

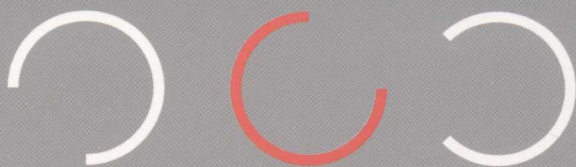
高 扬
著



金融市场

有效价差的估计方法

Estimation of the Effective Spread
on Financial Markets



清华大学出版社



高 扬
著



金融市场

有效价差的估计方法

Estimation of the Effective Spread
on Financial Markets



清华大学出版社
北 京

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

金融市场有效价差的估计方法/高扬著. —北京:清华大学出版社,2019
ISBN 978-7-302-52391-8

I. ①金… II. ①高… III. ①金融市场—研究 IV. ①F830.9

中国版本图书馆 CIP 数据核字(2019)第 038169 号

责任编辑:刘 晶

封面设计:汉风唐韵

责任校对:王凤芝

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市金元印装有限公司

经 销:全国新华书店

开 本:170mm×240mm 印 张:9.5 字 数:142 千字

版 次:2019 年 4 月第 1 版 印 次:2019 年 4 月第 1 次印刷

定 价:69.80 元

产品编号:073540-01

作者简介

高扬,2015年7月获北京大学经济学博士学位,现为北京工业大学经济与管理学院副教授,硕士生导师。目前主持1项国家自然科学基金青年项目及2项省部级项目,参与多项国家自然科学基金项目,并在国内外核心期刊上发表论文二十余篇。

研究方向为金融市场微观结构、金融时间序列等。

前言

PREFACE

流动性是金融市场微观结构研究中的一个关键问题,对于资产定价、市场有效性、企业财务以及风险管理等各类金融决策均有重要的意义,而流动性的测度则是这一类问题的基础。本书将统计学与金融热点问题相结合,研究市场微观结构中流动性的度量方法,特别是关于交易成本的度量方法及其统计推断,为流动性的应用提供经济计量学的理论基础。由于流动性是个多维度的概念,交易成本是其中重要的一个方面,度量交易成本的基本方式是买卖价差(Bid-Ask spread),而有效价差(Effective Spread)是常用的买卖价差指标,因此本书主要分四部分来分析交易成本特别是有效价差的推断方法。

(1) 首先从理论上分析比较两类有效价差估计的统计性质,即 Roll 的协方差估计(Roll,1984)及最近由 Corwin 和 Schultz(2012)提出的基于最高价和最低价得到的价差估计。与以往文献中采用估计价差与基准价差的相关系数来衡量和比较不同估计优劣表现的做法有所不同,本书通过推导并对比两种估计的偏差、均方误差及其在大样本下的性质,从而在理论上证明基于最高价和最低价的价差估计精度的确高于 Roll 的估计,并通过随机模拟分别在理想和非理想状况下进行验证。此外,借助 Bootstrap 方法给出的置信区间分析对 S&P 500 成分股票等的应用研究也从实证角度证实了 High-Low 估计优于 Roll 协方差估计的结论。

(2) 在 Roll 价格模型的基础上,本书基于价格极差的信息得到新的一阶矩条件,结合 Corwin 和 Schultz(2012)文章中价格极差的二阶矩条件,提出 5 种区别于 Corwin 和 Schultz 所提估计的新的 High-Low 矩估计,并推导出五种 High-Low 估计包括偏差、方差、相合性以及渐近正态性在内的统计性质,随后通过直接考察理论性质或者进行随机模拟研究等比较新的 High-Low 矩估计与

