

2010年12月22日

光大证券
EVERBRIGHT SECURITIES

股指期货微观市场初探——高频数据、交易的解读

◇ 股指期货高频交易系列报告

- **研究高频数据的意义。**高频数据代表着一个超大容量的数据群。在一个流动性较好的市场，单个交易日的高频数据量大致相当于30年按日统计的交易数据量(Dacorogna et al. (2001))。单纯从统计角度来看，独立观测值的数量越多，自由度也就越多，从而估计的参数也越精确。大批量的数据不仅可以提供更多的市场信息，也能够让我们更好地比较模型、区分不同模型的优劣。
- **沪深300股指期货高频数据来源与形式。**目前，中金所发布的股指期货高频数据可以通过各类证券行情软件和期货交易软件查看，但一般的软件不提供数据下载，而可提供下载的软件数据也往往止于分钟数据。为了更好地对这些高频数据进行下载，或得到实时的高频数据传输，一般需专业的数据经营商如WIND、彭博或路透数据终端的支持。
- **针对沪深300股指期货高频数据微观结构的探索。**沪深300股指期货单个交易日的成交量走势具有一定的周期性。由于期货开盘先于现货，收盘也晚于现货，在没有现货价格支撑的情况下，期指成交量在上午开盘和下午收盘时普遍较低。此外，在上午临近收盘和下午刚开盘的时刻其成交量也显著低于日内平均值，体现出投资者通常短线操作，在午间休市时期较为谨慎的投资心态。值得注意的是，成交量在下午收盘前的45分钟内呈现稳步减少的态势，而在最后5分钟内又有显著拉升。
- **沪深300高频交易机会。**通过观察市场微观结构，我们对样本进行简单的模拟信号测试，交易策略的盈利能力惊人！在选取的样本区间内，信号所产生的平均每日盈利点数达到190点，日回报率达到5%以上。

金融工程分析师：

刘道明(执业证书：S0930210060005)

021-22169109

liudaoming@ebcn.com

联系人：

李聚华，CFA

021-22169321

lijuhua@ebcn.com

金融工程研究团队

刘道明 李聚华 于栋华

张美云 熊伟 倪蕴韬

相关报告：

《光大证券股指期货连续指数—树立连续指数新标杆》

目录

一、	研究高频数据的意义	7
1.	利用完整的市场数据充分解读市场可提供的信息	7
2.	数量模型取代主观判断，减少非理性的判断失误	7
3.	利用计算机完成交易策略的执行	8
4.	算法交易和程序化交易降低交易成本	8
1)	算法交易	8
2)	程序化交易	8
5.	高频交易创造高回报投资机会	9
二、	沪深 300 股指期货高频数据来源与形式	10
1.	高频数据的来源	10
2.	数据存储的形式	13
三、	针对沪深 300 股指期货高频数据微观结构的探索	14
1.	样本选取	14
2.	样本内价格回报特性	14
3.	成交价日内走势	15
4.	成交量日内走势	18
5.	成交价格日内波动幅度走势	19
6.	买卖价差日内走势	20
7.	成交量与报价量日内平均比例	21
8.	高频波动率计算	23
1)	传统波动率估计方法	23
2)	传统日内波动率的估计方法：已实现波动率 (Realized Volatility, RV)	25
3)	TSRV (Two Scale Realized Volatility) 日内波动率估计的方法	26
9.	市场冲击成本模型介绍与应用	27
1)	冲击成本的形成	27
2)	影响冲击成本的因素	28
3)	冲击成本计算例证	29
四、	沪深 300 股指期货高频交易机会	31
1.	市场有效性检验	31
2.	高频交易的利润机会	32
3.	高频交易机会的探索	33
1)	交易信号的产生	33
2)	交易模型测试	34
4.	交易策略的实战可行性分析	35
1)	交易成本	35
2)	交易执行	35
3)	中金所限制	35
五、	股指期货高频交易的未来展望	36
1.	市场容量的增长	36
2.	机构参与度的增加	36

