

学校编号: 10384

分类号密级

学号: 15620081152088UDC

厦 门 大 学

硕士学位论文

**限价指令簿信息含量：中国 A 股市场的  
证据**

**Information content of the limit order book:**

**Evidence from Chinese Stock Market**

**戴嵩**

指导教师姓名: 郑振龙教授

专业名称: 金融工程

论文提交日期: 2011 年 4 月

论文答辩时间: 2011 年 月

学位授予日期: 2011 年 月

答辩委员会主席:

评阅人:

2011 年 4 月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为( )课题(组)的研究成果,获得( )课题(组)经费或实验室的资助,在( )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（     ） 1. 经厦门大学保密委员会审查核定的保密学位论文，  
于     年     月     日解密，解密后适用上述授权。

（     ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年     月     日



## 摘 要

近年来,随着高频数据可得性的增强,我国股票市场的微观结构成为金融学术界研究的热点问题之一。我国 A 股市场属于典型的指令驱动市场,这种市场结构与传统报价驱动市场的一个重大区别是,交易者可以根据自己对股票价格的估值,递交不同于当前最优买卖报价的限价指令。许多国内外研究表明,指令驱动市场上的限价指令簿中包含了许多交易中无法反映的信息,而这些信息可能会影响投资者的交易策略、股票价格的形成机制(比如买卖价差)以及日内短期收益。本文的主要写作目的便是研究限价指令簿信息与买卖价差及内日短期收益之间的关系。作者在本文的第四章与第五章中分别报告了以上两部分研究的模型设定与实证结果。

首先,本文在第四章分析了限价指令簿信息对买卖价差构成成分的影响。作者在此部分研究中采用的基础模型是被广泛应用于买卖价差成分研究中的 **MRR** 模型。作者在原模型基础上,引入限价指令簿信息指标作为模型参数的状态变量,将原来的非时变模型转化为灵活的时变系数模型,并基于此模型进行了实证研究。本部分实证得出两个主要结论:(1) 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;(2) 而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。由于新加入的状态变量并没有破坏原模型的结构,模型参数中的非时变部分仍然高度显著,且绝对大小合理,而且通过时变 **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。

其次,本文在第五章参考 Cao、Hansch 和 Wang(2009)的文章,通过多元回归的方法检验了滞后期限价指令簿信息(买卖压力与订单总数)与股票日内收益之间的关系。与 Cao、Hansch 和 Wang 三位作者不同的是,为了排除指令簿各档位信息存在的高度多重共线性问题,本文通过简单加和及主成份分析两种方法整合提取指令簿的总体信息(原文使用的则是分档信息),来检验指令簿总体信息指标与短期收益之间的关系。实证结果表明,通过简单加和与通过主成份分析提

取出的买卖压力指标具有高度的一致性，两者均对短期收益有显著影响：上期卖出净压力越大，当期收益越低；而代表市场活跃程度的订单总量指标同样对短期收益影响显著：上期订单量越大，当期收益越高。同时，作者还检验了最优报价档位信息与非最优报价档位信息对短期收益的不同影响。实证结果显示，非最优档位的买卖压力比最优档位的买卖压力对短期收益的解释能力更强，而最优档位的订单量则比非最优档位的订单量更具解释能力。

**关键词：**限价指令簿信息；时变 MRR 模型；日内收益

## Abstract

As the availability of high frequency data of Chinese stock market increases, the micro structure of this market has become a hot research field among scholars. Both Shanghai and Shenzhen Stock exchange in China are order-driven markets. One of the remarkable difference between this market structure and the traditional quote-driven market is that there would be plenty of limit order quotes that are different from the best bid and ask prices, which may contain additional information beyond the transaction prices and order flows. This kind of information may influence the traders' investment strategies, the intraday price formation mechanism (i.e. the bid-ask spread) as well as the intraday short term return. The purpose of this dissertation is to investigate the relationship between those information and the bid-ask spread as well as the short term return. The author reported the study methodologies and empirical results in Chapter 4 and Chapter 5 respectively.

In Chapter 4 of this dissertation, the author studies the relationship between the information content of the limit order book and bid-ask spread. The author modified the MRR model which is widely adopted in analyzing the components of bid-ask spread by introducing limit order book information proxies as state variables and turned the original model into a more flexible time-varying model to analyze how order book information would affect the components of the bid-ask spread. The empirical study implies two main conclusions. (1) The selling pressure will significantly affect the liquidity cost parameter of the original MRR model and that this effect is asymmetric between buyer and seller initiated orders; (2) The order quantity of the whole order book can reveal additional information beyond the transaction flows and have significant impact on the information cost parameter of the original model. Moreover, the empirical results indicate that the intraday pattern of the implied spread which is derived from the time-varying model is highly consistent with that of the real bid-ask spread. All the facts above imply that the formation mechanism and intraday pattern of bid-ask spread in Chinese stock market can be well explained by the modified time-varying MRR model.

In Chapter 5 of this dissertation, the author tests the relationship between the limit book information and the intraday short term return by following Cao, Hansch and Wang(2009)'s multivariate regression method. In order to avoid the multicollinearity effect among the selling(buying)pressure proxies of different steps of the order book, the author extracts the total selling pressure proxy of the entire order book by adopting simple summation method and principal components analysis method respectively. The empirical results indicate that the lagged total selling pressure proxies are negatively related to the current short term returns as well as the lagged quantity of the orders in the entire order book is positively related to the current short term return. In addition, the author distinguishes the information of the best bid-ask step from the information of faraway steps, and test their predictability to the future short term return. The results show that selling pressure implied by faraway steps has stronger predictability while the order quantity of the best bid-ask steps seems to be more informational.

**Key words:** Limit Order Book; Time Varying MRR Model; Intraday Return



# 目录

<b>第一章引言</b>	<b>1</b>
1.1 研究背景	1
1.2 本文研究目的和创新之处	2
1.3 文章结构	3
<b>第二章文献综述</b>	<b>4</b>
2.1 传统市场微观结构研究	4
2.2 限价指令簿相关研究	6
2.3 国内相关研究	7
<b>第三章我国股票市场微观结构及指令簿数据概述</b>	<b>9</b>
3.1 我国股票市场微观结构概述	9
3.2 数据说明	10
3.3 限价指令簿描述性统计量	10
<b>第四章限价指令簿信息与买卖价差成分关系研究</b>	<b>14</b>
4.1 研究方法	14
4.1.1 基本研究思路	14
4.1.2 MRR 模型概述	14
4.1.3 模型扩展——加入限价指令簿信息	15
4.1.4 模型估计方法	18
4.2 实证结果	19
4.2.1 各变量描述性统计量	20
4.2.2 模型参数估计结果分析	22
4.2.3 模型隐含价差分析	27
4.3 本章小结	28
<b>第五章限价指令簿信息与日内收益关系研究</b>	<b>29</b>
5.1 研究方法	29
5.1.1 研究基本思路	29

5.1.2 实证模型设定.....	29
5.2 实证结果.....	32
5.3 本章小结.....	34
第六章全文结论.....	35
参考文献.....	36
致谢.....	40

## Contents

<b>Chapter 1 Introduction.....</b>	<b>错误！未定义书签。</b>
<b>1.1 Research Backgroud .....</b>	<b>错误！未定义书签。</b>
<b>1.2 Purpose of This Dissertation.....</b>	<b>错误！未定义书签。</b>
<b>1.3 Framework.....</b>	<b>错误！未定义书签。</b>
<b>Chapter 2 Literature Review .....</b>	<b>错误！未定义书签。</b>
<b>2.1 Research on Market Microstructure of Dealer Market .....</b>	<b>错误！未定义书签。</b>
<b>2.2 Research on Limit Order Book .....</b>	<b>错误！未定义书签。</b>
<b>2.3 Current Research in China.....</b>	<b>错误！未定义书签。</b>
<b>CHAPTER 3 Introduction on Market and Data .....</b>	<b>错误！未定义书签。</b>
<b>3.1 Market Overview .....</b>	<b>错误！未定义书签。</b>
<b>3.2 The Data .....</b>	<b>错误！未定义书签。</b>
<b>3.3 Descriptive Statistics of The Limit Order Book .....</b>	<b>错误！未定义书签。</b>
<b>Chapter 4The Limit Order Book Information and Bid-Ask Spread</b>	
<b>Relationship Study .....</b>	<b>错误！未定义书签。</b>
<b>4.1 Study Methodology.....</b>	<b>错误！未定义书签。</b>
4.1.1 Overview.....	<b>错误！未定义书签。</b>
4.1.2 Introduction on MRR Model .....	<b>错误！未定义书签。</b>
4.1.3 Model Modification: Introducing Limit Order Book Information .....	<b>错误！未定义书签。</b>
4.1.4 Estimation Methodology.....	<b>错误！未定义书签。</b>
<b>4.2 Empirical Results .....</b>	<b>19</b>
4.2.1 Descriptive Statistics of Model Variables .....	<b>错误！未定义书签。</b>
4.2.2 Analysis on Estimation Results.....	<b>22</b>
4.2.3 Analysis on Model Implied Spreads .....	<b>错误！未定义书签。</b>
<b>4.3 Conclusions of This Chapter .....</b>	<b>错误！未定义书签。</b>
<b>Chapter 5 The Limit Order Book Information and Intraday Return</b>	
<b>Relationship Study .....</b>	<b>错误！未定义书签。</b>
<b>5.1 Study Methodology.....</b>	<b>错误！未定义书签。</b>
5.1.1 Overview.....	<b>错误！未定义书签。</b>
5.1.2 The Models .....	<b>错误！未定义书签。</b>
<b>5.2 Empirical Results .....</b>	<b>错误！未定义书签。</b>