Corwin 和 Schultz 的 High-Low 估计的估计精度。此外,本书还提出基于上述矩条件的多个广义矩估计,并通过模拟结果验证广义矩估计的优势。

(3) 无论是 Roll 的协方差估计,还是 Corwin 和 Schultz 文中的 High-Low 估计,抑或本书第二部分提出的新的 High-Low 矩估计或者广义矩估计,其本质都是矩估计。因此本书基于 Roll 的价格模型,从对数价格极差的分布函数出发,利用其近似正态特征,提出一种有效价差的近似极大似然估计,并给出拟极大似然估计的统计性质。通过数值模拟,比较这一新的估计与 Roll 的协方差估计、贝叶斯估计以及 Corwin 和 Schultz 的 High-Low 估计在各种不同状况下的精度。模拟的结果表明,无论是理想状态还是非理想状态下,极大似然估计和 High-Low 估计的精度均高于协方差和贝叶斯估计;当波动率相对较小的时候,极大似然估计的精度优于 High-Low 估计。另外,在交易不连续的非理想情形下,极大似然估计要比 High-Low 估计更加稳健。

(4) 由于低频价差估计的方法种类较多,结合前三个方面的结果,针对基于 Roll 的价格模型得到的多种低频有效价差的估计方法,本书以中国股票和债券市场的交易数据为例,对其交易成本尤其是有效价差的估计方法进行实证比较,研究不同的有效价差估计方法对中国金融市场的适用性。

目 录

CONTENTS

第一章 导读	1
第一节 研究背景	1
第二节 本书内容及结构	5
第三节 低频估计的文献综述	6
一、Roll 的估计	7
二、贝叶斯估计	9
三、Holden 估计	9
四、High-Low 估计	11
五、LOT 估计	13
六、FHT 估计	14
七、低频价差估计的比较	15
第二章 两种有效价差估计的渐近性质	16
第一节 引言	16
第二节 理论性质	17
一、Roll 估计的渐近性质	18
二、High-Low 估计的渐近性质	18
三、两种估计的比较	20
第三节 模拟研究	21
一、Roll 估计与 High-Low 估计的预测误差比较	21
二、RMSE 随 s 及 σ 的变化规律	27

第四节 有效价差的 Bootstrap 区间估计	30
一、Bootstrap 区间估计	30
二、实际例子	31
第五节 结论	33
第三章 基于价格极值的有效价差估计	34
第一节 引言	34
第二节 新的 High-Low 估计	35
第三节 理论性质	36
一、 $\hat{s}_1$ 的统计性质	36
二、 $\hat{s}_2$ 的统计性质	37
三、 $\hat{s}_3$ 的统计性质	38
四、 $\hat{s}_4$ 的统计性质	39
五、 $\hat{s}_5$ 的统计性质	41
六、统计性质比较	42
第四节 模拟研究	45
一、六种 High-Low 估计误差模拟结果	45
二、RMSE 随 s 和 σ 的变化趋势	50
第五节 广义矩估计	60
第六节 结论	66
第四章 有效价差的极大似然估计	68
第一节 引言	68
第二节 极大似然估计	69
第三节 理论性质	72
第四节 模拟研究	72
一、理想状况下误差变化	72
二、非理想状况下误差变化	79

三、估计误差随 s 及 σ 的变化	81
第五节 结论	87
第五章 中国股票市场流动性度量方法的比较	88
第一节 引言	88
第二节 高频基准价差	89
一、报价价差	90
二、有效价差	90
三、已实现价差	90
第三节 数据选取及说明	91
第四节 实证分析	92
一、描述性统计分析	92
二、估计误差比较	94
三、相关性比较	98
第五节 结论	101
第六章 中国债券市场流动性度量方法的比较	103
第一节 引言	103
第二节 流动性度量	105
一、高频交易成本	105
二、高频价格响应	106
三、低频交易成本	107
第三节 数据选取及描述性统计分析	107
第四节 实证结果	110
一、相关性分析	110
二、估计误差	112
第五节 结论	114

第七章 结论、不足之处及进一步研究方向	115
第一节 主要结论	115
第二节 不足之处	116
第三节 进一步研究方向	117
参考文献	119
附录	124
附录 A 第二章相关定理证明	124
一、引理 2.1 的证明	124
二、定理 2.1 的证明	125
三、引理 2.2 的证明	125
四、定理 2.2 的证明	126
五、推论 2.1 的证明	127
附录 B 第三章相关定理证明	129
一、引理 3.1 的证明	129
二、定理 3.1 的证明	129
三、引理 3.2 的证明	130
四、定理 3.2 的证明	131
五、引理 3.3 的证明	132
六、定理 3.3 的证明	133
七、引理 3.4 的证明	134
八、定理 3.4 的证明	134
九、引理 3.5 的证明	135
十、定理 3.5 的证明	136
十一、推论 3.1 的证明	137
附录 C 第四章相关定理证明	140
一、定理 4.1 的证明	140
二、定理 4.2 的证明	141

