当前股票潜在价值加上做市商对提供流动性要求的补偿,再加上随机扰动项,即:

$$p_t = \mu_t + \phi x_t + \xi_t = \mu_{t-1} + \theta(x_t - E[x_t | x_{t-1}]) + \phi x_t + \varepsilon_t + \xi_t \quad (2)$$

其中 $\phi > 0$, 表示做市商对提供流动性所要求报酬的大小,而 ξ_t 是均值为零的扰动项,且与 ε_t 不相关。

由式(2)可以看出,股票交易价格与其上期内在价值的偏离主要由 ϕx_t 和 $\theta(x_t - E[x_t | x_{t-1}])$ 构成,也即流动性成本与信息成本为买卖价差的主要成分。另外, Madhavan, Richardson and Roomans(1997)还根据模型推导出了隐含价差,其期望值等于 $2(\theta + \phi)$ 。

(三) 模型扩展——加入限价指令簿信息

虽然 MRR 模型具有简单直观的经济含义和较为简单的估计形式,但其最初针对的是做市商市场,且参数形式(信息不对称成本与流动性成本)均为非时变的。我国大陆股票市场属于典型的指令驱动市场,投资者除了可以在最优买卖档位投放订单之外,还可以根据自己对股票价值的估计递交限价指令。所以除了最优报价档位的订单流之外,限价指令簿内的其他信息可能会对日内价格的变动及买卖价差产生影响。基于此点,本文希望将从指令簿中提取的相关指标,在分笔数据的频度下作为状态变量加入模型参数,将原本的静态模型转化为更为灵活的时变系数模型,以研究买卖价差成分与指令簿信息之间的关系。

1. 对原模型信息不对称成本参数(θ)的扩展

近年来,有大量文献分析了信息交易者在指令驱动市场中的订单投放策略。Kaniel and Liu(2006)研究了信息交易者与耐心的非信息交易者(patient uninformed investors)的最优订单投放策略,他们指出:如果私有信息持续的时间足够长,则投放限价指令是信息交易者的最优策略。Bloomfield, O'Hara and Saar(2005)指出,由于信息交易者知晓资产的真实价格,他们递交的限价指令没有被捡便宜(picking-off)的风险,所以提交限价指令是其最优策略。同时,Goettler, Parlour and Rajan(2009)通过求解均衡数值解指出,平均来看,投机者是市场流动性的提供者,而限价指令则主要由信息交易者提交。

综合以上学者的论述,本文认为市场上的信息不对称程度并不完全由 MRR 模型中提出的市价指令流所体现,它还反映在限价指令簿中订单总量上。限价指令簿中订单总量越多,说明信息交

^① 本文中所述的日内股价的构成因素与日内买卖价差的构成因素并非同一概念,前者的范围大于后者:买卖价差是日内股价的一个构成因素,前者还包括新公共信息带来的冲击。

易者所占比重大, 整个市场的信息不对称程度越小。所以在市价指令流提供的信息不变的情况下, MRR 模型中反映信息不对称成本的参数应该是指令簿中订单总量的减函数, 即指令簿中的订单越多, 信息不对称程度越低, 交易者对信息不对称所要求的补偿也越低。于是我们可以将指令簿中的订单总量作为信息不对称参数的状态变量, 本文将其形式设定为:

$$\theta_t = \delta + \frac{\kappa}{q_t} \quad (3)$$

其中 q_t 为当期指令簿中的订单总量, κ 为信息不对称成本参数对订单总量的反映程度, δ 代表原 MRR 模型中参数的非时变部分。

2. 对原模型流动性成本参数(ϕ)的扩展

指令驱动市场与传统报价驱动市场一个最重要的区别就在于, 指令驱动市场上没有专门提供流动性的做市商, 市场上的流动性完全由递交限价指令的交易者提供。相应地, 交易者也会对其提供的流动性收取相应的报酬, 这部分成本就包含在 MRR 模型的参数 ϕ 中。