3. 交易机会的挖掘.....37

图目录

图 1: EXCEL 的 Bloomberg 插件数据导入步骤一	10
图 2: EXCEL 的 Bloomberg 插件数据导入步骤二	10
图 3: EXCEL 的 Bloomberg 插件数据导入步骤三	11
图 4: EXCEL 中 Bloomberg 的股指期货高频数据下载结果	11
图 5: Bloomberg 股指期货实时高频数据下载订购步骤一	11
图 6: Bloomberg 股指期货实时高频数据下载订购步骤二	12
图 7: WIND 股指期货高频数据下载步骤一	12
图 8: WIND 股指期货高频数据下载步骤二	12
图 9: WIND 股指期货高频数据下载步骤三	13
图 10: WIND 股指期货高频数据下载结果	13
图 11: 5 月主力合约 4 月 16 日分笔数据价格走势	15
图 12: 5 月主力合约 4 月 16 日一分钟间隔价格走势	15
图 13: 5 月主力合约 4 月 16 日五分钟间隔价格走势	16
图 14: 6 月主力合约 6 月 1 日分笔数据价格走势	16
图 15: 6 月主力合约 6 月 1 日一分钟间隔价格走势	16
图 16: 6 月主力合约 6 月 1 日五分钟间隔价格走势	17
图 17: 5 月至 8 月主力合约 10 分钟间隔价格走势	17
图 18: 5 月至 8 月主力合约 30 分钟间隔价格走势	18
图 19: 5 月至 8 月主力合约 60 分钟间隔价格走势	18
图 20: 5 月至 8 月主力合约日内一分钟平均成交量	19
图 21: 5 月至 8 月主力合约日内五分钟平均成交量	19
图 22: 5 月至 8 月主力合约日内一分钟平均波动幅度	20
图 23: 5 月至 8 月主力合约日内五分钟平均波动幅度	20
图 24: 5 月至 8 月主力合约日内一分钟平均买卖价差	21
图 25: 5 月至 8 月主力合约日内五分钟平均买卖价差	21
图 26: 5 月至 8 月主力合约日内一分钟平均成交量与买一量之比	22
图 27: 5 月至 8 月主力合约日内一分钟平均成交量与卖一量之比	22
图 28: 5 月至 8 月主力合约成交量与最优买卖报价量频率统计	23
图 29: 5 月至 8 月主力合约最优买入报价量与最优卖出报价量频率统计	23
图 30: 日收益率分解图	24
图 31: 光大股指期货主力合约连续指数历史波动率比较	24
图 32: 已实现波动率 (RV) 的计算过程	25
图 33: 已实现波动率与样本量之间的关系	25
图 34: TSRV 的计算过程	26
图 35: RV 与 TSRV (K=10min) 的比较	27
图 36: TSRV 与标准差法波动率的比较	27
图 37: 买入成交造成冲击成本的演化图	28
图 38: 冲击成本与成交速度的关系	29
图 39: 信号产生后价格点位的累积走势 (子样本 4/16/2010-4/22/2010)	33

图 40：交易信号盈利点数模拟走势（子样本 4/16/2010-4/22/2010）34

图 41：交易信号全样本盈利统计35

表目录

表 1：股指期货不同频率下的回报率统计.....14

表 2：股指期货委托报价（成交前）29

表 3：股指期货委托报价（成交时）30

表 4：股指期货委托报价（成交后）30

表 5：股指期货主力合约 2010/4/16 开盘交易价格变化31

表 6：不同频率下交易策略的最优表现比较.....32

表 7：交易信号全样本盈利统计.....34

一、 研究高频数据的意义

1. 利用完整的市场数据充分解读市场可提供的信息

一个著名的攀岩者曾在接受采访时被问到，“为什么要挑战那些危险性极高的山峰”，而他的答案是“因为他们存在着。”如果要问我们为何要选择高频数据进行投资分析，我们也想用同样的答案来回答——因为他们存在着。

市场数据的原始形态是以时间间隔不规则的高频数据的形式所存在的，每一笔交易或报价都可看作是信息的逻辑单元。高流动性的资本市场每日提供上亿笔的交易和报价信息，不同程度地反映当日市场参与者对市场的观点及看法。

目前，已有越来越多的投资者需要参考日内交易数据做出投资决定，高频数据为他们提供了最佳的信息平台，理应成为最重要的金融分析对象之一。然而，大部分的金融学术期刊研究仍集中于低频率、间隔等距的时间序列。对此，我们认为主要存在有两方面的因素。第一，高频数据的收集、处理、存储、调用以及应用的成本较大，这也是为何目前绝大多数的金融数据存储形式以单个交易日或者更低的频率为单位。第二、大多数的统计方法和原理集中于齐次的、等间隔的时间序列，对于事件发生间隔随机的时间序列的研究较少。令人遗憾的是，这些存在于金融市场的原始数据均不是等间隔的形式，需要人工处理后才能变成齐次的时间序列。直到近年来，随着计算机科技的超速发展，数据的获取与处理问题日趋简化，高频数据的应用才得到推广。目前，国际上大多数交易所和数据运营商都已能够提供高频的分笔数据，NYSE 提供的 Level 2 TAQ 数据就包含了纽交所和纳斯达克所有上市公司的每笔交易数据和报价单。