<b>5.3 Conclusions of This Chapter .....</b>	<b>错误！未定义书签。</b>
<b>Chapter 6 Conclusions.....</b>	<b>错误！未定义书签。</b>
<b>Reference.....</b>	<b>错误！未定义书签。</b>
<b>Acknowledgement.....</b>	<b>40</b>

## 第一章 引言

### 1.1 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。一些著名的研究成果，包括 Fama 和 French(1992)提出的包含小市值因子和账面价值比因子的三因素模型，Carhart(1997)在 FF 三因子模型基础上提出的动量效应模型，Barberis、Shleifer 和 Vishny(1998)基于保守性偏差和代表性偏差的 BSV 模型，以及 Daniel、Hirscheifer 和 Subramanyam(1998)基于过度自信和自归因偏差提出的 DHS 模型等，都是通过研究低频度上股价变化与不同因素之间的关系，来解释股票价格变化的原因。

而对于股票日内价格变化的研究，起步则相对较晚。这主要是由 20 世纪 80 年代以前高频数据的可得性较差，且计算工具效率不高所导致的。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 和 Mendelson(1980)、O'Hara 和 Oldfield(1986)、Madhavan 和 Smidt(1993)等人则在此领域进行了更为深入的研究。他们指出流动性因素（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 和 Milgrom(1985)、Easley 和 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。此类研究虽然较好的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)、Seppi(1997)、Foucault(1999)等人的文章。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。

## 1.2 本文研究目的和创新之处

本文的目的是研究限价指令簿信息与日内价格变化（买卖价差）及日内收益的关系。首先，作者希望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。此外，作者还对日内收益与滞后期限价指令簿信息指标建立了回归模型，以检验两者之间的关系。

本文的创新点在于：（1）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；（2）在研究指令簿信息与日内收益关系时，利用简单加和法与主成份分析法提取了滞后期指令簿总买卖压力指标，解决了不同档位信息中存在的高度多重共线性问题。

### 1.3 文章结构

文章的结构大致如下：第二部分为文献综述，第三部分介绍了我国 A 股市场微观结构及限价指令簿特征，第四部报告了研究限价指令簿信息与股票价格日内变动（买卖价差）关系的模型及实证结果，第五部分报告了研究了限价令簿信息与日内收益之间的关系的方法与实证结果，第六部分为本文结论。

## 第二章 文献综述

由于本文主要研究的限价指令簿信息与日内买卖价差及短期收益之间的关系，所以在本章中，作者首先介绍了研究买卖价差形成机制与构成的经典文献，这构成了本章的第一节；随后，作者介绍了与指令驱动市场及限价指令簿相关的研究，这构成了本章的第二节；在本章的最后部分，作者介绍了国内学者的相关研究成果。

### 2.1 传统市场微观结构研究

传统的市场微观结构研究大都是在做市商市场框架下解释日内价格形成机制及买卖价差构成。总体来看，学者们通常从信息不对称与流动性成本（做市商存货成本）两个角度来分析买卖价差。

早在 1971 年，Bagehot(1971)就提出买卖价差中应该包含信息不对称成本。他认为，市场上的交易者分为持有正确信息的知情交易者和认为自己持有信息的非知情交易者。前者希望通过做市商与后者进行交易，并依靠自己的信息优势获取利润。而作为做市商来说，他们会在与知情交易者交易的过程中蒙受一定的损失，并且会将其转嫁到非知情交易者身上，这表明买卖价差中包含了信息不对称成本。而 Glosten 和 Milgrom(1985)的研究则在此基础上引入了一个表示知情交易者比例（即信息不对称比例）的变量，且规定资产价值服从两点分布，其差值代表了资产价值本身的不确定性。他们的模型表明，买卖价差与信息不对称程度呈正比，与资产价值的不确定性呈反比。同时他们还指出，由于信息不对称的存在，即使市场上没有其他成本，买卖价差也不会消失。

而 Kyle(1985)的研究则从另一个角度解释了信息不对称与买卖价差的关系。在他的模型中，存在一个对信息具有垄断性的交易者。该交易者会在其私有信息转化为公共信息之前利用信息优势进行交易并以此获利。而做市商则通过观察订单流的情况来制定交易价格。Kyle 证明，在仅允许递交市价指令(market order)的情况下，买单数量与卖单数量的差值，即订单的不平衡性会成为资产真实价值的噪音，使得做市商错误估计资产价格，从而引起价格



偏离价值，导致买卖价差的出现。而 Holden 和 Subrahmanyam(1992)以及 Pfleiderer(1988)的研究则分别通过引入多个具有长期私有信息的竞争性知情交易者和允许部分非知情交易者具有选择入市时机的权利对 Kyle 的信息模型进行了拓展。

在买卖价差与流动性成本的关系方面，Demsetz(1968)的文章是最早研究成果之一。他在文章中指出，做市商在交易所中的角色是提供流动性，并根据市场条件的变化，被动地对价差进行调整。而买卖价差则是对他们在竞争中提供流动性的补偿。同时，Demsetz 在横截面上对价差决定因素进行了归因分析。他发现，成交量、过往收益率的波动率、交易价格及公司规模可以在横截面上解释买卖价差的大部分变化。

而 Smidt(1971)的文章则对 Demsetz 关于做市商行为的假设提出了质疑。Smidt 在文章中指出，做市商所扮演的角色并非简单而被动地提供流动性。相反地，他们会根据自己存货量的波动，主动调整价差，以补偿自己的存货风险。尽管此时做市商的主要角色仍然是流动性的供给者，但他们同时也会为了保证存货的快速周转率以及降低买单卖单的不平衡而对价格进行主动调整。Smidt 的模型表明，如果做市商的存货水平一旦低于或高于理想水平，他们对价差的调整就会使得交易价格在短期内背离证券的期望价值。

Garman(1976)的文章则在 Smidt(1971)研究的基础上，较为规范地建立了买卖价差与做市商存货水平之间关系的模型。Garman 的模型中假设，做市商的存货水平等于其初始存货水平与各期订单流的加合（其中买单符号为正，而卖单符号为负）且其资本存在上限。同时，做市商被动调整卖价与买价，以使当期的买单与卖单绝对值的期望相等（即使得当期市场出清）。由此可以推出，做市商的存货水平服从漂移项为 0 的随机游走过程。所以在做市商自有资本存在上限的情况下，当时间趋向于正无穷时，做市商破产的概率为 1。于是为了避免破产，做市商不得不根据自己的存货水平主动调整价差，以弥补持有存货带来的风险。同时，该模型还表明，提高做市商自有资本最低水平，是减少短期价格过度波动的一个有效途径。而 Stoll(1978)以及 Amihud 和 Mendelson(1980)的研究进一步说明，由于做市商在控制存货水平的过程中可能遭到一定的损失，所以即使客观的交易成本可以忽略，买卖价差依然会

因为做市商的存货控制行为而存在。

除此之外，还有许多学者在综合考虑了各种因素之后，建立了解释日内股价形成机制的模型。此类模型最早由 Roll(1984)提出，该模型通过观察成交价格的序列相关特性估计买卖价差。而 Stoll(1989)则首先从逆向选择成本、委托单处理成本及存货成本三个方面考虑买卖价差的形成，即日内价格的形成机制。他指出，做市商在市场上提供流动性且面临与知情交易者交易时可能承受的损失，所以他们需要制定足够大的买卖价差以抵消由此带来的风险及成本。但买卖价又不能太大，否则会出现无人愿意与其交易的情况。