原模型中非时变的流动性成本参数是否会受到限价指令簿信息的影响呢? 作者认为, 该参数会受到限价指令簿所表现出的卖出(买入)净压力的影响。

首先, 本文通过修改 Cao, Hansch and Wang (2009) 提出的加权价格指标, 定义了一个代表指令簿卖出(买入)净压力的指标 WP, 其具体形式为:

$$WP = \frac{\sum_{j=1}^5 (Qbuy_j Pbuy_j + Qsell_j Psell_j)}{\sum_{j=1}^5 (Qbuy_j + Qsell_j)} / MIDprice - 1 \quad (4)$$

其中 $Qbuy_j$ 和 $Pbuy_j$ ($i=1, \dots, 5$) 分别表示第 1 到 5 档买单的量价, $Qsell_j$ 和 $Psell_j$ ($i=1, \dots, 5$) 分别表示第 1 到 5 档卖单的量价, $MIDprice$ 是买卖中间价。该指标可按照以下方式理解: 等式右边括号中的部分是以订单量为权重的加权价格, 若卖方订单量较大, 则加权价格会向卖方倾斜, 大于中间价 $MIDprice$, 加权价格除以中间价并减 1 后为正, 此时表现为卖方净压力。相反, 如果买方订单量较大, 则加权价格小于中间价 $MIDprice$, 此时指标为负, 表现为买方净压力。

从直观上来看, 当指令簿表现为卖方净压力时, 卖方交易者竞争比较激烈, 当他们为买单(交易发生在最优卖档)提供流动性时, 会索取较低的流动性报酬。相反, 当买方交易者为卖单提供流动性时(交易发生在最优买档), 则会索取较高的流动性报酬。所以本文认为, 当交易为买单时, ϕ_t 是 WP 的减函数, 而交易为卖单时, ϕ_t 是 WP 的增函数。为了区分这种买卖单中存在的不对称性, 本文在 ϕ_t 的具体函数形式中加入了虚拟变量, 将其设定为:

$$\phi_t = \gamma + \beta_1 WP_t + \beta_2 I_{2,t} WP_t \quad (5)$$

其中 γ 为原模型流动性成本参数中的非时变部分, $I_{2,t}$ 为虚拟变量, 当交易为买单的时候为 0, 交易为卖单的时候为 1。于是, 当交易为买单时, $\phi_t = \gamma + \beta_1 WP_t$, 而当交易为卖单时, $\phi_t = \gamma + (\beta_1 + \beta_2) WP_t$ 。根据前文论述的流动性成本与净卖出(买入)压力的关系, 作者认为 $\beta_1 < 0, \beta_2 > 0$ 且其绝对值要大于 β_1 。特别地, 当买卖压力相等时, ϕ_t 与原 MRR 模型中参数的大小相同。如果假设卖方净压力与买方净压力与订单流独立, 且期望为 0, 即:

$$E(WP_t) = E(WP_t | I_2 = 1) = E(WP_t | I_2 = 0) = 0 \quad (6)$$

则 ϕ_t 的无条件均值可写为:

$$\begin{aligned} E(\phi_t) &= E(\phi_t | I_2 = 0)P(I_2 = 0) + E(\phi_t | I_2 = 1)P(I_2 = 1) \\ &= E(\gamma + (\beta_1 + \beta_2) WP_t | I_2 = 1)P(I_2 = 1) + E(\gamma + \beta_1 WP_t | I_2 = 0)P(I_2 = 0) \\ &= \gamma + (\beta_1 + \beta_2)E(WP_t | I_2 = 0) + \beta_1 E(WP_t | I_2 = 1) \\ &= \gamma \end{aligned} \quad (7)$$

即 ϕ_t 的无条件均值等于原模型的非时变参数,这与我们的直观感觉是吻合的。

(四)模型估计方法

1.待估模型形式

将 MRR 模型原参数加入指令簿信息作为状态变量后,模型形式变为:

$$p_t = \mu_{t-1} + \left(\delta + \frac{\kappa}{q_t}\right)(x_t - E[x_t | x_{t-1}]) + (\gamma + \beta_1 WP_t + I_{2,t} \beta_2 WP_t)x_t + \varepsilon_t + \xi_t \quad (8)$$

