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地实现，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和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]

$$\begin{cases} p_t^o = p_t + \dfrac{s}{2}Q_t \\ p_t = p_{t-1} + u_t \end{cases} \tag{1-1}$$

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

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

再据假设 $\mathrm{Cov}(u_t, u_{t-j})=0(j\neq 0)$ 及 $\mathrm{Cov}(u_t, \Delta Q_{t-j})=0(j\neq 0)$ 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。

满足

$$s = 2\sqrt{-\mathrm{Cov}(\Delta p_t^o, \Delta p_{t-1}^o)} \tag{1-3}$$

利用上述关系，显然可以利用观测价格差分序列的样本自协方差函数来估计买卖价差，具体如下：

$$\hat{s}_{\mathrm{Roll}} = \begin{cases} 2\sqrt{-\widehat{\mathrm{Cov}}(\Delta p_t^o, \Delta p_{t-1}^o)}, & \text{如果}\widehat{\mathrm{Cov}} < 0 \\ 0, & \text{如果}\widehat{\mathrm{Cov}} \geqslant 0 \end{cases} \tag{1-4}$$

上述估计通常被称为买卖价差的协方差估计或者简称 Roll 的估计。Roll(1984)放松其部分假设，允许基础价值变动，并在此基础上考虑更长的时间间隔的情形。这两种情形下得到的价格变动协方差与基本模型完全一致。由 Roll(1984)模型的前提假设可以看出，$s=2\sqrt{-\mathrm{Cov}(\Delta p_t^o, \Delta p_{t-1}^o)}$实际是在仅有订单处理成本情况下对买卖价差的估计，因此是有偏的被低估的估计。Harris(1990)推导出在 Roll 模型的理想条件下价差估计的偏差和方差，并解释在实际数据处理中 Roll 估计不准确的原因。

2. 考虑交易序列相关性的改进的 Roll 估计

Choi 等(1988)认为，有效买卖价差的估计还应考虑交易方向的序列相关性，令 $\delta = P(Q_t = 1 \mid Q_{t-1} = 1) = P(Q_t = -1 \mid Q_{t-1} = -1)$，模型与 Roll 模型一致，类似于 Roll 估计，易得

$$s = 2\sqrt{-\mathrm{Cov}(\Delta p_t^o, \Delta p_{t-1}^o)/(1-\delta)} \tag{1-5}$$

对 δ 的估计采用极大似然法，似然函数为 $L(\delta) = \prod_{i=1}^{n}[\delta I_i + (1-\delta)(1-I_i)]$，其中当 $Q_i = Q_{i-1}$时，$I_i = 1$；当 $Q_i = -Q_{i-1}$时，$I_i = 0$。Choi 等(1988)提出两种估计 I_i 的方法：一是利用高频报价直接计算 I_i；另一种是在高频报价无法获得时采用迭代的方法估计 I_i。

3. 包含不交易情形的 Roll 估计

Holden(2009)在 Roll 模型的基础上考虑到存在不交易的可能性，对 Roll 模型进行改进。假设某天发生交易的概率为 μ，不发生交易的概率为 $1-\mu$，假设交易方向等可能，即成交价格在买价和卖价的概率均为 $\mu/2$，则

$$s = 2\sqrt{\mathrm{Cov}(\Delta p_t^o, \Delta p_{t-1}^o)/\mu} \tag{1-6}$$

其中，$\hat{\mu}=\mathrm{TD}/(\mathrm{NTD}+\mathrm{TD})$，TD 为一段时间内发生交易的天数，NTD 为不发生交易的天数。

二、贝叶斯估计

由于 Roll 模型中的协方差有时为正值，得出的估计值会存在问题，Hasbrouck(2004,2009)对 Roll 模型中的参数 c、σ^2 给出基于 Gibbs 抽样的贝叶斯估计方法。模型及模型假设同 Roll 估计，样本量大小为 T，定义 $c=s/2$，观测到价格的样本为$\{p_1^o,\cdots,p_T^o\}$，潜变量为交易方向指示变量$\{Q_1,\cdots,Q_T\}$以及真实价格$\{p_1,\cdots,p_T\}$，对参数 c、σ^2 的先验分别为 $N^+(0,\sigma_{c,\mathrm{prior}}^2)$、$IG(\alpha,\beta)$（逆伽马分布）。Gibbs 抽样的步骤如下：

(1) 先初始化 $c^{[0]},\sigma^{2[0]},Q_1^{[0]},\cdots,Q_T^{[0]}$；

(2) 将式(1-2)看作回归方程，在 $Q_t=Q_t^{[i-1]}(t=1,\cdots,T)$，$\sigma^2=\sigma^{2[i-1]}$ 的条件下得到 c 的后验分布，从该分布中抽取 $c^{[i]}$；