后来的研究者在 Stoll 模型的基础上进行了一系列的拓展和改进。其中 Glosten 和 Harris(1988)提出了一个基于交易指标变量的模型来解释买卖价差。其贡献在于对交易方向进行了区分，并用 NYSE 的数据进行了实证研究。而 George、Kaul 和 Nimalendran(1991)的研究引入了时变的收益率，但同时假设存货成本为零。Madhavan、Richardson 和 Roomans(1997)则同时将信息不对称成本、流动性成本和买卖方向引入了模型，并使用 NYSE 的数据解释了日内价格及买卖价差的形成机制。

## 2.2 限价指令簿相关研究

早期的限价指令簿相关文献通常在信息对称的前提下研究指令簿对价格的影响。其中 Cohen(1981)的模型研究了在多个做市商的竞争性市场上，限价指令簿如何影响交易者的指令提交策略。其研究指出，决定市场价差的是限价指令而非市价指令，非零的价差是市场的均衡性质；价差水平则取决于交易者在市价指令和限价指令之间的选择，而选择的结果是由限价指令的执行概率所决定的。在交易不活跃的市场，限价指令的执行概率较低，价差较大。

Handa 和 Schwartz(1996)的研究分析了投资者使用限价指令交易时所面临的风险。他们在文章中认为，当负面消息到来时，限价指令成交概率较高，递交限价指令的投资者会因此蒙受损失；而当正面消息到来时，限价指令成交的概率较低，递交限价指令的投资者难以获得预期的利润。所以交易者是否提交限价指令，是由其交易对手为知情交易者的概率决定的，且对于流动

性交易者来说，提交限价指令并非占优策略。但值得注意的是，Handa 和 Schwartz 所提出的模型只是一个静态均衡模型，无法解释动态市场中投资者的行为以及投资策略的选择。

在 Handa 与 Schwartz 提出限价指令簿的静态均衡模型之后，许多学者开始进一步研究动态市场中的交易者订单投放策略。其中，Parlour(1998)的研究主要着眼于报价深度与交易者订单投放策略的关系；Foucault(1999)研究了交易者提交限价指令时所面临的订单无法成交的风险；Foucault、Kadan 和 Kandel(2005)分析了交易者耐心程度、订单投放策略与等待时间之间的关系。而 Handa、Schwartz 和 Tiwari(2003)则在模型中加入了信息不对称因素，讨论指令驱动市场中的价格形成机制。他们的研究表明，投资者异质信念、不同交易者占比及信息不对称等因素均会影响限价指令簿中的报价和买卖价差。Rosu(2009)的模型则放松了交易者订单投放行为的假设，允许投资者能够随时对订单进行修改及撤销，在连续时间框架下研究了市场价格的形成机制。他的研究表明，较高和市场活跃程度与投资者竞争水平会导致较小的买卖价差及价格冲击；价格的短期变化程度要大于其永久变化程度。Geotler, Parlour 和 Rajan(2009)的则通过数值模拟的方法证明，限价指令簿的深度(depth)具有信息含量，且较大的卖方净压力(sellingpressure)会导致价格的下降。

在实证研究方面，Sandas(2001)在 Glosten(1994)模型的基础上，考虑了时间优先原则和离散价格，使用斯德哥尔摩交易所的限价指令簿数据，对限价指令与市场流动性及逆向选择成本之间的关系进行了检验。Cao、Hansch 和 Wang(2009)的文章研究了澳大利亚股票交易所限价指令簿的信息含量。三位作者通过估计含有指令簿信息的向量误差修正模型（VECM）及多元线性回归模型，指出指令簿信息对价格及短期收益均有一定程度的解释力度。

## 2.3 国内相关研究

随着高频数据可得性的增加，我国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。雷觉铭、曾勇和李平(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。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。陈瑜(2008)基于 Sandas(2001)及 Glosten 和 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 GH 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。

## 第三章 我国股票市场微观结构及指令簿数据概述

### 3.1 我国股票市场微观结构概述

我国上海证券交易所与深圳证券交易所都是典型的指令驱动市场。投资者通过递交限价指令的方式将自己希望交易证券的价格与买卖数量提交给交易所，并进入交易所的限价指令簿。交易所计算机按照价格优先、时间优先的原则自动撮合成交。在这种交易制度下，并不存在专门提供流动性的做市商，市场上的流动性全部由限价指令簿提供。换言之，在指令驱动市场中，投资者是流动性的直接提供者，他们的行为决定了市场流动性的好坏。同时，上交所与深交所的指令簿都是（部分）公开的，所有投资者在进行订单提交决策的时候，都可以观察到限价指令簿中的量价分布情况。所以指令簿的特征可能会对投资者的递单策略产生影响。

上交所的交易时间为每天（周六、周日与法定节假日除外）的上午 9:15-11:30 和下午 13:00-15:00。其中上午 9:15-9:25 是集合竞价时段。在此阶段，交易所系统对所有有效订单进行一次集中撮合，并确定开盘价。在 9:25-9:30 之间，交易者无法进一步递交订单也无法撤单。9:30 开始，市场进入连续竞价阶段。而下午开盘前没有集合竞价，13:00 直接进入连续竞价阶段。深交所的开盘集合竞价时间与上交所相同，而连续竞价时段则为上午 9:30-11:30 和下午 13:00-14:57。下午 14:57-15:00 之间的三分钟，为深交所收盘集合竞价阶段。交易者在连续竞价阶段可以自由提交订单或撤单，其中订单数量通常以“手”为单位，每手等于 100 股，而由于配股送股等原因而产生的零股，则可以单独交易；订单价格以“元”为单位，且最小变化单位为 0.01 元。

此外值得注意的是，我国沪深股市均采用了涨跌停板制度。除上市首日的新股外，其他非 ST、\*ST 股票每日的涨跌幅相对前日收盘价不得超过 10%，ST 与 \*ST 的股票涨跌幅相对前日收盘价不得超过 5%。涨跌幅超过涨跌幅限制的指令均为视为无效指令，交易所计算机对其进行自动撤单处理。为了保持数据涨跌幅的一致性，本文的样本并未纳入 ST 与 \*ST 股票。

### 3.2 数据说明

本文所使用的是来自万得(wind)高频数据库的 A 股分笔高频数据。其中的字段包括股票代码、每笔交易的成交时间（具体到秒）、成交价格、成交量、成交额、成交方向以及实时五档最优买卖档位的价格及订单量等。本文选取了上证 50 指数与深圳成指的成份股共 90 只股票作为上海与深圳市场的代理股票，纳入实证样本，样本期长度为 2007 年 1 月 4 日到 2007 年 12 月 31 日。为了去除每天开盘与收盘阶段的噪音，本文在选取数据时参考了 Cao、Hansch 和 Wang(2009)的方法，删除了每天开盘最初 5 分钟与收盘前 5 分钟的数据。同时，由于本文所采用的数据对于涨跌停股票的指令簿数据记录是残缺的，所以在实证处理的过程中，本文删除了涨跌停阶段及量价明显错误的交易记录。

### 3.3 限价指令簿描述性统计量

对于一个限价指令簿，研究者们所关心的变量主要有以下几个：每笔成交量、买卖价差、各档位对应的订单量以及报价差异。

其中每笔成交量(volume)和各档位所对应的订单量（本文使用  $Q_{buy1} \sim Q_{buy5}$  表示买一到买五档对应的订单量， $Q_{sell1} \sim Q_{sell5}$  表示卖一到卖五档对应的订单量）反映了该时刻市场的活跃程度，该变量可直接从数据库中得到。

买卖价差(spread)反映了该时刻最优买价与最优卖价的差距，其中包含了市场摩擦所带来的成本。通常来说，买卖价差可表示为绝对价差或相对价差。其中绝对价差的定义为最优卖价与最优买价之间的绝对差距，即：

$$ABS_{spread} = P_{sell1} - P_{buy1} \quad (3.1)$$

但值得注意的是，由于样本内股票的价格相差很大，所以其平均绝对价差的水平也会有较大差异。所以本章选用相对价差作为买卖价差的代理指标，其定义为最优卖价与最优买价之间的绝对差距除以中间价，即：

$$RLTV_{spread} = \frac{2 \times (P_{sell1} - P_{buy1})}{P_{sell1} + P_{buy1}} \quad (3.2)$$