Madhavan, Richardson and Roomans (1997) 在其文章中证明, $E[x_t | x_{t-1}] = \rho x_{t-1}$, 其中 ρ 为订单流 x_t 的一阶自相关系数, 且 $p_{t-1} = \mu_{t-1} + \phi_{t-1} x_{t-1} + \xi_{t-1}$ 。将其带入式 (4.8) 中, 得到模型的可估计形式为:

$$\begin{aligned} \Delta p_t = & \alpha + (\gamma + \delta)x_t - (\gamma + \delta\rho)x_{t-1} + \kappa \frac{x_t}{q_t} - \kappa\rho \frac{x_{t-1}}{q_t} + \beta_1(WP_t x_t - WP_{t-1} x_{t-1}) \\ & + \beta_2(I_{2,t} WP_t x_t - I_{2,t-1} WP_{t-1} x_{t-1}) + \varepsilon_t + \xi_t - \xi_{t-1} \end{aligned} \quad (9)$$

其中待估参数为 $\alpha, \beta_1, \beta_2, \kappa, \delta, \gamma$ 以及 ρ 。

2.估计方法选择

为避免残差条件异方差及订单流非正态分布所带来的影响, 本文选择广义矩估计方法 (GMM) 对模型进行估计。

广义矩估计方法是普通矩估计方法的一般化, 其原理是假设参数满足一定的理论矩条件, 然后用样本矩代替总体矩来进行参数估计。同时, 广义矩估计是一个大样本估计方法, 在大样本条件下, GMM 估计量是渐进有效的, 且渐进服从 t 分布, 而在小样本中, GMM 估计量往往是无效的。由于本文采用的是全年分笔高频数据, 所以样本点的数量可以得到有效保证, 无需担心估计量的无效性。广义矩估计的另一个优点是, 在估计的时候只需要给出一些总体矩条件的表达式 (如回归方程残差与自变量或工具变量无关), 而不必给出各变量的概率密度函数。这样就避免最小二乘回归所要求的正态性假设。值得注意的是, 当估计时选择的矩条件多于待估参数个数时, 广义矩估计系统会产生过度识别现象。为充分利用样本提供的所有信息, 最终的估计结果必须尽量满足各个矩条件方程, 将整个方程系统的冲突程度降到最低。广义矩估计解决过度识别问题的标准是最小化样本矩之间的加权距离。假设 φ 为待估参数向量, 样本矩条件为 φ 的函数, $g_T(\varphi)$ 。在过度识别情况下, 待估参数向量的估计值可表示为:

$$\hat{\varphi}_{GMM} = \arg \min_{\hat{\varphi}} \{g_T(\hat{\varphi})' W_T g_T(\hat{\varphi})\} \quad (10)$$

其中 W_T 是一个权重向量, 用来定义不同矩条件在系统中的重要程度。通常来讲, 在估计的第一阶段, W_T 被设计为单位向量, 以获得参数的一致估计量; 在估计的第二阶段, W_T 被设计为样本方差-协方差矩阵的逆矩阵, 从而得到参数的渐进有效估计量。

针对上节中推导出的可估计模型, 其残差为 $u_t = \varepsilon_t + \xi_t - \xi_{t-1}$ 。故其矩条件可以写为如下形式:

$$E \begin{bmatrix} x_t x_{t-1} - x_t^2 \rho \\ u \\ u x_t \\ u x_{t-1} \\ u \frac{x_t}{q_t} \\ u \frac{x_{t-1}}{q_t} \\ u (WP_t x_t - WP_{t-1} x_{t-1}) \\ u (I_{2,t} WP_t x_t - I_{2,t-1} WP_{t-1} x_{t-1}) \end{bmatrix} = 0 \quad (11)$$

其中的第一个矩条件定义了订单流的一阶自相关,第二个矩条件要求模型残差均值为 0(或理解为模型残差与工具变量 1 不相关),后面几个矩条件则保证了模型残差与自变量无关,即普通最小二乘回归的约束条件。