(3) 通过 $c^{[i]}$ 计算回归的残差，进而得出 σ^2 的后验分布，抽取 $\sigma^{2[i]}$；

(4) 分别从 $f(Q_1\mid c^{[i]},\sigma^{2[i]},Q_2^{[i-1]},\cdots,Q_T^{[i-1]}),\cdots,f(Q_T\mid c^{[i]},\sigma^{2[i]},Q_1^{[i]},Q_2^{[i]},\cdots,Q_{T-1}^{[i]})$中抽取 $Q_1,\cdots,Q_T$；

重复进行(2)～(4)步 n 次，对 $c^{[1]},\cdots,c^{[n]}$取均值得到有效价差的估计。事实上，Gibbs Roll 的方法提高了 Roll 模型采用低频交易数据估计买卖价差的准确性，但是一定程度上增大了计算的负担。

三、Holden 估计

Holden(2009)以及 Goyenko 等(2009)根据价格聚类、序列相关等基本思想提出 Holden、Effective tick 等估计。

1. Holden 估计

Holden 对式(1-2)的价格变化过程做了一个自然的拓展，他认为 s_t 是变化的，由于买卖价差一般可分解为三个组成部分——逆向选择成本、订单处理成本以及存货成本，Holden 假设逆向选择成本和存货成本占价差的比例为 α，订单处

理成本的比例为 $1-\alpha$ 中。Holden 提出的模型为

$$\Delta p_t = \frac{s_t}{2}Q_t - (1-\alpha)\frac{s_{t-1}}{2}Q_{t-1} + u_t, u_t \sim N(\bar{u}, \sigma_u^2) \tag{1-7}$$

假设某天发生交易的概率为 μ，不发生交易的概率为 $1-\mu$，并且假定 s_t 从集合 $s_j, j=1,2,\cdots,J$ 中以对应的概率 $\gamma_j, j=1,2,\cdots,J$ 随机抽取，不妨设 $s_1, s_2,\cdots,s_J$ 已由小到大排序。例如考虑 1/8 美元价格细分的情形，s_t 有 γ_1 的概率取 $s_1=1/8$ 价差，γ_2 的概率取 $s_2=1/4$ 价差，γ_3 的概率取 $s_3=1/2$ 价差，γ_4 的概率取 $s_4=1$ 价差，则

$$\text{Holden} = \frac{\sum_{j=1}^{J}\hat{\gamma}_j s_j}{\bar{P}} \tag{1-8}$$

定义 C_t 为第 t 天观测到的价格类别，具体来说，若第 t 天无交易，则 $C_t=0$。在交易日里，$C_t=1,2,3,4$ 表示对应的交易价格为奇数倍的 1/8、1/4、1/2 以及整数。定义第 t 天不可观测的方向半价差 $H_t=s_tQ_t/2$，Q_t 为交易方向指示变量，此处 H_t 有 9 种可能的取值：0 美元，±1/16 美元，±1/8 美元，±1/4 美元，±1/2 美元。对于三个连续交易日，观测到价格为 (P_t, P_{t+1}, P_{t+2})，对应的价格类别为 (C_t, C_{t+1}, C_{t+2})。若 H 是所有给定价格类别后的半价差 (H_t, H_{t+1}, H_{t+2}) 的集合，则利用极大似然法可得到参数 $(\mu, \gamma_1, \gamma_2, \cdots, \gamma_{J-1}, \bar{u}, \sigma_u^2, \alpha)$ 的估计。似然函数如下：

$$\begin{aligned} & P(P_t, P_{t+1}, P_{t+2} \mid \mu, \gamma_1, \gamma_2, \cdots, \gamma_{J-1}, \bar{u}, \sigma_u^2, \alpha) = \\ & \sum_{(H_t, H_{t+1}, H_{t+2}) \in H} \{P(C_t) \cdot P(C_{t+1}) \cdot P(C_{t+2}) \cdot P(H_t \mid C_t) \cdot P(H_{t+1} \mid C_{t+1}) \cdot P(H_{t+2} \mid C_{t+2}) \\ & \qquad \cdot \varphi(P_{t+1} - H_{t+1} - (P_t - (1-\alpha)H_t)) \cdot \varphi(P_{t+2} - H_{t+2} - (P_{t+1} - (1-\alpha)H_{t+1}))\} \end{aligned} \tag{1-9}$$