第一章

导 读

第一节 研究背景

金融市场微观结构研究主要是分析和探索金融市场中金融资产价格的形成机制以及市场的交易机制等,通过对市场交易机制设计的改善,可以促使金融市场更加有效地实现资源配置。流动性是衡量金融市场质量最重要的准则之一,因为适度的流动性能够提高市场效率,降低交易成本,从而促进市场的交易。因而流动性是金融市场微观结构研究里的一个关键问题,对于资产定价、公司金融以及风险管理等各类金融决策均有重要的意义,而流动性的测度则是这一类问题的基础。

流动性具体是指在市场上一笔较大额的交易能够快速实现,并且对市场的价格产生尽可能小的影响能力。在一个流动性较强的市场上,交易者应该能够以观测到的市场价格比较迅速地来完成任意数量资产的交易;反之,在一个缺乏流动性的市场上,交易者为了完成一个较大数额的交易,必须改变市场刚刚形成的价格或者通常需要等待较长的时间。因此,流动性的测度往往是从多个维度进行的,根据 Kyle(1985)对流动性的定义,衡量流动性一般有四个维度,即宽度、深度、即时性和弹性。其中,宽度是很重要的一个方面,其主要衡量的是交易成本指标,而度量交易成本的基本方式是买卖价差(Bid-ask Spread)。买卖价差越小,立即执行交易的成本就越低,从而市场流动性就越好。

买卖价差等流动性指标对于金融计量领域内的许多实证研究都具有重要的意义。例如,在资产定价研究领域内,Chordia 等(2000)指出,多种价差测度是系统地变化的;Goyenko(2006)证实了几种价差测度均是可以被定价的;Fujimoto(2004),Korajczyk 和 Sadka(2008)、Hasbrouck(2009)等都检验了买卖

价差在美国金融市场中的定价作用;而 Bekaert 等(2007)则指出买卖价差在新兴市场定价问题中所起到的作用。此外,包括买卖价差在内的各种流动性测度在公司金融研究中也逐渐变得更加重要,例如 Pham 等(2003)、Dennis 和 Strickland(2003)、Cao 等(2004)、Lipson 和 Mortal(2004)、Schrand 和 Verrecchia(2004)、Lesmond 等(2008)及其他许多学者都研究了股票的流动性在公司金融领域里的应用问题。此外,Helfin 和 Shaw(2000)、Lipson 和 Mortal(2004)、Lerner 和 Schoar(2004)等还探索了流动性对资本结构、证券发行形式以及其他金融决策问题的重要意义。由此可见,研究买卖价差等流动性的测度问题的确是金融市场微观结构中的一个关键问题。

依据买卖价差的定义,其又可以区分为报价价差、已实现价差、有效价差等。具体来讲,早期对买卖价差成分的研究主要是以报价驱动市场为对象,报价价差(Quoted Spread)是做市商报出的卖价(Ask Price)与买价(Bid Price)的差值,是可以直接观测到的数值,它反映做市商对订单执行成本的预期。在指令驱动市场,不少学者把指令簿价差(即最佳买卖报价价差和报价深度加权价差)作为报价价差。已实现价差(Realized Spread)衡量订单执行价格和订单执行后一段时间的买卖报价中点之间的差额。已实现价差反映订单执行后的市场影响成本。有效价差(Effective Spread)则是指订单的成交价格与订单到达时市场的均衡价格之间的差值。有效价差衡量订单的实际执行成本,在一定程度上克服买卖价差不能反映订单在买卖价差之外和之内成交的情况(即高估或低估执行成本)。考虑到实际的交易很多发生在报价范围内,有效价差被认为比报价价差更好地刻画了实际的交易成本(Petersen 和 Fialkowski,1994)。有效价差本质上是由成交价与均衡价格之间的差距决定的,而均衡价格无法直接观测到,因此需要根据观测到的价格信息对有效价差进行估计。如果能够获取高频的交易数据,通常利用买卖报价的中间价作为均衡价格的代理值,并由此计算出有效价差;在无法获取高频数据的情形下,需要利用低频观测数据对有效价差进行估计。