三、实证结果

在具体的实证过程中,本文采用了 Madhavan, Richardson and Roomans(1997)文章中的方法,进行分时段估计。本文将每天的交易时间分为 4 个时段,分别是上午 9:35–10:30, 10:30–11:30, 下午 13:00–14:00 及 14:00–14:55。这样做一方面可以让我们通过观察不同时段参数估计的结果,检验模型的稳健性,另一方面可以让我们直观地比较模型隐含价差与真实价差日内走势的异同。

(一) 各变量描述性统计量

在报告正式的模型估计结果之前,本文首先报告了模型中所包含各个变量的描述性统计量,以便读者对其有一个直观的认识。由于样本中各只股票的特征(如价格、可流通股本数量、成交量、不同交易记录间价格的绝对变化量等)有较大差距,所以本文并未将所有股票交易数据放在一起(pooling)统计,而是将各只股票分开统计,并报告了不同变量均值与标准差等统计特征的分布情况。其中表 1 报告了全天未分段的各变量描述性统计量,而表 2 则报告了日内不同时段各变量均值的分布情况。

表 1 模型各变量描述性统计特征

变量	均值	最大	最小	标准差
$ \Delta p $ 均值	0.018426	0.101169	0.004359	0.014278
Δp 的方差	0.005278	0.138129	6.94E-05	0.016605
x_t 均值	0.026499	0.064297	-0.09473	0.020776
WP_t 均值	4.83E-05	0.000345	-0.00042	0.000143
$ WP_t $ 均值	0.001218	0.002921	0.00065	0.00039
$\frac{1}{q_t}$ 均值	1.1E-05	5.07E-05	1.23E-07	9.02E-06

从表 1 可以看出, $|\Delta p|$ 的平均水平在 0.018 左右,大概相当于 1.8 个最低报价单位。由于分笔记录间的价格变化在一定程度上反映了买卖价差的大小,所以百分位数量级的价格变化均值是合理的。而 Δp 方差的数量级与其水平相当,这说明各只股票之间价格变动水平的差距较大,因为价值 100 元的股票与价值 5 元的股票价格变化幅度很可能不在一个数量级。我们之所以没有选择报告相对价格变动(即分笔收益率)的描述性统计量,是因为原模型因变量是价格变化的绝对值而非相对值,且模型中两个参数反映的也是绝对价差的构成。正是为了避免这种价格差异所带来的影响,本文才选择了将样本股票分开进行 GMM 估计的实证形式。

而订单流方向变量的均值为 0.026,较接近于 0,且与其绝对值水平 1 差距为 38 倍,这说明总体来看买单卖单的分布未出现明显的不对称性。

同样,对于净卖出(买入)压力指标 WP_t 来说,其均值在 $10E-5$ 的数量级,而绝对值 $|WP_t|$ 的均值在 $10E-3$ 数量级,两者相差 26 倍左右。所以总体来看,指令簿内的买卖压力较为平衡。这也从侧面印证了我们关于 $E(WP_t)=0$ 的假设的合理性。

$\frac{1}{q_t}$ 均值和最大值的数量级都在 $10E-5$,而最小值的数量级为 $10E-7$,标准差数量级在 $10E-6$ 。

这种情况与前面 Δp 类似,是由不同股票本身可流通股数差异造成的。但通过观察可以发现, $\frac{1}{q_i}$ 的均值比标准差大一个数量级,说明其离散性不大,基本集中在 $10E-5$ 这个数量级上。

表 2 模型各变量均值日内变化

变量	9:35-10:30	10:30-11:30	13:00-14:00	14:00-14:55
$ \Delta p $	0.021259	0.017996	0.017298	0.016924
x_i	0.023351	0.020267	0.037243	0.024857
WP_i	-7.6E-06	5.37E-06	7.16E-05	0.000132
$ WP_i $	0.001285	0.001223	0.001195	0.001162
$\frac{1}{q_i}$	1.21E-05	1.1E-05	1.07E-05	1.03E-05
$Spread$	0.037606	0.032908	0.032023	0.030897
$RLTVspread$	0.001473	0.001331	0.001292	0.001252
Δp 的方差	0.005333	0.010879	0.002506	0.002262