其中，$\varphi(\cdot)$ 是均值为 $\bar{u}$、方差为 σ_u^2 的正态分布的密度函数。若考虑不交易的情形，再利用无交易里的买卖报价，Holden 提出了另一种形式的价差估计如下：

$$\text{Holden2} = \frac{\mu\sum_{j=1}^{J}\hat{\gamma}_j s_j + \begin{cases} (1-\mu)\dfrac{1}{\text{NTD}}\sum_{t=1}^{\text{NTD}} QS_t, & \text{NTD} > 0 \\ 0, & \text{NTD} = 0 \end{cases}}{\bar{P}} \tag{1-10}$$

其中，NTD 为没有交易的天数，$QS_t=A_t-B_t$ 为没有交易时的报价价差，仍然通过极大似然法来估计参数。

2. Effective 估计

Holden(2009)、Goyenko 等(2009)根据价格分类的特征同时提出 Effective tick 估计。

令 s_t 为第 t 天收盘交易的有效价差，假设 S_t 从集合 s_j，$j=1,2,\cdots,J$ 中以对应的概率 γ_j，$j=1,2,\cdots,J$ 随机抽取，不妨设 $s_1,s_2,\cdots,s_J$ 已由小到大排序。令 N_j 为交易价格对应价差为 s_j 时的交易量，F_j 为观测到的交易价格对应价差为 s_j 时的概率，则

$$F_j=\frac{N_j}{\sum_{j=1}^{J}N_j},\quad j=1,2,\cdots,J \tag{1-11}$$

令 U_j 为 s_j 对应的概率(Unconstrained)，则 $U_j=\begin{cases}2F_j, & j=1\\ 2F_j-F_{j-1}, & j=2,\cdots,J-1,\\ F_j-F_{j-1}, & j=J\end{cases}$

故

$$\hat{\gamma}_j=\begin{cases}\mathrm{Min}[\mathrm{Max}\{U_j,0\},1], & j=1\\ \mathrm{Min}\left[\mathrm{Max}\{U_j,0\},1-\sum_{k=1}^{j-1}\hat{\gamma}_k\right], & j=2,3,\cdots,J\end{cases} \tag{1-12}$$

则

$$\text{Effective Tick1}=\frac{\sum_{j=1}^{J}\hat{\gamma}_j s_j}{\bar{P}} \tag{1-13}$$

这里只考虑发生交易的价格，即 $\bar{P}$ 为一段时间内发生交易的价格平均值。Holden 在其文章中还提出其他三种 Effective Tick 估计，包括将所有价格均视为交易价格、将交易价格和非交易价格分开考虑、加入交易的概率等三种情形。

四、High-Low 估计

Corwin 和 Schultz(2012)提出了一种基于日内最高价和最低价估计买卖价

差的方法。该估计主要基于两个基本的认知：一是最高价通常是由买方发起的交易形成的，而最低价则一般是由卖方发起的交易形成的；二是一个时段里对数最高价和最低价的差值，即所谓的价格极差(Range)，既包含波动率的成分又包含买卖价差的成分，而且它所蕴含的波动率成分与时段的长度成正比，而对应的买卖价差部分则与时段的长度无关。

记第 t 天里交易的最高价、最低价的真实值分别为 H_t 和 L_t，对应的能够观测到的最高价和最低价分别记为 H_t^o 和 L_t^o。此外，将第 t 天和第 $t+1$ 天相邻两天里的最高价和最低价的真实值记为 $H_{t,t+1}$、$L_{t,t+1}$，对应的观测值则分别记为 $H_{t,t+1}^o$ 和 $L_{t,t+1}^o$。在式(1-1)的设置下，观测到的最高价与最低价与真实的最高价与最低价以及买卖价差的关系可以具体表达为

$$\ln H_t^o = \ln H_t + \frac{s}{2},\quad \ln L_t^o = \ln L_t - \frac{s}{2} \tag{1-14}$$

易见

$$R_t^o \triangleq \ln H_t^o - \ln L_t^o = (\ln H_t - \ln L_t) + s \triangleq R_t + s \tag{1-15}$$

其中，分别用 R_t 和 R_t^o 来表示实际的价格极差和观测到的价格极差，显然两者的差值即给出买卖价差。

由式(1-15)得到

$$E[R_t^{o2}] = E[R_t^2] + 2sE[R_t] + s^2 \tag{1-16}$$

根据 Feller(1951)以及 Parkinson(1980)等给出的具有零漂移项的布朗过程的极差的统计性质，可以得到

$$\gamma_1 \triangleq E[R_t^{o2}] = k_2\sigma^2 + 2k_1\sigma + s^2 \tag{1-17}$$

其中，$k_1=\sqrt{8/\pi}\approx 1.5958$，$k_2=4\ln 2\approx 2.7726$ 为两个常数。

类似地，对连续两天的价格极差 $R_{t,t+1}^o=\ln H_{t,t+1}^o-\ln L_{t,t+1}^o$ 重复上述分析并注意到 $E[R_{t,t+1}]=\sqrt{2}\sigma$ 以及 $E[R_{t,t+1}^2]=2\sigma^2$，得到

$$\gamma_2 \triangleq E[R_{t,t+1}^{o2}] = k_2 2\sigma^2 + 2\sqrt{2}k_1\sigma + s^2 \tag{1-18}$$