本书关注的是基于低频数据对金融市场有效价差的估计问题。尽管目前日内高频数据已经被广泛使用,我们仍然经常采用基于低频数据对有效价差的估计进行资产定价、公司金融以及市场有效性检验等各类金融决策问题的研究,这

主要基于三方面的原因：一是某些情况下研究者所考察的时间区间内日内高频数据尚无法获取，或者在研究涉及多个国际市场时，高频数据无法同时获得。例如，虽然美国市场能够获取从1983年开始的高频交易数据，但还有很多国家的金融市场日内高频交易数据仍是无法获得的。二是即使能够获得高频交易数据，部分报价数据和交易数据之间也可能存在不匹配的情况，直接利用其计算有效价差并不可靠。三是运用日度等低频数据估计交易成本可以减少处理海量高频数据的时间和工作量，特别是如果研究的重点聚焦在金融市场的平均交易成本，采用低频数据估计的交易成本就能收到不错的效果。所以，目前仍然有许多学者研究如何基于低频的交易数据进行有效价差的估计。

基于低频价格信息估计有效价差的最经典的结果来自 Roll(1984)。在 Roll 的价格模型里，成交价格(或者说观测价格)是在均衡价格的基础上按照交易由买方还是卖方发起相应地加上或减去有效价差，由此根据成交价格改变量的样本自协方差函数即可构造出对有效价差的估计。此后许多学者对 Roll 的工作进行了各种改进。例如，Choi 等(1988)认为在 Roll 的价格模型中应该考虑到交易方向的序列相关性；Holden(2009)进一步考虑了存在不交易的可能性及其影响；Hasbrouck(2004, 2009)针对 Roll 的模型给出了基于 Gibbs 抽样的贝叶斯(Bayes)估计方法等。此外，Glosten 和 Harris(1988)、Madanvan 等(1997)、Huang 和 Stoll(1997)等通过将交易指示量与价格变化和交易规模等联系起来拓展了 Roll 的模型，分别提出了一些新的价格模型，并由此给出了估计和分解买卖价差的方法。Holden(2009)拓展了 Huang 和 Stoll 的模型，并根据序列相关、价格聚类等基本思想提出了几种新的价差估计形式。Corwin 和 Schultz(2012)则在 Roll 模型的基本假设下依据每个交易日最高价和最低价的差值(即日内价格极差)提出了估计买卖价差的 High-Low 方法。上述方法均在假定价格模型为 Roll 模型的基础上对价差进行估计，而 Lesmond 等(1999)从构建收益率模型的角度出发，基于在非零收益率的交易日里知情交易发生以及在零收益率的交易日里知情交易未发生的假设提出了 LOT 等估计有效价差的思路；Fong 等(2017)在 LOT 模型的基础上进一步基于交易成本对称的假设，提出了 FHT 估计等。

在上述有效价差估计方法的基础上,Lesmond 等(2004)、Gehrig 和 Fohlin (2006)、Kim 等(2007)、Lipson 和 Mortal (2004)、Bharath 等(2009)、Asparouhova 等(2010)以及其他学者采用了 Roll 协方差估计以及基于 Roll 模型的一些拓展估计作为流动性的度量,研究了日内高频交易数据能够获取之前时,流动性在资产定价、公司金融以及市场有效性等问题中的应用。Amihud 等(2003)、Chakrabarti 等(2005)、Griffin 等(2010)则考察了 Roll 等有效价差估计在多个国际市场间的应用问题。此外,Lesmond 等(2004)、Mei 等(2005)、Bekaert 等(2007)以及 Griffin 等(2010)还进行了 LOT 估计的应用研究。