注:表 2 在模型变量之外加入了绝对价差($Spread$)与相对价差($RLTVspread$)日内均值的分布,以观察价差在日内的变化趋势。

通过观察表 2 可以看出,除 WP_i 外,各变量均值在日内变化幅度不大,均维持在同一数量级。而 WP_i 均值本身的符号并不统一,且数量级仍然很小。值得注意的是,价格变化绝对值、绝对价差与相对价差在日内均表现为弱“L”型,这与之前学者(杨之曙,2003;屈文洲和吴世农,2002;孙培源和施东辉,2002;穆启国和吴冲锋,2004)的研究结论是一致的。

以上是作者对模型中各变量描述性统计量的介绍,在下面一节中,作者将报告具体的模型估计结果。

(二)模型参数估计结果分析

表 3-表 6 分别报告了 4 个时段模型参数的估计值。由于本文是将样本内股票分开进行估计的,所以报告中呈现的是各个参数在横截面上的描述性统计量与相关检验的结果。

通过分析表 3-表 6,我们可以得到以下结论:

1.截距项均值的符号在四个时段并不一致(9:35-10:30 及 14:00-14:55 两个时段为负,其他时段为正)且数量级较小。同时,该系数在前三个时段皆未能通过均值异于零的横截面检验,且在个股中显著的比例较低(第一时段为 35.56%,第二时段为 47.62%,第三时段为 28.89%,第四时段为 35.56%)。通常来说,模型因变量中不能被自变量所解释的部分会包含在残差项或截距项中。截

表 3 模型估计结果 1(9:35-10:30)

系数	均值	最大值	最小值	标准差	横截面 t 值	符号正确且显著的个股百分比
α	4.21E-05	0.001169	-0.00034	0.000221	1.802397	35.56%
β_1	-1.23996	0.534442	-4.93756	1.210875	-9.71468*	74.44%
β_2	2.41991	9.179242	-0.90672	2.251924	10.19452*	80.00%
δ	0.004002	0.019003	0.000211	0.00326	11.6491*	100.00%
γ	0.008299	0.050175	0.003874	0.00612	12.86529*	100.00%
κ	128.7896	1097.837	-30.2367	145.3006	8.408813*	91.11%
ρ	0.058525	0.140206	-0.01952	0.037136	14.95066*	88.89%

注:横截面 t 值是用来检验模型系数均值是否显著异于零,“*”表示系数均值在 1%的水平上显著异于零。

表 4 模型估计结果 2(10:30-11:30)

系数	均值	最大值	最小值	标准差	横截面 t 值	符号正确且显著的个股百分比
α	-2E-05	0.000761	-0.0004	0.000186	-0.99513	47.62%
β_1	-0.90808	15.65263	-15.4066	3.983954	-2.12603**	66.67%
β_2	1.638378	27.1528	-29.1475	7.79487	1.960492**	67.86%
δ	0.00356	0.013379	0.000318	0.002452	13.54445*	97.62%
γ	0.008469	0.023663	0.004125	0.004341	18.19605*	100.00%
κ	99.71598	488.5083	-290.048	107.759	8.631195*	89.29%
ρ	0.074695	0.140532	0.003407	0.029729	23.43545*	98.81%

注: (1) 横截面 t 值是用来检验模型系数均值是否显著异于零, “*”和“**”分别表示系数均值在 1% 和 5% 的显著性水平上显著异于零; (2) 在此时间段中系数估计出现了比较明显的异常值(outlier), 例如有的大于 10000, 为剔除异常值的影响, 本文删除了首尾各三条异常数据, 故样本容量为 84。

表 5 模型估计结果 3(13:00-14:00)

系数	均值	最大值	最小值	标准差	横截面 t 值	符号正确且显著的个股百分比
α	7.39E-06	0.000713	-0.00042	0.000139	0.502991	28.89%
β_1	-1.5025	0.490327	-21.2632	2.570074	-5.54613*	73.33%
β_2	2.971185	37.11676	-1.61514	4.737648	5.949604*	75.56%
δ	0.003374	0.014669	0.00015	0.002765	11.57656*	100.00%
γ	0.008299	0.050175	0.003874	0.00612	12.86529*	100.00%
κ	130.0033	2268.266	-423.093	255.8428	4.820615*	90.00%
ρ	0.086095	0.178801	0.005144	0.034406	23.73884*	98.89%