根据式(1-17)和式(1-18)式，可以通过迭代算法解出 s 和 σ 并将其表示为 γ_1 和 γ_2 的函数。Corwin 和 Schultz(2012)注意到 $k_2\approx k_1^2$，建议可以忽略两者之间的差异，从而得到 s 和 σ 的显式解

$$s = \frac{\sqrt{2\gamma_1} - \sqrt{\gamma_2}}{\sqrt{2} - 1}, \quad \sigma = \frac{\sqrt{\gamma_2} - \sqrt{\gamma_1}}{(\sqrt{2} - 1)k_1} \tag{1-19}$$

那么，用观测价格极差的样本矩来取代上式中的总体的矩 γ_1 及 γ_2，即可得到相应的买卖价差与波动率的估计

$$\hat{s}_{hl} = \frac{\sqrt{2\hat{\gamma}_1} - \sqrt{\hat{\gamma}_2}}{\sqrt{2} - 1}, \quad \hat{\sigma}_{hl} = \frac{\sqrt{\hat{\gamma}_2} - \sqrt{\hat{\gamma}_1}}{(\sqrt{2} - 1)k_1} \tag{1-20}$$

其中，

$$\begin{aligned} \hat{\gamma}_1 &= \frac{1}{n}\sum_{t=1}^{n} R_t^{o2} = \frac{1}{n}\sum_{t=1}^{n}\left[\ln\frac{H_t^o}{L_t^o}\right]^2, \\ \hat{\gamma}_2 &= \frac{1}{n-1}\sum_{t=1}^{n-1} R_{t,t+1}^{o2} = \frac{1}{n-1}\sum_{t=1}^{n-1}\left[\ln\frac{H_{t,t+1}^o}{L_{t,t+1}^o}\right]^2 \end{aligned} \tag{1-21}$$

五、LOT 估计

Lesmond 等(1999)提出一种新的估计有效价差的方法，他们假设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，基于此，提出 LOT、Zeros 等估计，基本模型为

$$R_{jt}^* = \beta_j R_{mt} + \varepsilon_{jt} \tag{1-22}$$

其中，R_{jt}^* 是股票 j 在第 t 天的真实收益率，R_{mt} 为第 t 天的市场收益率，β_j 是股票 j 对 R_{mt} 的敏感度，ε_{jt} 是第 t 天的公共信息冲击。假设 $\varepsilon_{jt} \sim N(0, \sigma_j^2)$，令 $\alpha_{1j} \leqslant 0$ 为卖出股票 j 的相对交易费用，$\alpha_{1j} \geqslant 0$ 为买入股票 j 的相对交易费用，R_{jt} 是股票 j 第 t 天的观测收益率，则股票 j 的观测收益率 R_{jt} 满足

$$\begin{aligned} R_{jt} &= R_{jt}^* - \alpha_{1j}, R_{jt}^* < \alpha_{1j} \\ R_{jt} &= R_{jt}^*, \alpha_{1j} < R_{jt}^* < \alpha_{2j} \\ R_{jt} &= R_{jt}^* - \alpha_{2j}, R_{jt}^* > \alpha_{2j} \end{aligned} \tag{1-23}$$

因此得到的有效价差估计为

$$\mathrm{LOT} = \alpha_{2j} - \alpha_{1j} \tag{1-24}$$

通过极大似然法估计出模型的参数，极大似然函数如式(1-25)所示：

$$L(\alpha_{1j},\alpha_{2j},\beta_j,\sigma_j \mid R_{jt},R_{mt})$$
$$=\prod_1 \frac{1}{\sigma_j}\varphi\left[\frac{R_{jt}+\alpha_{1j}-\beta_j R_{mt}}{\sigma_j}\right]\times \prod_0 \left[\phi\left(\frac{\alpha_{2j}-\beta_j R_{mt}}{\sigma_j}\right)-\phi\left(\frac{\alpha_{1j}-\beta_j R_{mt}}{\sigma_j}\right)\right]$$
$$\times \prod_2 \frac{1}{\sigma_j}\varphi\left[\frac{R_{jt}+\alpha_{2j}-\beta_j R_{mt}}{\sigma_j}\right] \tag{1-25}$$

s. t. $\alpha_{1j}\leqslant 0,\alpha_{2j}\geqslant 0,\beta_j\geqslant 0,\sigma_j>0$,其中 $\varphi(\cdots)$和 $\phi(\cdots)$分别为标准正态分布的密度概率和分布函数。