各档间的报价差异则反映了整个指令簿中价格的离散程度。本文将其定义为相邻两档买（卖）价格的差距除以其平均值，即：

$$RLTVbuygap_i = \frac{2 \times (Pbuy_{i+1} - Pbuy_i)}{(Pbuy_{i+1} + Pbuy_i)} \quad (3.3)$$

$$RLTVsellgap_i = \frac{2 \times (Psell_{i+1} - Psell_i)}{(Psell_{i+1} + Psell_i)} \quad (3.4)$$

以上各变量的描述性统计量如表 1 所示：

表 1 限价指令簿变量描述性统计量

变量名称	均值	标准差	最小值	最大值
Volume	16439.53	68050.24	1	65665075
Qbuy1	36185.33	175273.3	1	22070221
Qbuy2	54880.38	283265.5	1	23845010
Qbuy3	66802.44	368396.4	1	23854010
Qbuy4	61665.17	330338.2	1	25357350
Qbuy5	66693.56	351259.8	1	23631420
Qsell1	38834.48	189039.7	1	19406681
Qsell2	53272.02	306024.8	1	19969952
Qsell3	66494.98	396007.5	1	24180323
Qsell4	61641.2	354976.3	1	25030923
Qsell5	66927.86	374024.6	1	51880762
RLTVspread	0.001254	0.001175	0.000043	0.11027
RLTVbuygap1	0.000815	0.000813	0.000044	0.082024
RLTVbuygap2	0.000725	0.000649	0.000044	0.066116
RLTVbuygap3	0.000703	0.000604	0.000044	0.05726
RLTVbuygap4	0.000698	0.000592	0.000044	0.066225
RLTVsellgap1	0.000872	0.000886	0.000043	0.068966
RLTVsellgap2	0.000768	0.000696	0.000043	0.044834
RLTVsellgap3	0.000741	0.00065	0.000043	0.048689
RLTVsellgap4	0.000729	0.000627	0.000043	0.04381

通过观察表 1 中限价指令簿变量的描述性统计量，可以得出以下结论：

综合表 1 所列出的四个统计量来看，各变量的均值与最大最小值的差异明显，且标准差较大，说明各变量的变化幅度很大。作者认为导致这一现象的原因可能是不同股票的价格、可流通股本总量和交易活跃程度差异较为明显，所以仅分析各统计量的绝对值意义不大。更值得我们关注的是各变量均值在不同报价档位间的变化趋势，即整个指令簿的形状。

从表 1 可以看出，各档位所对应的订单量的均值皆在单笔成交量均值的 2 倍以上，这说明除去交易信息以外，限价指令簿可能蕴含了更为丰富的信息。同时，随着档位远离最优报价，其对应的订单数量呈上升的趋势，这说明投资者对股价的估值并非集中在最优报价档位上，他们可能更希望利用自己的私有信息及限价指令的投放来获取利润。这也进一步印证了作者关于限价指令簿信息含量的预测。另外值得注意的是，随着档位远离最优报价，相邻两档之间的报价差异呈递减趋势，且均小于相对价差。这说明随着档位远离最优报价，报价间的离散程度呈减弱趋势。图 1 与图 2 分别刻画了平均订单量与相对价格差距随档位变化的趋势。

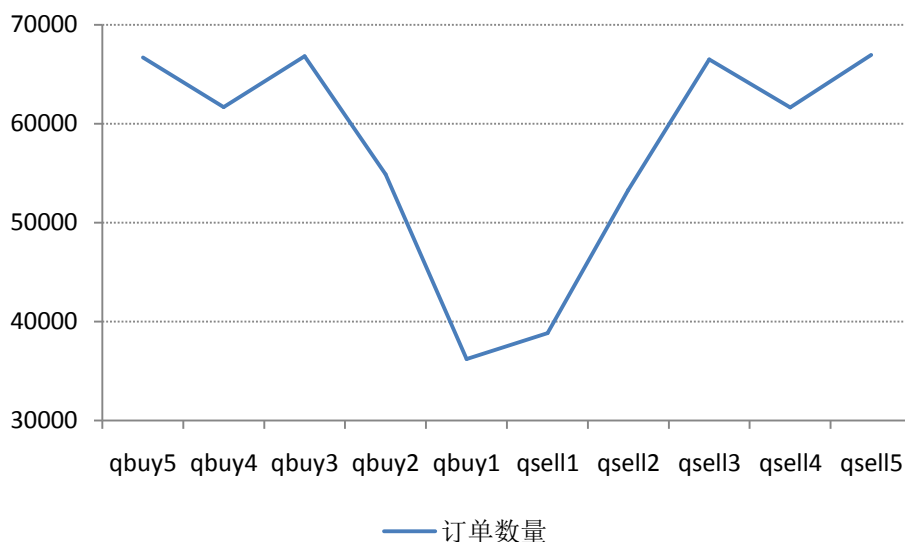


图 1 各档位订单量变化趋势



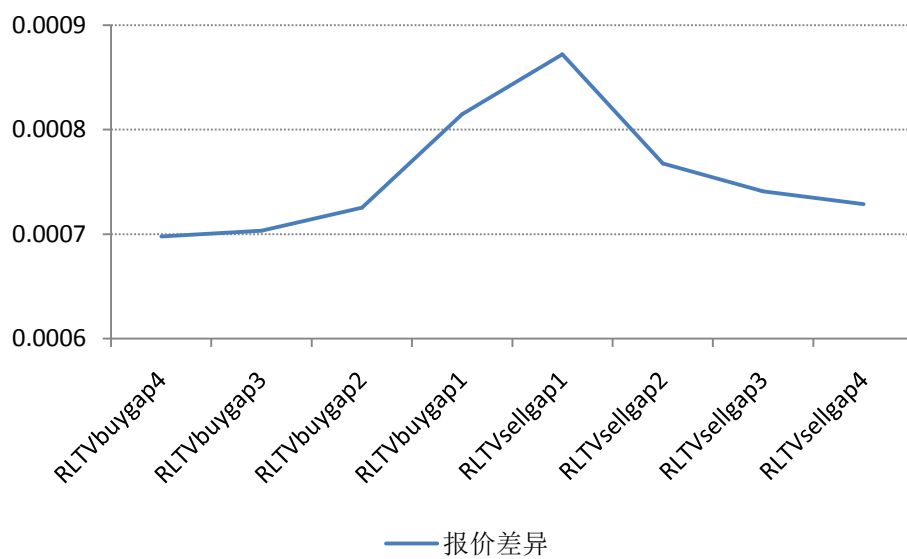


图 2 各档位相对价格差距变化趋势

通过以上分析，读者可以对样本股票限价指令簿特征有一个直观的认识。本文接下来两章实证研究中所使用的各种指标，均是基于指令簿中的价格与订单量来构建的。

## 第四章 限价指令簿信息与买卖价差成分关系研究

### 4.1 研究方法

#### 4.1.1 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 和 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。

#### 4.1.2 MRR 模型概述

MRR 模型是 Madhavan、Richardson 和 Roomans(1997)三位学者提出的用来研究日内股价形成机制的模型，该模型由 Glosten 和 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<sup>①</sup>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 和 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，我们可以得到做市商对当期股票潜在价值的判断为：

$$\mu_t = \mu_{t-1} + \theta(x_t - E[x_t | x_{t-1}]) + \varepsilon_t \quad (4.1)$$

其中  $\mu_t$  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； $\varepsilon_t$  表示新的公共信息，是一个均值为零的白噪声(Whitenoise)

<sup>①</sup>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。

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，即：

$$\begin{aligned} p_t &= \mu_t + \phi x_t + \xi_t \\ &= \mu_{t-1} + \theta(x_t - E[x_t | x_{t-1}]) + \phi x_t + \varepsilon_t + \xi_t \end{aligned} \quad (4.2)$$

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

由式 (4.2) 可以看出，股票交易价格与其上期内在价值的偏离主要由  $\phi x_t$  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 和 Roomans(1997)还根据模型推导出了隐含价差，其期望值等于  $2(\theta + \phi)$ 。

#### 4.1.3 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 ( $\theta$ ) 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 和 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 和 Saar(2005) 的文章指出，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 (pickoff) 的风险，所以提交限价指令是其最优策略。同时，

Goettler、Parlour 和 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,说明信息交易者所占比重越大,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  $\theta$  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。于是我们可以将订单簿中的订单总量作为信息不对称参数的状态变量,本文将形式设定为:

$$\theta_t = \delta + \frac{\kappa}{q_t} \quad (4.3)$$

其中  $q_t$  为当期订单簿中的订单总量, $\kappa$  为信息不对称成本参数对订单总量的反映程度, $\delta$  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 ( $\phi$ ) 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。相应的,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  $\phi$  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 和 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:

$$WP = \left( \frac{\sum_{j=1}^5 (Q_{buy_j} P_{buy_j} + Q_{sell_j} P_{sell_j})}{\sum_{j=1}^5 (Q_{buy_j} + Q_{sell_j})} \right) / MIDprice - 1 \quad (4.4)$$