注: 横截面 t 值是用来检验模型系数均值是否显著异于零, “*”表示系数均值在 1% 的水平上显著异于零。

表 6 模型估计结果 4(14:00-14:55)

系数	均值	最大值	最小值	标准差	横截面 t 值	符号正确且显著的个股百分比
α	-9.2E-05	0.00021	-0.00064	0.000125	-6.96511*	35.56%
β_1	-1.63812	0.602452	-14.0521	2.048226	-7.58733*	76.67%
β_2	3.105877	26.01493	-1.68187	3.942274	7.474097*	74.44%
δ	0.003024	0.012304	5.48E-05	0.002405	11.92851*	100.00%
γ	0.008077	0.042342	0.003934	0.005417	14.14495*	100.00%
κ	147.622	2463.499	-896.678	294.6275	4.753341*	91.11%
ρ	0.082223	0.15917	-0.0006	0.036312	21.48168*	98.89%

注: 横截面 t 值是用来检验模型系数均值是否显著异于零, “*”表示系数均值在 1% 的水平上显著异于零。

距项显著程度不高, 可以从侧面印证模型解释变量具有较好的解释能力。综合以上分析, 本文认为从整体来看, 模型自变量对因变量 Δp 有较好的解释能力。

2. WP_i 前面的系数 β_1 均值为负, β_2 均值为正, 且 β_2 的绝对值大于 β_1 。这与本文在前面所做出的预测一致, 即买单与卖单中流动性成本对指令簿净卖出(买入)压力的反应是不对称的。从横截面数据来看, 在四个时段中, β_1 与 β_2 两个系数的均值在横截面上皆显著异于零, 且在个股中显著比例较高 (其中 β_1 在四个时段中符号正确且显著的比例分别为: 74.44%、66.67%、73.33%与

76.67%, β_2 在四个时段中符号正确且显著的比例分别为:80%、67.86%、75.56%与 74.44%)。这样的结果表明,从样本整体来看,指令簿卖出(买入)净压力对 MRR 模型中流动性成本具有显著影响,且这种影响在买单与卖单中是非对称的:当交易为买单时, $\phi_i = \gamma + \beta_1 WP_i$; 当交易为卖单时, $\phi_i = \gamma + (\beta_1 + \beta_2) WP_i$ 。根据表 3-表 6 所示的回归结果以及表 2 所示的描述性统计量可以看出, β_1 的均值为 -1.32, $(\beta_1 + \beta_2)$ 的均值为 1.21, 按照 $|WP_i| = 0.00128$ 的水平大致估算,同样的买卖压力在买单和卖单中对流动性成本造成的不对称冲击之差大概为 0.00324 左右,数量级与相同,且绝对值略小。作者认为这样的结果从直观上来看是合理的。

3. $\frac{1}{q_i}$ 前面的系数在四个时段中均值皆为正,且在横截面上显著异于零。这与本文前面的预测同样是一致的,即限价指令簿中的订单量可以反映交易订单流之外的信息不对称程度,指令簿中订单量越大,信息不对称程度越低,价差中的信息不对称成本也越低。从个股层面来看, κ 符号正确且显著的比例很高(分别为 91.11%、89.29%、90.00%和 91.11%),这进一步印证了模型拓展部分假设的正确性。从数值大小上来看, κ 在四个时段中的均值为 126.53, 而 $\frac{1}{q_i}$ 的均值为 $1.1E-05$, 所以可以大致估算指令簿信息对信息不对称成本的影响为 0.00115。该结果与 δ 数量级相同,且绝对值略小。作者认为该结果从直观上来看也是合理的。

4. 代表原 MRR 模型中非时变参数的 δ 与 γ 在估计中表现出非常高的稳定性(全部四个时段的 360 个个股层面估计中,只有 1 个 δ 不显著, γ 则全部显著大于零),且数量级($10E-3$)正确。这从侧面表明本文的拓展没有破坏原模型的性质,改进是合理的。