关于 LOT 估计一个重要的问题是 3 个区域的划分,Lesmond 等(1999)根据 R_{jt}、R_{mt} 的范围划分:区域 0,$R_{jt}=0$;区域 1,$R_{jt}\neq 0,R_{mt}<0$;区域 2,$R_{jt}\neq 0$,$R_{mt}>0$。Goyenko 等(2009)提出了一种新的划分方法:区域 0,$R_{jt}=0$;区域 1,$R_{jt}<0$;区域 2,$R_{jt}>0$。

六、FHT 估计

Fong 等(2017)在 LOT 模型的基础上,通过对收益率和交易成本进行以下两方面的假设,得到了一种新的交易成本的估计,即 FHT 估计:①真实收益率 R_{jt}^* 服从均值为 0,方差为 σ_j^2 的正态分布;②交易成本 s_j 是对称的。根据这两方面假设,可以将 LOT 模型变形为 FHT 模型:

$$R_{jt}=\begin{cases} R_{jt}^*+s_j/2, & \text{如果 } R_{jt}^*<-s_j/2 \\ 0, & \text{如果 } -s_j/2\leqslant R_{jt}^*\leqslant s_j/2 \\ R_{jt}^*-s_j/2, & \text{如果 } R_{jt}^*>s_j/2 \end{cases} \tag{1-26}$$

其中,R_{jt} 是第 t 天的观测收益率,根据 R_{jt}^* 的正态性假设,计算零收益率发生的概率。同时,由于 Zeros 度量给出了实际中观测到的零收益率发生的频率,基于矩估计的想法可得

$$\Phi\left(\frac{s_j}{2\sigma_j}\right)-\Phi\left(-\frac{s_j}{2\sigma_j}\right)=\text{Zeros} \tag{1-27}$$

即

$$FHT=2\sigma_j\Phi^{-1}\left(\frac{1+\text{Zeros}}{2}\right) \tag{1-28}$$

七、低频价差估计的比较

针对上述的各种基于低频交易数据的有效价差估计，Hasbrouck(2004，2009)计算了多种基于低频数据构造的价差估计与基准买卖价差指标间的相关系数，得到基于 Gibbs 抽样的贝叶斯估计优于其余估计方法的结论；Goyenko 等(2009)比较包括上述大部分估计方法在内的低频流动性指标与高频基准指标间的相关性强弱和估计误差大小，给出一些适合美国市场的低频流动性指标；类似地，Holden(2009)也采用相关性和估计误差的评价标准比较上述几种低频价差估计方法的流动性度量效果；Corwin 和 Schultz(2012)比较 High-Low 估计、Roll 估计等四种价差估计方法与高频基准价差之间的相关性大小，发现 High-Low 估计的表现优于 Roll 等其他估计。Marshall 等(2012)、Karnaukh 等(2015)以及 Schestag 等(2016)分别对美国的商品期货市场、外汇市场以及债券市场也进行了类似的研究。

目前对于中国金融市场流动性度量效果的比较研究还比较少，而且主要集中在股票市场，例如 Ahn 等(2012)比较了包括中国市场在内的全球 21 个新兴市场的低频价差估计等流动性指标的度量效果；张峥等(2014)根据 1999 年到 2009 年的中国股票市场的数据，比较了不同低频流动性指标的度量效果。

此外，在已有的研究中，比较不同有效价差低频指标度量效果大多是通过比较低频流动性指标与高频基准指标之间的相关系数，相关系数越高说明低频流动性指标的度量效果越好，较少有学者从低频流动性指标的估计精度的角度比较其度量效果。常用的低频流动性指标，如 Roll 的协方差估计、Hasbrouck 的 Gibbs Roll 估计、High-Low 估计和 FHT 估计等，除了可以通过它们的大小变化得到流动性强弱变化的规律，还可以准确得到当前的交易成本或者买卖价差的具体数值。由于市场微观结构关于流动性的研究中，通常不仅需要得到流动性强弱的动态变化规律，还需要了解交易成本或者买卖价差的具体数值，因此获得一些衡量具体交易成本或价差结果的指标就显得尤为重要。

第二章 两种有效价差估计的渐近性质

第一节 引 言

通过前面的参考文献回顾中可以看出，有效价差的低频估计方法种类很多，如何比较不同估计方法的流动性度量效果是一个值得研究的问题。为了比较各种低频价差估计方法对流动性的度量效果，将基于高频交易数据计算的价差作为基准，并且比较基准价差与各种低频估计间的相关性强弱是不少学者通常采用的评价方法。例如，Hasbrouck(2004,2009)计算了多种基于低频数据构造的价差估计与基准买卖价差指标间的相关系数，得到基于 Gibbs 抽样的贝叶斯估计优于其余估计方法的结论；Holden(2009)也采用相关性和估计误差的评价标准比较了几种低频价差估计方法的流动性度量效果；Corwin 和 Schultz(2012)比较了 High-Low 估计、Roll 估计等四种价差估计方法与高频基准价差之间的相关性大小，发现 High-Low 估计的表现优于 Roll 等其他估计。