为了比较各种低频价差估计方法对流动性的度量效果,不少学者往往借助由高频数据计算或估计出来的买卖价差作为基准,并通过计算和比较不同低频估计与基准价差之间的相关系数来对其进行评价,而这种直接比较相关性的方法可能有至少两方面的不足:一是该比较方法必须从实证的角度出发,也就是说会过分依赖于某个具体的金融市场或者资产的高频以及低频的交易数据。如果变换为其他市场或资产的交易数据,这种相关性的差距或次序并不一定保持不变。另外,如果是针对不同方式定义的流动性度量,考察并比较它们与基准度量的相关性是可以理解的,但是如果只是针对同一个模型里相同参数的不同估计,比如 High-Low 估计及协方差估计都是对 Roll 的模型里有效价差参数的估计,那么按照统计和计量的逻辑,通过分析估计量的精度(比如偏差和均方误差)和统计性质来比较不同的估计显然是更加合理的做法。然而,很少有研究直接从分析估计量的精度和统计性质出发,因此,本书着重于研究低频价差估计的统计性质,首先主要研究 Roll 估计与 High-Low 估计的统计特征,并给出这两种估计的理论性质及模拟比较结果。

无论是 Corwin 和 Schultz(2012)的文章还是本书的研究结果,均证实了利用价格极差信息的 High-Low 估计比单纯采用收盘价信息的价差估计存在显著优势。受此启发,我们也进一步根据价格极差信息提出新的低频价差估计,包括(简单)矩估计、广义矩估计以及极大似然估计等,并且给出新估计的理论性质,同时进行随机模拟研究或者实证分析,从理论性质、模拟结果以及实证分析三个角度证实本书提出的有效价差估计方法是比较理想的价差估计。最后,根据基

于 Roll 的价格模型得到的多种低频有效价差的估计方法,本书分别对中国股票市场 and 债券市场进行实证研究,探索究竟哪种估计方法适用于中国的金融市场。

第二节 本书内容及结构

在上一节中,围绕着低频有效价差的估计,提出了本书研究的问题,这些问题可以归纳为四个方面,对这四个方面的研究将分别对应于后面互相关联的四个章节。下面根据这些问题的顺序,简要描述本书主要的研究工作,并给出本书具体的结构安排。

第一章中对低频有效价差的估计方法进行了文献综述。本书的基本模型即为文献综述中 Roll 的价格模型,下面各章研究都是以 Roll 的基本价格模型为基础的。

第二章的研究从低频价差估计的统计性质出发,由于 Roll 协方差估计和最近提出的 High-Low 估计都属于计算简便、较易获得并且使用广泛的估计,第二章首先计算了这两种估计偏差、均方误差等,并研究其相合性、渐近正态性等渐近性质,进而在理论上直接给出对这两种价差估计的统计评价。然后,采用随机模拟的方法对理论结果进行验证,随机模拟的设置分为理想状况和非理想状况,理想状况即价格可以连续地观测到,而非理想状况包括存在隔夜收益率以及价格不能被频繁地观测到等,并且本书给出两种价差估计的均方误差随波动率或买卖价差的变化趋势。最后,借助 Bootstrap 方法,考察两种估计方法在有限样本下的置信区间,从实际角度验证了 High-Low 估计的精度更高。

从第二章的研究可以得出利用价格极差信息的 High-Low 估计比单纯采用收盘价信息的价差估计的估计精度更高,而 Corwin 和 Schultz(2012)所提的 High-Low 估计主要是基于两个二阶矩条件的矩估计,因此,第三章采用更多的矩条件提出基于价格极差的新的矩估计或者广义矩估计。具体来讲,首先根据四个矩条件提出五个新的基于最高价格和最低价的矩估计,然后计算这五个矩估计的估计偏差、均方误差等,研究其相合性、渐近正态性等渐近性质,进而从理论上直接给出包括 Corwin 和 Schultz(2012)的 High-Low 估计在内的共六个基

于极差的矩估计的统计评价。理论结果表明, Corwin 和 Schultz 的估计并不是最优的 High-Low 估计。接下来, 本章对六种 High-Low 估计也进行模拟研究, 随机模拟的设置与第二章类似。此外, 本章还提出基于上述多个矩条件的广义矩估计等, 并通过模拟结果检验广义矩估计的估计精度等。