5. 订单流的一阶自相关系数 ρ 在横截面与个股层面高度稳定,其均值水平在 0.06-0.1 之间。Madhavan, Richardson and Roomans (1997) 证明,买单(卖单)连续出现的概率,即 $E(x_i = x_{i-1} | x_{i-1})$ 等于 $\frac{\rho + 1}{2}$ 。由此可以计算出买单(卖单)连续出现的概率略大于 0.5。该结论与 MRR 文章中的结论一致。

(三) 模型隐含价差分析

根据前面的介绍, MRR 模型中推导出的单只股票的隐含价差(implied spread)的期望值可以写做 $2[E(\theta_i) + E(\phi_i)]$ 。同时,根据 $E(\phi_i) = \gamma$, 可推得隐含价差为 $2(\kappa \times E[\frac{1}{q_i}] + \delta + \gamma)$ 。如果用样本矩替代总体矩,隐含价差可表示为 $2[\kappa \times \overline{(1/q_i)} + \delta + \gamma]$ 。

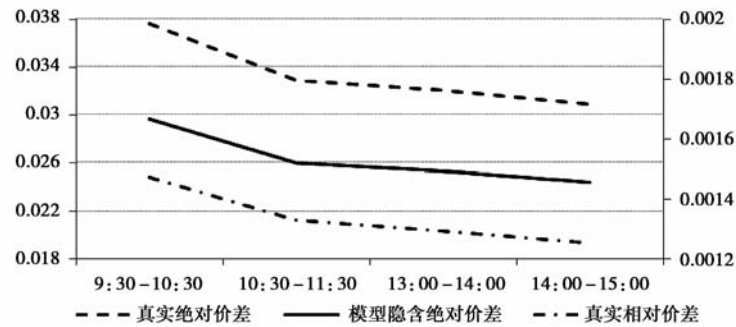
表 7 报告了 4 个时段中隐含价差、真实绝对价差及真实相对价差在横截面均值上的日内变化情况。

表 7 隐含价差、真实绝对价差及真实相对价差横截面均值变化趋势

变量	9:35-10:30	10:30-11:30	13:00-14:00	14:00-14:55
隐含价差	0.0296	0.02601	0.025311	0.024382
<i>Spread</i>	0.037606	0.032908	0.032023	0.030897
<i>RLTVspread</i>	0.001473	0.001331	0.001292	0.001252

注:变量 spread 与 RLTVspread 的含义与表 2 相同。

由表 7 可以看出,隐含价差与真实绝对价差相比略小,这与 Madhavan, Richardson and Roomans (1997) 的结论一致。同时,模型隐含价差在日内呈现弱“L”型,与真实绝对价差和真实相对价差的日内走势基本一致,如图 1 所示。这表明改进后模型所推导出的隐含价差表现出了真实价差的特征,进一步印证了模型的合理性。



注: 左侧坐标轴为模型隐含价差及真实绝对价差刻度, 右侧坐标轴为真实相对价差刻度。

图 1 隐含价差、真实绝对价差及真实相对价差均值日内走势

四、结论

本文主要研究内容为限价指令簿信息与买卖价差成分之间的关系。作者在 MRR 模型基础上引入限价指令簿信息代理指标作为参数状态变量, 将原本非时变的模型扩展为时变模型。

本文实证结果表明, 限价指令簿中体现出的净卖出(买入)压力指标对原 MRR 模型中的流动性成本参数有显著影响, 且相同的净卖出(买入)压力在买单与卖单中对流动性成本的影响是不对称的。而限价指令簿中的订单总量则可以反映出交易订单流冲击之外的信息不对称程度, 且对原 MRR 模型中的信息不对称成本参数有显著影响。同时, 状态变量的加入并未破坏原模型的属性, 原模型参数的非时变部分的估计结果非常稳定, 且符号与数量级均正确。由时变模型估计出的隐含价差不论在绝对大小还是日内走势上, 都体现了真实价差的特性, 说明扩展后的模型可以较好地解释买卖价差的形成机制。