然而，正如本书前面研究背景所提到的，容易看出这种通过计算相关系数来比较不同低频价差估计的方法存在两个方面的缺陷。首先，因为该方法一定是从实证的角度出发，所以会过分依赖于某个具体的金融资产或者市场高频或低频的交易数据。然而如果变换为其他资产或市场的交易数据，这种相关性的次序或差距就有可能不再成立。其次，如果考察不同方式定义的流动性度量方法，比较其与基准度量指标间的相关性强弱是能够理解的，但是如果比较基于同一模型中同一参数的不同估计方法，比如基于 Roll 价格模型的得到 Roll 协方差估计和 High-Low 估计均是对同一有效价差参数的估计，那么如果从计量和统计

的逻辑思维出发，通过分析不同估计量的偏差和均方误差等估计精度和统计性质来比较不同估计方法的优劣，显然是更合理的做法。

因此，本章将以低频价差估计方法的统计性质作为出发点，计算并比较不同价差估计方法的估计误差（偏差、均方误差等），并给出估计的相合性、渐近正态性等渐近性质，从而直接从理论上给出针对不同价差估计方法的统计评价。因为 Roll 协方差估计与 Corwin 和 Schultz（2012）提出的 High-Low 估计都是使用较多、计算简便且易得的低频价差估计方法，所以本章主要基于 Roll 协方差估计和 High-Low 估计进行统计特征的研究，并针对这两种估计给出其理论性质及模拟和实证分析的比较结果。由于 Harris（1990）曾对 Roll 的协方差估计进行过一些统计性质的研究，本章的研究也可看作对 Harris 工作的拓展。

本章的结构如下：第二节给出了 Roll 估计和 High-Low 估计的理论性质，第三节通过随机模拟对这两种估计的度量效果进行比较，第四节基于 Bootstrap 区间分析对两种估计的流动性度量效果进行实证研究，最后的结论在第五节。

第二节　理论性质

Corwin 和 Schultz（2012）通过随机模拟及实例分析指出，High-Low 估计与高频基准价差间的相关性要强于 Roll 估计相对应的结果。Hasbrouck（2004，2009）、Holden（2009）等也采用类似的方法比较过不同有效价差估计方法的好坏。但是，正如本章前一小节所提到的，尽管在实际中比较估计值与基准值之间的相关性可能会具有一定的启发性，但是却无法准确衡量估计的精确度。实际上，相关性很强的变量之间的差距可能很大，所以只比较相关性的方法并不符合统计推断的逻辑。因此，本小节将直接研究 Roll 及 High-Low 估计的统计性质，从理论上计算和推导上述两种估计的偏差和均方误差，进而直接比较两种估计方法的优劣。本小节在式（1-1）及式（1-14）的基础上研究两种有效价差估计 $\hat{s}_{roll}$ 和 $\hat{s}_{hl}$ 的统计性质，具体包括计算其偏差和均方误差的大小、推导其相合性和渐近分布等。主要的结果如以下的几个引理和定理所示，证明的具体细节参见附录 A。

一、Roll 估计的渐近性质

引理 2.1 在 Roll 模型的假设下：

$$c = \mathrm{Cov}(\Delta p_t^o, \Delta p_{t-1}^o), \hat{c} = \frac{1}{n}\sum_{t=2}^{n}(\Delta p_t^o - \overline{\Delta p^o})(\Delta p_{t-1}^o - \overline{\Delta p^o}),\text{则}$$

(1) $bias(\hat{c}) = E\hat{c} - c = -\frac{1}{n}(3c + \sigma^2) + o\left(\frac{1}{n}\right)$

$$\mathrm{Var}(\hat{c}) = \frac{1}{n}[(\eta + 4)c^2 + \sigma^4 - 4c\sigma^2] + o\left(\frac{1}{n}\right)$$

(2) $p\lim\limits_{n\to\infty}\hat{c} = c, \sqrt{n}(\hat{c} - c) \xrightarrow{d} N(0, (\eta + 4)c^2 + \sigma^4 - 4c\sigma^2)$

其中存在均值为 0，方差为 σ_e^2 的不相关序列$\{e_t\}$，使得 $\Delta p_t^o = e_t + \theta e_{t-1}$，且 $\eta = Ee_t^4/\sigma_e^4$。