在第二章和第三章研究的基础上可以发现, 无论是 Roll 协方差估计还是 High-Low 估计, 实际上都是矩估计, 而且第三章提出的估计也均为矩估计或者广义矩估计。因此, 第四章则主要利用价格极差的对数近似服从正态分布的特性, 给出有效价差的一种新的近似极大似然估计方法。首先给出基于遗传算法的近似极大似然估计方法, 然后与前两章一致, 给出极大似然估计的渐近性质。接下来则针对协方差估计、贝叶斯估计以及 High-Low 的估计和本章的极大似然估计的精度通过随机模拟进行比较, 随机模拟的设置仍与前两章一致。最后给出一个简要的结论。

第五章和第六章结合前面三章中的低频价差估计方法, 分别基于中国股票和债券市场的高频交易数据对我国金融市场中有效价差估计的适用性进行实证研究。

最后一章给出本书所研究内容的一个简要的总结, 并对本书的不足之处进行阐述, 指出下一步改进和研究的方向。

第三节 低频估计的文献综述

第一节研究背景中提到的基于低频价格信息估计有效价差的方法, 最经典的结果来自于 Roll(1984)。在 Roll 的价格模型里, 观测到的价格(或者说成交价格)是在均衡价格的基础上按照交易是由买方还是卖方发起相应地加上或减去有效价差的一半, 并由此根据成交价格差分的样本自协方差函数即可构造出对有效价差的估计。此后许多学者对 Roll 的工作进行了各种改进。例如, Choi 等(1988)认为, 在 Roll 的价格模型中应该考虑到交易方向的序列相关性; Holden(2009)进一步考虑了存在不交易的可能性及其影响; Hasbrouck(2004, 2009)针对 Roll 的模型给出了基于 Gibbs 抽样的贝叶斯估计方法等; Holden

(2009)在拓展了 Roll 模型根据序列相关、价格聚类等基本思想提出了几种新的价差估计形式等。

Corwin 和 Schultz(2012)则基于 Roll 模型的基础上根据日内的价格极差提出了估计买卖价差的 High-Low 方法。易见,上述各种价差估计方法都是基于价格模型的基础上提出的,而 Lesmond 等(1999)则依据非零收益率的交易日里知情交易发生以及在零收益率的交易日里知情交易未发生的假设下给出了基于收益率的模型,并由此提出 LOT 等估计有效价差的思路;Fong 等(2017)在 LOT 模型的基础上进一步基于交易成本对称的假设,提出 FHT 估计等。下面将分别介绍上述几种有效价差的估计方法。

一、Roll 的估计

1. Roll 协方差估计

Roll(1984)提出了有效价差的协方差估计。其采用的价格模型为^①

$$\begin{cases} p_t^o = p_t + \frac{s}{2}Q_t \\ p_t = p_{t-1} + u_t \end{cases} \quad (1-1)$$

其中, p_t^o 为第 t 天最后一笔交易的对数观测价格(即第 t 天的对数收盘价); p_t 为第 t 天证券的对数基础价值或者真实价格; u_t 为第 t 天均值为 0 且方差为 σ^2 的公共信息冲击,且假定 u_t 序列不相关; s 为有效价差; Q_t 为交易方向指示变量, $Q_t=1$ 时代表买方发起的交易, $Q_t=-1$ 代表卖方发起的交易。Roll 进一步假设 Q_t 取 +1 或 -1 的概率相等,都是 1/2, Q_t 序列是独立的,即每次交易的方向没有关联,并且与 u_t 相互独立。

由 Roll 的模型易得

$$\Delta p_t^o = \frac{s}{2} \Delta Q_t + u_t \quad (1-2)$$

再据假设 $\text{Cov}(u_t, u_{t-j}) = 0 (j \neq 0)$ 及 $\text{Cov}(u_t, \Delta Q_{t-j}) = 0 (j \neq 0)$ 可得买卖价差

^① Roll(1984)的模型中是直接采用的原始价格。本书为了与 Corwin 和 Schultz (2012)的估计方法进行比较,统一将模型设置成对数价格。