参考文献

- 陈瑜(2008):《中国股票市场限价指令簿量价关系及逆向选择成本研究》, 电子科技大学博士学位论文。
- 戴洁(2005):《中国股票市场限价委托单簿与委托单流的实证研究》, 清华大学博士学位论文。
- 穆启国、吴冲锋(2004):《深圳证券交易所买卖价差的构成分析》,《系统工程理论与方法应用》,第3期。
- 雷觉铭、曾勇(2006):《上海股市买卖价差成分分析》,《系统工程》,第6期。
- 屈文洲、吴世农(2002):《中国股市微观结构的特征分析——买卖报价价差模式及影响因素的实证研究》,《经济研究》,第1期。
- 孙培源、施东辉(2002):《微观结构、流动性与买卖价差——一个基于上海股市的经验研究》,《世界经济》,第4期。
- 杨之曙(2003):《上海股市日内流动性——深度变化实证研究》,《金融研究》,第6期。
- Amihud, Y. and H. Mendelson (1980): "Dealership Market: Market-Making with Inventory", *Journal of Financial Economics*, 8, 31-53.
- Bloomfield, R., M. O'Hara and G. Saar(2005): "The 'Make or Take' Decision in an Electronic Market: Evidence on the Evolution of Liquidity", *Journal of Financial Economics*, 75, 165-199.
- Cao, C., O. Hansch and X. Wang (2009): "The Information Content of an Open Limitorder Book", *Journal of Futures Markets*, 29, 16-41.
- Demsetz, H.(1968): "The Cost of Transacting", *Quarterly Journal of Economics*, 82, 33-53.
- Easley, D. and M. O'hara (1987): "Price, Trade Size, and Information in Securities Markets", *Journal of Financial Economics*, 19, 69-90.

- Foucault, T.(1999): "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market", *Journal of Financial Markets*, 2, 99-134.
- Garman, M. (1976): "Market Microstructure", *Journal of Financial Economics*, 3, 257-275.
- Glosten, L. (1994): "Is the Electronic Open Limit Order Book Inevitable?", *Journal of Finance*, 49, 1127-1161.
- Glosten, L. and P. Milgrom (1985): "Bid, Ask and Transaction Prices in a Market-maker Market with Heterogeneously Informed Traders", *Journal of Financial Economics*, 14, 1127-1161.
- Glosten, L. and L. Harris (1988): "Estimating the Component of the Bid-Ask Spread", *Journal of Financial Economics*, 21, 123-142.
- Goettler, R. , C. Parlour and U. Rajan (2009): "Informed Traders and Limit Order Markets", *Journal of Financial Economics*, 93, 67-87.
- Kaniel, R. and H. Liu(2006): "What Orders do Informed Traders Used", *Journal of Business*, 79, 1867-1913.
- Kyle, A. (1985): "Continuous Auctions and Insider Trading", *Econometrica*, 53, 1315-1335.
- Madhavan, A., M. Richardson and M. Roomans (1997): "Why do Security Prices Change? A Transaction-Level Analysis of Nyse Stocks", *Review of Financial Studies*, 10, 1035-1064.
- Madhavan, A. and S. Smidt(1993): "An Analysis of Changes in Specialist Inventories and Quotations", *Journal of Finance*, 48, 1595-1628.
- O'hara, M. and G. S. Oldfield(1986): "The Microeconomics of Market Making", *Journal of Financial and Quantitative Analysis*, 21 (04), 361-376.
- Rock, K.(1996): "The Specialist's Order Book and Price Anomalies", *Review of Financial Studies*, 9, 1-20.
- Sandas, P.(2001): "Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market", *Review of Financial Studies*, 14, 705-734.
- Seppi, D. (1997): "Liquidity Provision with Limit Orders and a Strategic Specialist", *Review of Financial Studies*, 10, 103-150.
- Smidt, S. (1971): "Which Road to an Efficient Stock Market: Free Competition or Regulated Monopoly?", *Financial Analysts Journal*, 27, 18-69.

(责任编辑:罗 滢)