根据引理 2.1，同时由式(1-3)和式(1-4)可得 $c = -s^2/4$ 以及 $\hat{s}_{roll} = 2\sqrt{-\hat{c}}$，因此可推知 Roll 估计的性质如定理 2.1 所示。

定理 2.1 在 Roll 模型的假设下：

(1) $E(\hat{s}_{roll}) < s$ 且

$$bias(\hat{s}_{roll}) = -\frac{1}{n}\left[\left(1 + \frac{\eta}{8}\right)s + \frac{2\sigma^4}{s^3}\right] + o\left(\frac{1}{n}\right)$$

$$\mathrm{Var}(\hat{s}_{roll}) = \frac{1}{n}\left[\left(\frac{\eta}{4} + 1\right)s^2 + 4\sigma^2 + \frac{4\sigma^4}{s^2}\right] + o\left(\frac{1}{n}\right)$$

(2) $p\lim\limits_{n\to\infty}\hat{s}_{roll} = s, \sqrt{n}(\hat{s}_{roll} - s) \xrightarrow{d} N(0, \sigma^2(\hat{s}_{roll}))$

其中 $\sigma^2(\hat{s}_{roll}) = \left(\frac{\eta}{4} + 1\right)s^2 + 4\sigma^2 + \frac{4\sigma^4}{s^2}$。

定理 2.1 对 Harris(1990)当中的有关结果进行了拓展。

二、High-Low 估计的渐近性质

引理 2.2 在基于式(1-1)以及条件式(1-14)的假设下：

$$\hat{\gamma}_1 = \frac{1}{n}\sum_{t=1}^{n}\left[\ln\frac{H_t}{L_t} + s\right]^2, \quad \hat{\gamma}_2 = \frac{1}{n-1}\sum_{t=1}^{n-1}\left[\ln\frac{H_{t,t+1}}{L_{t,t+1}} + s\right]^2$$

由此可推知 $\hat{\gamma}_1$ 和$\hat{\gamma}_2$ 的性质如下：

(1) $\operatorname{Var}(\hat{\gamma}_1)=\dfrac{1}{n}[(k_4-k_2^2)\sigma^4+4(k_3-k_1k_2)s\sigma^3+4(k_2-k_1^2)s^2\sigma^2]$

$$\begin{aligned}\operatorname{Var}(\hat{\gamma}_2)=\frac{1}{n-1}\{&4(1+2\rho_{22})(k_4-k_2^2)\sigma^4\\&+8\sqrt{2}[k_3-k_1k_2+(\rho_{12}+\rho_{21})\sqrt{k_4-k_2^2}\sqrt{k_2-k_1^2}]s\sigma^3\\&+8(1+2\rho_{11})(k_2-k_1^2)s^2\sigma^2\}\end{aligned}$$

$$\begin{aligned}\operatorname{Cov}(\hat{\gamma}_1,\hat{\gamma}_2)=\frac{1}{n}[&4\delta_{22}(k_4-k_2^2)\sigma^4\\&+4\sqrt{2}(\delta_{21}+\sqrt{2}\delta_{12})\sqrt{k_4-k_2^2}\sqrt{k_2-k_1^2}s\sigma^3\\&+8\sqrt{2}\delta_{11}(k_2-k_1^2)s^2\sigma^2]\end{aligned}$$

(2) $\sqrt{n}\begin{pmatrix}\hat{\gamma}_1-\gamma_2\\ \hat{\gamma}_2-\gamma_2\end{pmatrix}\xrightarrow{d}N(0,\Sigma)$

其中，$\Sigma=\begin{pmatrix}\sigma_{11}&\sigma_{12}\\ \sigma_{21}&\sigma_{22}\end{pmatrix}$，且 $\sigma_{11}=(k_4-k_2^2)\sigma^4+4(k_3-k_1k_2)s\sigma^3+4(k_2-k_1^2)s^2\sigma^2$

$$\begin{aligned}\sigma_{12}=\sigma_{21}=&4\delta_{22}(k_4-k_2^2)\sigma^4+4\sqrt{2}(\delta_{21}+\sqrt{2}\delta_{12})\sqrt{k_4-k_2^2}\sqrt{k_2-k_1^2}s\sigma^3\\&+8\sqrt{2}\delta_{11}(k_2-k_1^2)s^2\sigma^2\end{aligned}$$

$$\begin{aligned}\sigma_{22}=&4(1+2\rho_{22})(k_4-k_2^2)\sigma^4+8\sqrt{2}[k_3-k_1k_2\\&+(\rho_{12}+\rho_{21})\sqrt{k_4-k_2^2}\sqrt{k_2-k_1^2}]s\sigma^3+8(1+2\rho_{11})(k_2-k_1^2)s^2\sigma^2\end{aligned}$$