其中  $Q_{buy_j}$  和  $P_{buy_j}$  ( $i=1, \dots, 5$ ) 分别表示第 1 到 5 档买单的量价,  $Q_{sell_j}$  和  $P_{sell_j}$  ( $i=1, \dots, 5$ ) 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， $\phi_t$  是 WP 的减函数；而交易为卖单时， $\phi_t$  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  $\phi_t$  的具体函数形式中加入了虚拟变量，将其设定为：

$$\phi_t = \gamma + \beta_1 WP_t + \beta_2 I_{2,t} WP_t \quad (4.5)$$

其中  $\gamma$  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  $\beta_1$ 。特别的，当买卖压力相等时， $\phi_t$  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：

$$E(WP_t) = E(WP_t | I_2 = 1) = E(WP_t | I_2 = 0) = 0 \quad (4.6)$$

则  $\phi_t$  的无条件均值可写为：

$$\begin{aligned} E(\phi_t) &= E(\phi_t | I_2 = 0)P(I_2 = 0) + E(\phi_t | I_2 = 1)P(I_2 = 1) \\ &= E(\gamma + \beta_1 WP_t | I_2 = 0)P(I_2 = 0) + E(\gamma + (\beta_1 + \beta_2) WP_t | I_2 = 1)P(I_2 = 1) \\ &= \gamma + \beta_1 E(WP_t | I_2 = 0)P(I_2 = 0) + (\beta_1 + \beta_2) E(WP_t | I_2 = 1)P(I_2 = 1) \\ &= \gamma \end{aligned} \quad (4.7)$$

即  $\phi_t$  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。

#### 4.1.4 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：

$$p_t = \mu_{t-1} + \left(\delta + \frac{\kappa}{q_t}\right)(x_t - E[x_t | x_{t-1}]) + (\gamma + \beta_1 WP_t + I_{2,t} \beta_2 WP_t) x_t + \varepsilon_t + \xi_t \quad (4.8)$$

Madhavan、Richardson 和 Roomans(1997)在其文章中证明， $E[x_t | x_{t-1}] = \rho x_{t-1}$ ，

其中  $\rho$  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：

$$\begin{aligned} \Delta p_t = & \alpha + (\gamma + \delta)x_t - (\gamma + \delta\rho)x_{t-1} + \kappa \frac{x_t}{q_t} - \kappa\rho \frac{x_{t-1}}{q_t} + \beta_1(WP_t x_t - WP_{t-1} x_{t-1}) \\ & + \beta_2(I_{2,t} WP_t x_t - I_{2,t-1} WP_{t-1} x_{t-1}) + \varepsilon_t + \xi_t - \xi_{t-1} \end{aligned} \quad (4.9)$$

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

##### (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  $\varphi$  为待估参数向量，样本矩条件为  $\varphi$  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：

$$\hat{\phi}_{GMM} = \arg \min_{\hat{\phi}} \{g_T(\hat{\phi})' W_T g_T(\hat{\phi})\} \quad (4.10)$$

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， $W_T$  被设计为样本据方差-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：

$$E \begin{pmatrix} x_t x_{t-1} - x_t^2 \rho \\ u \\ u x_t \\ u x_{t-1} \\ u \frac{x_t}{q_t} \\ u \frac{x_{t-1}}{q_t} \\ u(WP_t x_t - WP_{t-1} x_{t-1}) \\ u(I_{2,t} WP_t x_t - I_{2,t-1} WP_{t-1} x_{t-1}) \end{pmatrix} = 0 \quad (4.11)$$

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。

## 4.2 实证结果

在具体的实证过程中，本文采用了 Madhaven、Richardson 和 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。

#### 4.2.1 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 2 报告了全天未分段的各变量描述性统计量, 而表 3 则报告了日内不同时段各变量均值的分布情况。

表 2 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
$\Delta p$ 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 2 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  $\Delta p$  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  $x_t$  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。

同样的，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，订单簿内的买卖压力较为平衡。这也从侧面印证了 4.1 节中我们关于  $E(WP_t) = 0$  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  $\Delta p$  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， $\frac{1}{q_t}$  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 3 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
$x_t$	0.023351	0.020267	0.037243	0.024857
$WP_t$	-7.6E-06	5.37E-06	7.16E-05	0.000132
$ WP_t $	0.001285	0.001223	0.001195	0.001162
$\frac{1}{q_t}$	1.21E-05	1.1E-05	1.07E-05	1.03E-05
<i>Spread</i>	0.037606	0.032908	0.032023	0.030897
<i>RLTVspread</i>	0.001473	0.001331	0.001292	0.001252
$\Delta p$ 的方差	0.005333	0.010879	0.002506	0.002262

注：表 3 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 3 可以看出，除  $WP_t$  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  $WP_t$  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)以及穆启国和吴冲锋(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。

#### 4.2.2 模型参数估计结果分析

表 4-表 7 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。

表 4 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
$\alpha$	4.21E-05	0.001169	-0.00034	0.000221	1.802397	35.56%
$\beta_1$	-1.23996	0.534442	-4.93756	1.210875	-9.71468*	74.44%
$\beta_2$	2.41991	9.179242	-0.90672	2.251924	10.19452*	80.00%
$\delta$	0.004002	0.019003	0.000211	0.00326	11.6491*	100.00%
$\gamma$	0.008299	0.050175	0.003874	0.00612	12.86529*	100.00%
$\kappa$	128.7896	1097.837	-30.2367	145.3006	8.408813*	91.11%
$\rho$	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% 的水平下显著异于零。

表 5 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
$\alpha$	-2E-05	0.000761	-0.0004	0.000186	-0.99513	47.62%
$\beta_1$	-0.90808	15.65263	-15.4066	3.983954	-2.12603**	66.67%
$\beta_2$	1.638378	27.1528	-29.1475	7.79487	1.960492**	67.86%
$\delta$	0.00356	0.013379	0.000318	0.002452	13.54445*	97.62%
$\gamma$	0.008469	0.023663	0.004125	0.004341	18.19605*	100.00%
$\kappa$	99.71598	488.5083	-290.048	107.759	8.631195*	89.29%
$\rho$	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% 的显著性水平下显著异于零。

注 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  $\kappa$  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 6 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
$\alpha$	7.39E-06	0.000713	-0.00042	0.000139	0.502991	28.89%
$\beta_1$	-1.5025	0.490327	-21.2632	2.570074	-5.54613*	73.33%
$\beta_2$	2.971185	37.11676	-1.61514	4.737648	5.949604*	75.56%
$\delta$	0.003374	0.014669	0.00015	0.002765	11.57656*	100.00%
$\gamma$	0.008299	0.050175	0.003874	0.00612	12.86529*	100.00%
$\kappa$	130.0033	2268.266	-423.093	255.8428	4.820615*	90.00%
$\rho$	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% 的水平下显著异于零。

表 7 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
$\alpha$	-9.2E-05	0.00021	-0.00064	0.000125	-6.96511*	35.56%
$\beta_1$	-1.63812	0.602452	-14.0521	2.048226	-7.58733*	76.67%
$\beta_2$	3.105877	26.01493	-1.68187	3.942274	7.474097*	74.44%
$\delta$	0.003024	0.012304	5.48E-05	0.002405	11.92851*	100.00%
$\gamma$	0.008077	0.042342	0.003934	0.005417	14.14495*	100.00%
$\kappa$	147.622	2463.499	-896.678	294.6275	4.753341*	91.11%
$\rho$	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%的水平下显著异于零。

通过分析表 4-表 7，我们可以得到以下结论：

1、截距项  $\alpha$  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  $\Delta p$  有较好的解释能力。

2、 $WP_i$  前面的系数  $\beta_1$  均值为负， $\beta_2$  均值为正，且  $\beta_2$  的绝对值大于  $\beta_1$ 。这与本文在 4.1.3 节所做出的预测一致，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， $\beta_1$  与  $\beta_2$  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  $\beta_1$  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%， $\beta_2$

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， $\phi_t = \gamma + \beta_1 WP_t$ ；当交易为卖单时， $\phi_t = \gamma + (\beta_1 + \beta_2) WP_t$ 。根据表 4-表 7 所示的回归结果以及表 3 所示的描述性统计量可以看出， $\beta_1$  的均值为 -1.32， $(\beta_1 + \beta_2)$  的均值为 1.21，按照  $|WP_t| 0.00128$  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  $\gamma$  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、 $\frac{1}{q_t}$  前面的系数  $\kappa$  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 4.1.3 节中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， $\kappa$  符号正确且显著的比例很高（分别为 91.11%、89.29%、90.00%和 91.11%），这进一步印证了本文 4.1.3 节中模型拓展部分假设的正确性。从数值大小上来看， $\kappa$  在四个时段中的均值为 126.53，而  $\frac{1}{q_t}$  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  $\delta$  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  $\delta$  与  $\gamma$  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  $\delta$  不显著， $\gamma$  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  $\rho$  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 和 Roomans(1997)的文章中证明，买单（卖单）连续出现的概率，即  $E(x_t = x_{t-1} | x_{t-1})$  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。

### 4.2.3 模型隐含价差分析

根据 4.1.2 中的介绍，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  $2[E(\theta_t) + E(\phi_t)]$ 。同时，根据 4.1.3 节中推导出的  $E(\phi_t) = \gamma$ ，可推得隐含价差为  $2(\kappa \times E[\frac{1}{q_t}] + \delta + \gamma)$ 。如果用样本矩替代总体矩，隐含价差可表示为  $2[\kappa \times \overline{(1/q_t)} + \delta + \gamma]$ 。

表 8 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 8 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 3 相同

由表 8 可以看出，隐含价差与真实绝对价差相比略小，这与 Madhavan、Richardson 和 Roomans(1997)文章中的结论一致。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 3 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。

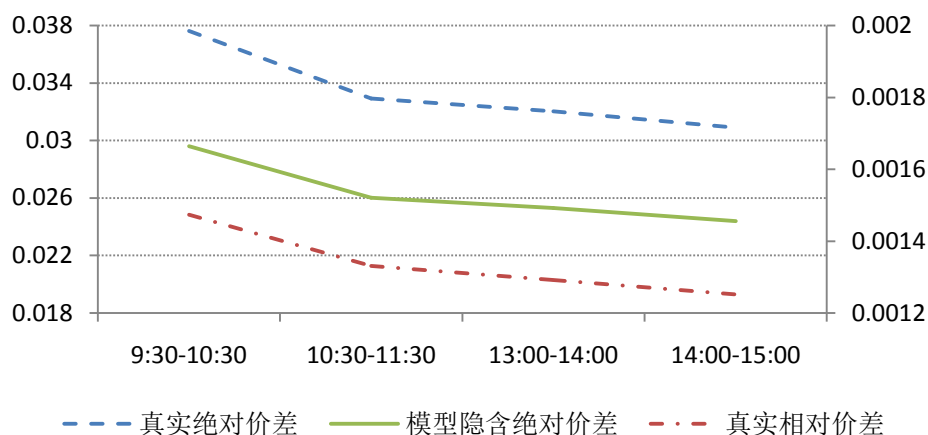


图 3 隐含价差、真实绝对价差及真实相对价差均值日内走势

注：图 3 中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。

### 4.3 本章小结

本章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。

本章实证结果表明，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。

在讨论完限价指令簿信息与日内买卖价差成分之间的关系后，作者将会在下一章中进一步研究限价指令簿信息与日内收益的之间关系。



## 第五章 限价指令簿信息与日内收益关系研究

### 5.1 研究方法

#### 5.1.1 研究基本思路

众多研究表明，限价指令簿中的量价信息会对股票日内收益产生影响。Harris(1990)的研究中将递交限价指令的交易者区分为成本驱动交易者(pre-committedtraders)与价值驱动交易者(value-motivatedtraders)，并认为两者递交的订单均传递出与未来股价变化相关的信息。Harris 在文章中指出，当买方（卖方）订单量大于卖方（买方）订单量时，未来的短期股价将会上升（下降），即收益为正（负）。而 Huang 和 Stoll(1994)，Chordia、Roll 和 Subrahmanyam(2002)以及 Chordia 和 Subrahmanyam(2004)分别通过实证分析，证明最优报价档订单深度不平衡程度与短期收益有关。另外，Cao、Hansch 和 Wang(2009)则利用澳大利亚市场限价指令簿信息研究了短期收益与订单簿中各档对应的买卖压力及价格冲击指标的关系。本章主要参考了 Cao、Hansch 和 Wang(2009)的实证思路，研究了中国大陆股票市场限价指令簿信息与股票日内收益的关系。

#### 5.1.2 实证模型设定

Cao、Hansch 和 Wang(2009)的文章指出，订单簿中各档表现出的滞后期买卖压力，会对当期短期收益产生影响。三位作者将每天的交易时段分为若干个长度为 5 分钟的子样本，并计算子样本期内的收益率及反映指令簿信息含量的各种指标。他们将 5 分钟收益定义为  $r_t = (p_t - p_{t-1}) / p_t$ ，其中  $p_t$  为当前股价， $p_{t-1}$  为 5 分钟前的股价。而指令簿各档对应的买卖压力定义为  $NSR_j = \frac{(Q_{sell_j} - Q_{buy_j})}{Q_{sell_j} + Q_{buy_j}}$ ，

其中  $Q_{sell_j}$  与  $Q_{buy_j}$  分别表示 5 分钟内第  $j$  档卖单与买单的总量。下标  $j$  表示订单所在档位：最优报价档为 1， $j$  随档位远离最优的程度递增。当  $NSR$  为正时，第  $j$  档表现为净卖出压力； $NSR$  为负时，第  $j$  档表现为净买入压力。Cao、Hansch

和 Wang 在实证中共进行了 10 个回归，将第 1 档到第 10 档的 NSR 逐次加入回归方程，检验其显著性。他们发现，短期收益与 NSR 呈负相关关系，即上一期卖出压力越大，当期收益率越低，且随着纳入回归方程指标数量的增加，方程调整后 R 方依次递增；同时，随着档位远离最优报价，NSR 在个股层面的显著性呈递减趋势，由此认为近档位的买卖压力对短期收益的解释能力更强。但本文作者认为，三位作者得出的第二个结论，即近档位的买卖压力对短期收益的解释能力比远档位更强，可能并非由远档位订单信息含量较差所导致。引起这一现象的原因可能是不同指标之间存在多重共线性(multicollinearity)。

为了证实以上猜测，作者分别计算了样本内股票各档买卖压力的相关系数矩阵。计算结果显示，各档买卖压力的相关系数从 20%-80%不等，且大多集中在 50%-70%的区间。这说明各档买卖压力存在较强的正向相关关系，表现为高度多重共线性。该结论证实了作者之前猜测的合理性。

为了避免多重共线性对回归系数显著性的影响，本文采用了两种方法对不同档位的买卖压力指标进行处理。

其中第一种方法较为直观，即简单地将所有档位的信息集合在一起，构造指

标  $NSRT = \frac{\sum_{j=1}^5 Q_{sell_j} - \sum_{j=1}^5 Q_{buy_j}}{\sum_{j=1}^5 (Q_{sell_j} + Q_{buy_j})}$ ，用来表示整个限价指令簿所表现出的买卖压力。

与之前的 NSR 指标类似：当 NSRT 为正时，指令簿呈现净卖出压力；NSRT 为负时，指令簿呈现净买入压力。作者关于短期收益与 NSRT 之间关系的预测与 Cao、Hansch 和 Wang(2009)相同，即短期收益与 NSRT 负相关。

除了直接加和各档位的信息之外，本文还采用了主成份分析法(PrincipalComponentAnalysis,PCA)来提取影响各档买卖压力共同变化的因素（即主成份）。在此步处理中，作者将原本存在高度多重共线性的 5 个指标转化成了 5 个新的且相互正交的主成份指标。经过观察发现，各档买卖压力对第一主成份变化的反应系数均为正，且大小相近。作者由此认为第一主成份可以较好解释各档买卖压力的水平变化，并将其视为指令簿总体买卖压力的代理指标，作为解释变量之一纳入下文的回归方程中。

另外，本文还将指令簿中的订单总数  $q_t$  纳入了回归方程。作者认为，指令簿

中的订单总数除了可以衡量市场信息不对称程度之外，还可以作为整个市场活跃度的指标。将此指标纳入回归方程，是为了检验市场整体活跃程度（而非单纯的交易活跃程度）对日内收益的影响。

为了控制短期收益受自身相关性的影响，本文对短期收益建立了 AR (1) 模型并取其残差，作为回归中的被解释变量；为了排除交易层面对短期收益的影响，本文参照 Cao、Hansch 和 Wang 的做法，将五分钟的平均买卖价差作为控制变量纳入回归方程。