上述表达式中的 k_1,k_2,k_3,k_4 以及 $\rho_{ij},\delta_{ij}(i,j=1,2)$ 均为常数，其中 k_1,k_2 如前所述，$k_3=(2\pi)^{3/2}/3\approx5.2499$，$k_4=9\zeta(3)\approx10.8185$，$\zeta(x)$ 为 zeta 函数，具体的细节可参见 Feller (1951)和 Parkinson(1980)。ρ_{ij},δ_{ij} 分别定义为 $\rho_{ij}=Corr(R_{t,t+1}^i,R_{t+1,t+2}^j)$，$\delta_{ij}=Corr(R_t^i,R_{t,t+1}^j)$，$i,j=1,2$。因为 ρ_{ij} 和 δ_{ij} 两式中均不存在未知参数，因此我们可以通过数值随机模拟得到参数 ρ 和 δ 任意精度的估计值。例如，我们通过模拟每天 10 000 个价格观测值，根据 50 000 天的数据可以得到的计算结果为 $\rho_{11}=0.3395$，$\rho_{12}=\rho_{21}=0.3323$，$\rho_{22}=0.3304$，$\delta_{11}=0.5801$，$\delta_{12}=0.5602$，$\delta_{21}=0.5830$，$\delta_{22}=0.5767$。

由引理 2.2 我们可以推导出有效价差的 High-Low 估计$\hat{s}_{hl}$的统计性质，如定理 2.2 所示。

定理 2.2 在式(1-1)以及式(1-14)的假设下，成立

(1) $E(\hat{s}_{hl}) > s$，且

$$bias(\hat{s}_{hl}) = \frac{2+\sqrt{2}}{8n}\left\{-\frac{\sigma_{11}}{(k_2\sigma^2+2k_1s\sigma+s^2)^{3/2}}+\frac{\sigma_{22}}{\sqrt{2}\,(2k_2\sigma^2+2\sqrt{2}k_1s\sigma+s^2)^{3/2}}\right\}+o\left(\frac{1}{n}\right)$$

$$\mathrm{Var}(\hat{s}_{hl}) = \frac{3+2\sqrt{2}}{4n}\left\{\frac{2\sigma_{11}}{k_2\sigma^2+2k_1s\sigma+s^2}+\frac{\sigma_{22}}{2k_2\sigma^2+2\sqrt{2}k_1s\sigma+s^2}-\frac{2\sqrt{2}\sigma_{12}}{(k_2\sigma^2+2k_1s\sigma+s^2)^{1/2}(2k_2\sigma^2+2\sqrt{2}k_1s\sigma+s^2)^{1/2}}\right\}+o\left(\frac{1}{n}\right)$$

(2) $p\lim\limits_{n\to\infty}\hat{s}_{hl} = s$，$\sqrt{n}(\hat{s}_{hl}-s)\xrightarrow{d}N(0,\sigma^2(\hat{s}_{hl}))$

其中，

$$\sigma^2(\hat{s}_{hl}) = \frac{3+2\sqrt{2}}{4}\left\{\frac{2\sigma_{11}}{k_2\sigma^2+2k_1s\sigma+s^2}+\frac{\sigma_{22}}{2k_2\sigma^2+2\sqrt{2}k_1s\sigma+s^2}-\frac{2\sqrt{2}\sigma_{12}}{(k_2\sigma^2+2k_1s\sigma+s^2)^{1/2}(2k_2\sigma^2+2\sqrt{2}k_1s\sigma+s^2)^{1/2}}\right\}$$

三、两种估计的比较

综合定理 2.1 和定理 2.2 中的结论，我们可以比较容易地得到以下推论。

推论 2.1 当 $n\to\infty$时，

(1) $|bias(\hat{s}_{roll})|>|bias(\hat{s}_{hl})|$，$\dfrac{|bias(\hat{s}_{roll})|}{|bias(\hat{s}_{hl})|}>\dfrac{16\sqrt[4]{6}(1+\eta/8)^{3/4}k_2^{3/2}}{3(2+\sqrt{2})\rho_{22}(k_4-k_2^2)}>11$

(2) $RMSE(\hat{s}_{roll})>RMSE(\hat{s}_{hl})$，$\dfrac{RMSE(\hat{s}_{roll})}{RMSE(\hat{s}_{hl})}>\left(\dfrac{[(\eta+4)^{1/2}+2]k_2}{(2+2\rho_{22}-4\delta_{22})(k_4-k_2^2)}\right)^{1/2}>2.45$

无论是从估计的偏差还是从均方误差来讲，Roll 的估计精度均比 High-Low 的估计精度差。