采用 NSRT 作为解释变量的回归方程可写为方程(5.1)所表示的形式：

$$\varepsilon_t = \alpha + \beta_{SPR} Spr_{t-1} + \beta_1 q_{t-1} + \beta_2 NSRT_{t-1} + v_t \quad (5.1)$$

其中  $\varepsilon_t$  为对短期收益建立 AR (1) 模型所获得的残差， $Spr_{t-1}$  为上期平均买卖价差， $q_{t-1}$  为上期订单簿中所含的订单总量， $NSRT_{t-1}$  为上期指令簿买卖压力， $v_t$  为模型扰动项。

采用从各档买卖压力指标中提取出的第一主成份作为解释变量的回归方程可写为方程(5.2)所表示的形式：

$$\varepsilon_t = \alpha' + \beta'_{SPR} Spr_{t-1} + \beta'_1 q_{t-1} + \beta'_2 PRIN1_{t-1} + v'_t \quad (5.2)$$

其中  $PRIN1_{t-1}$  表示从上一期各档买卖压力指标中提取出的第一主成份，其他变量含义与方程 (5.1) 相同。

此外值得注意的是，在指令驱动市场中，最优买卖档往往具有一定的市价指令属性，希望迅速成交的投资者会倾向于直接在最优买卖档投放订单。所以为了进一步检验排除市价指令因素的“纯限价指令簿”信息对短期收益的影响，本文设计了以下回归模型：

$$\varepsilon_t = \alpha'' + \beta''_{SPR} Spr_{t-1} + \beta_3 q_{1,t-1} + \beta_4 NSR_{1,t-1} + \beta_5 q_{2-5,t-1} + \beta_6 NSR_{2-5,t-1} + v_t'' \quad (5.3)$$

其中  $\varepsilon_t$  和  $Spr_{t-1}$  的含义与之前相同， $q_{1,t-1}$  为上期最优档订单总量， $q_{2-5,t-1}$  为上期 2~5 档订单总量， $NSR_{1,t-1}$  为上期最优档买卖压力， $NSR_{2-5,t-1}$  为上期 2~5 档呈现出的买卖压力， $v_t''$  为模型扰动项。

本章实证研究中使用的股票样本与第四章相同，而日内子样本长度为 5 分钟。

## 5.2 实证结果

表 9 和表 10 分别报告了 5.1 节中 (5.1) 和 (5.2) 两个回归方程的估计结果。

表 9 方程 (5.1) 估计结果

参数	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
$\alpha$	-0.00024	0.002245	-0.00227	0.000502	-4.5896*	35.56%
$\beta_1$	0.003256	0.025975	-0.01785	0.004816	6.413007*	83.33%
$\beta_2$	-0.00063	0.000566	-0.00229	0.000446	-13.4201*	76.67%
$\beta_{SPR}$	-0.00769	0.07528	-0.19674	0.037363	-1.9537**	20.00%

注：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%的显著性水平下显著异于零。

表 10 方程 (5.2) 估计结果

参数	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
$\alpha'$	-0.00024	0.002632	-0.00227	0.000522	-4.43318*	38.89%
$\beta_1'$	0.003226	0.025251	-0.01834	0.004779	6.403962*	83.33%
$\beta_2'$	-0.00015	0.000109	-0.0004	0.000112	-12.3596*	75.56%
$\beta_{SPR}'$	-0.00769	0.075667	-0.19633	0.037267	-1.95761**	21.11%

注：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%的显著性水平下显著异于零。

通过分析表 9 和表 10，我们可以得出以下结论：

1、 $\beta_1$ 和 $\beta_1'$ 的均值皆在横截面上显著为正，且在个股层面符号正确且显著的比例较高（两组回归中均为 83.33%）。该结果说明订单簿表现出的滞后期市场活

跃程度对当期收益有显著影响，市场活跃程度越高，未来的短期收益越高。

2、 $\beta_2$  与  $\beta_2'$  的均值在横截面上亦显著为负，而在个股中符号正确且显著的比例分别为 76.67% 与 75.56%。作者认为，这样的结果可以说明两方面的问题。首先，两组回归所得到的买卖总压力指标前面的系数估计结果十分接近，说明 5.1 节中所提出的两种信息整理方法是合理且可以相互支持的，简单加和法与主成份分析法均可以较好地反映指令簿各档买卖压力共同的水平变动情况；其次，两组回归结果都表明，滞后期订单簿所表现出的卖方压力越大，股票当期收益越低。这与作者在 5.1 节中的预测是一致的。

表 10 则报告了 5.1 节中方程 (5.3) 的回归结果。

表 11 方程 (5.3) 估计结果

参数	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
$\alpha''$	-0.00025	0.002258	-0.00212	0.000496	-4.86048*	38.89%
$\beta_3$	0.011494	0.064103	-0.02027	0.01184	9.209151*	57.78%
$\beta_4$	0.001509	0.022409	-0.01782	0.0042	3.408617*	20.00%
$\beta_5$	3.95E-05	0.001022	-0.00055	0.000296	1.265061	13.33%
$\beta_6$	-0.00064	0.000272	-0.00178	0.00036	-16.8689*	81.11%
$\beta_{spr}''$	-0.00744	0.074475	-0.18463	0.036529	-1.93349**	22.22%

注：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% 的显著性水平下显著异于零。

通过观察表 10，作者发现将最优及非最优买卖档位订单量进行区分后，最优买卖档的滞后期订单数量对未来短期收益有更好的解释能力（ $\beta_3$  在个股中符号正确且显著的比例为 57.78%，大于  $\beta_4$  20% 的显著比例）。这说明最优买卖档位的订单量更好地体现了市场的活跃程度，而由于非最优档的订单被立即执行的概

率较低，对市场活跃程度的贡献也相对较低。而另一方面，非最优档位所体现出的滞后期买卖压力对未来收益的解释能力强于最优档位买卖压力（ $\beta_5$  仅在 13.33% 的个数中显著，远小于  $\beta_6$  81.11% 的比例）。作者认为这表明非最优档体现出的买卖压力包含了更多关于未来股价走势的信息，而最优档订单中包含了部分流动性交易者的订单，所以并不能完全反映交易者对未来股价走势的估计。

### 5.3 本章小结

作者在本章中研究了滞后期指令簿信息与日内收益之间的关系。本章实证结果表明，简单加和与主成份分析两种方法均可以较好地解决限价指令簿内各档买卖压力指标所存在的高度多重共线性问题，并提取代表指令簿总体买卖压力的相应指标。同时，本章实证结果还表明，限价指令簿内的订单总量与整体买卖压力均对下一期的短期收益具有一定的预测作用。其中指令簿中的订单总量（市场活跃程度）与未来短期收益呈正相关关系，且最优买卖档位的订单量比非最优档位的订单量具有更强的解释能力。而净卖出压力则与未来短期收益呈负相关关系，且非最优档位表现出的买卖压力比最优档位表现出的买卖压力具有更强的解释力度。

## 第六章 全文结论

本文利用上证 50 与深圳成指成份股 2007 年全年高频数据研究了我国 A 股市场限价指令簿信息与日内买卖价差及短期收益之间的关系。

首先,作者将从限价指令簿中提取出的信息指作为状态变量加入了被广泛应用于买卖价差研究的 MRR 模型中,使原本的常参数模型变成了时变参数模型,并借此分析了限价指令簿信息对买卖价差构成成分的影响。实证结果显示,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;而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。同时,本文实证结果还表明,新加入的状态变量并没有破坏原模型的结构,原本参数中的非时变部分仍然高度显著,且绝对大小合理。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高度吻合,说明模型可以较好地解释和反映我国 A 股市场买卖价差的性质。

其次,作者参考 Cao、Hansch 和 Wang(2009)的方法,通过多元回归的方法检验了限价指令簿信息(买卖压力与订单总数)与股票日内收益之间的关系。与 Cao、Hansch 和 Wang 三位作者不同的是,为了排除指令簿各档位信息存在的高度多重共线性问题,本文通过简单加和及主成份分析两种方法整合提取指令簿的总体信息(原文使用的则是分档信息),来检验指令簿总体信息指标与短期收益之间的关系。实证结果表明,通过简单加和与通过主成份分析提取出的买卖压力指标具有高度的一致性,两者均对短期收益有显著影响:上期卖出净压力越大,当期收益越低;而代表市场活跃程度的订单总量指标同样对短期收益影响显著:上期订单量越大,当期收益越高。同时,作者还检验了最优报价档位信息与非最优报价档位信息对短期收益的不同影响。实证结果显示,非最优档位的买卖压力比最优档位的买卖压力对短期收益的解释能力更强,而最优档位的订单量则比非最优档位的订单量更具解释能力。

## [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》, 电子科技大学硕士学位论文.
- 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》, 清华大学博士论文.
- 雷觉铭和曾勇, 2006, 《上海股市买卖价差成分分析》, 《系统工程》 2006 年第六期 74-80 页.
- 穆启国和吴冲锋, 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 13 期 239-243 页.
- 屈文洲和吴世农, 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》第一期 56-63 页.
- 孙培源和施东辉, 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》第四期 69-72 页.
- 杨之曙, 2003, 《上海股市日内流动性——深度变化实证研究》金融研究第六期 25-37 页.
- Admati, A. R., and P. Pfleiderer, 1988, "A theory of intraday patterns: Volume and price variability", *Review of Financial Studies* 1, PP. 3-40.
- Amihud, Y., and H. Mendelson, 1980, "Dealership market: Market-making with inventory", *Journal of financial economics* 8, PP. 31-53.
- Bagehot, W., 1971, "The only game in town", *Financial Analysts Journal* March-April, PP. 12-14.
- Barberis, N., A. Shleifer, and R. Vishny, 1998, "A model of investor sentiment", *Journal of financial economics* 49, PP. 307-343.
- 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, PP. 165-199.
- Cao, C., O. Hansch, and X. Wang, 2009, "The information content of an open limitorder book", *Journal of Futures Markets* 29, PP. 16-41.
- Carhart, M. M., 1997, "On persistence in mutual fund performance", *the Journal of Finance* 52, PP. 57-82.
- Chordia, T., R. Roll, and A. Subrahmanyam, 2002, "Order imbalance, liquidity, and market returns", *Journal of financial economics* 65, PP. 111-130.



- Chordia, T., and A. Subrahmanyam, 2004, "Order imbalance and individual stock returns: Theory and evidence", *Journal of financial economics* 72, PP. 485-518.
- Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1981, "Transaction costs, order placement strategy, and existence of the bid-ask spread", *The Journal of Political Economy* 89, PP. 287-305.
- Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, "Investor psychology and security market under-and overreactions", *the Journal of Finance* 53, PP. 1839-1885.
- Demsetz, H., 1968, "The cost of transacting", *The Quarterly Journal of Economics* 82, PP. 33-53.
- Easley, D., and M. O'Hara, 1987, "Price, trade size, and information in securities markets", *Journal of financial economics* 19, PP. 69-90.
- Fama, E. F., and K. R. French, 1992, "The cross-section of expected stock returns", *the Journal of Finance* 47, PP. 427-465.
- Foucault, T., 1999, "Order flow composition and trading costs in a dynamic limit order market", *Journal of Financial markets* 2, PP. 99-134.
- Foucault, T., O. Kadan, and E. Kandel, 2005, "Limit order book as a market for liquidity", *Review of Financial Studies* 18, PP. 1171-1217.
- Garman, M. B., 1976, "Market microstructure", *Journal of financial economics* 3, PP. 257-275.
- George, T. J., G. Kaul, and M. Nimalendran, 1991, "Estimation of the bid-ask spread and its components: A new approach", *Review of Financial Studies* 4, PP. 623-656.
- 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, PP. 71-100.
- Glosten, L. R., 1994, "Is the electronic open limit order book inevitable?", *the Journal of Finance* 49, PP. 1127-1161.
- Glosten, L. R., and L. E. Harris, 1988, "Estimating the components of the bid/ask spread", *Journal of financial economics* 21, PP. 123-142.
- Goettler, R. L., C. A. Parlour, and U. Rajan, 2009, "Informed traders and limit order markets", *Journal of financial economics* 93, PP. 67-87.
- Handa, P., R. Schwartz, and A. Tiwari, 2003, "Quote setting and price formation in an order driven market", *Journal of Financial markets* 6, PP. 461-489.
- Handa, P., and R. A. Schwartz, 1996, "Limit order trading", *the Journal of Finance* 51, PP.

1835-1861.

- Harris, L. 1990, "Liquidity, trading rules, and electronic trading systems", New York University. Monograph series in finance and economics, Monograph 1990(4).
- Holden, C. W., and A. Subrahmanyam, 1992, "Long-lived private information and imperfect competition", *the Journal of Finance* 47, PP. 247-270.
- Huang, R. D., and H. R. Stoll, 1994, "Market microstructure and stock return predictions", *Review of Financial Studies* 7, PP. 179-213.
- Kaniel, R., and H. Liu, 2006, "What orders do informed traders used", *Journal of Business* 79, PP. 1867-1913.
- Kyle, A. S., 1985, "Continuous auctions and insider trading", *Econometrica* 53, PP. 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, PP. 1035-1064.
- Madhavan, A., and S. Smidt, 1993, "An analysis of changes in specialist inventories and quotations", *the Journal of Finance* 48, PP. 1595-1628.
- O'Hara, M., and G. S. Oldfield, 1986, "The microeconomics of market making", *Journal of Financial and Quantitative Analysis* 21, PP. 361-376.
- Parlour, C. A., 1998, "Price dynamics in limit order markets", *Review of Financial Studies* 11, PP. 789-816.
- Rock, K., 1996, "The specialist's order book and price anomalies", *Review of Financial Studies* 9, PP. 1-20.
- Roll, R., 1984, "A simple implicit measure of the effective bid-ask spread in an efficient market", *the Journal of Finance* 39, PP. 1127-1139.
- Rosu, I., 2009, "A dynamic model of the limit order book", *Review of Financial Studies* 22, PP. 4601-4641.
- Sandas, P., 2001, "Adverse selection and competitive market making: Empirical evidence from a limit order market", *Review of Financial Studies* 14, PP. 705-734.
- Seppi, D. J., 1997, "Liquidity provision with limit orders and a strategic specialist", *Review of Financial Studies* 10, PP. 103-150.
- Smidt, S., 1971, "Which road to an efficient stock market: Free competition or regulated monopoly?", *Financial Analysts Journal* 27, PP. 18-69.

Stoll, H. R., 1978, "The supply of dealer services in securities markets", *the Journal of Finance* 33, PP. 1133-1151.

Stoll, H. R., 1989, "Inferring the components of the bid-ask spread: Theory and empirical tests", *the Journal of Finance* 44, PP. 115-134.

## 致谢

时光荏苒，自 2004 年 9 月迈入南强校门，至今已近七载。厦大的凤凰花每年开两季，一季老生去，一季新生来。没想到转眼间，自己也要跟凤凰花道别，去到一个并不熟悉的南方城市开启新生活了。

七年间，在“自强不息，止于至善”这句校训的鞭策下，我完成了从一个毛头小子到知识青年的转变。七年间，我得到了许多老师、同学和朋友的帮助，他们在学习、生活乃至为人处世方面都深深的影响了我。在这里，我要对他们表示自己深深的谢意。

我首先要感谢我的导师郑振龙教授。对郑老师的第一印象，是他在新生入门指导课上的精彩演讲。正式这次演讲，激发了我对于金融工程的兴趣与热情，并决定将其作为自己的专业。在顺利考取研究生后，我有幸成为了郑老师门下的一名学生。三年的研究生生涯里，我在学习与生活的各个方面，都受到了郑老师无微不至的关心与指导。尤其是在我论文选题遇到困难不知所措的时候，郑老师的安慰、宽容和鼓励给了我莫大的力量。而在后来的论文写作过程中，郑老师也给我提出了宝贵的指导和修改意见，使论文得以顺利完成。

同时，我还要感谢金融工程专业的陈蓉老师、陈淼鑫老师和远在新西兰的林海老师。他们是我本科及研究生学习阶段多门专业课程的任课老师。正是他们的教导，让我获得了许多宝贵的专业知识。这些知识不仅对我的论文写作有很大的帮助，而且为我将来在证券公司金融工程部的职业生涯奠定了良好的基础。

我还要感谢我同门的邓弋威师兄、刘杨树师兄和吕凯师兄，他们在我论文选题写作的过程中，都提出了许多宝贵的指导和意见；感谢同班的王路跖、吴江宏、韩露和岳清慧等同学，谢谢你们陪我度过了三年的快乐时光；感谢杨凡颖，没有你的爱与陪伴，我无法熬过找工作和写论文那段最艰难的日子，你的安慰和鼓励让我无以回报。

最后我要感谢我的父母，感谢你们的养育之恩。没有你们一直以来的支持，我绝无可能取得今天这些成绩。

最近恰逢母校 90 周年校庆，在这里对厦大说一句生日快乐！