高频数据代表着一个超大容量的数据群。在一个流动性较好的市场，单个交易日的高频数据量大致相当于 30 年的按日统计的交易数据量 (Dacorogna et al. (2001))。单纯从统计角度来看，独立观测值的数量越多，自由度也就越多，从而估计的参数也越精确。大批量的数据不仅可以提供更多的市场信息，也能够让我们更好地比较模型、区分不同模型的优劣。而且，一些低频数据所不适用的统计方法在高频数据分析中也如鱼得水，如基于尾部分布的极端事件概率。按定义，极端事件的发生是罕见的，因此统计极端事件的发生概率通常会遭遇数据样本的局限性问题。但在高频数据分析中，分析师可以轻易地调用百万数量级的数据对此类事件进行分析。与此相类似，在对不同模型的优劣进行排序时，通过对大量的数据进行拟合测试，由此得出的结论也更加精准。

2. 数量模型取代主观判断，减少非理性的判断失误

传统投资方法结合基本面分析和技术面分析，分析模式多不固定且需要植

入主观判断方式。投资分析师根据自己的经验和知识，收集来自于各种渠道的数据，应用各类估值比较模型对市场和特定的证券产品进行分析解读。这种分析方法被市场认可并已沿用了上百年之久，无论在发达国家的资本市场（如美、英等国）或是新兴资本市场（如中国）都被广泛地接受。

进入上世纪 90 年代，随着计算机科技的普及，金融分析软件被大量地应用于证券市场分析，随后量化的投资方法开始崭露头角，并以其出色的表现得到市场的认可。相对于传统的投资研究方法，量化投资更偏重数据分析，以数量模型为基础，将客观的模型信号作为投资决策。因此，量化投资的研究方法剔除了人为的主观判断，能避免分析师受市场非正常波动的影响所做出的非理性决策。更重要的是，基于高频数据的量化模型能充分利用市场发布出的每一道信息，为分析师描述出更完整的市场状况，从而减少了因信息收集的失误或不完整所造成的错误判断。

3. 利用计算机完成交易策略的执行

传统的量化模型跟据市场变化提供买入、卖出或平仓的信号，而基于高频数据的数量模型可以提供系统而完整的投资决策。在高速计算机上运行的量化模型其参数设置可具体到仓位的头寸大小，止损点位，冲击成本的估计，甚至报价单的价格位设置和实时的风险控制。从一定角度上看，高频交易模型几乎可以完全替代人脑，成为名副其实的赚钱机器 (Money Machine)！

4. 算法交易和程序化交易降低交易成本

1) 算法交易

投资者在进行单笔大金额交易时，通常会面临两大难题：首先，市场的流动性是否提供足够的交易空间，如若流动性不足是否会提高交易成本，导致投资丧失潜在的利润空间。其次，集中的单笔大金额交易易被其他投资者察觉，提供市场额外的信号，其他投资者可利用这些信号挖掘投资机会，从而降低了潜在的投资回报空间。量化分析师利用高频数据所具有的庞大信息优势，提出了算法交易。这类模型利用高频数据拟合各类冲击成本的估算模型并优化交易执行的策略，使得单笔大金额交易能够在不被察觉的情况下被市场所消化，并有效降低此类交易所承受的各类成本（比如冲击成本、机会成本）。

2) 程序化交易

衍生品市场的发展为投资者提供了越来越多的风险套利机会。指数套利、期货和现货套利，期权和现货套利等投资策略时常涉及各类资产在同一时段的多笔交易。对于这种涉及短时间内多笔交易并行的投资策略，交易执行的精确度将严重影响到交易策略的成功性，因此交易的实施细则尤为重要，通常需要借助计算机来完成，程序化交易就由此诞生。基于高频数据的分析，程序化交易可决定怎样完成数十个或几十个股票的同时买卖，以

期减少买卖股票之间的相互作用和交易的误差。

5. 高频交易创造高回报投资机会

高频交易是一种计算机化交易策略的执行，其主要特征是持有仓位的时间极为短暂。在高频交易中，精心设计好的程序在高速计算机上运行，对实时的市场数据进行分析。程序中的各类算法用于捕捉市场上可能存在的持续时间从零点几秒到几小时之久的交易机会。高频交易市场的竞争通常围绕着交易算法和交易执行的速度以及看似微薄但极其稳定的利润来源，并不像长期投资者通常寻找持有期达数周，数月或数年的投资机会。因此，高频交易已被证明有潜力获得比传统买入并中长期持有的策略高出几千倍的夏普比率！其高回报、高胜率、低风险的特征是华尔街各大投资银行、对冲积基金所追逐的对象。2008 年全球金融市场受金融风暴的袭击，各大机构损失惨重，但高频交易却成为市场中唯一的大赢家。

典型的高频交易策略：高频交易策略的成功很大程度上是因为模型能够同时处理大量信息，而这是普通交易员无法做到的。目前全球市场上常见的高频交易策略有如下四类：做市商策略，磁带记录阅读交易，事件套利以及统计套利(Aldridge, Irene (2009))。我们在此对这些策略进行简单的介绍：

做市商策略：其一般的操作方法是连续放置少量的高于市场价格的卖出限价单和低于市场价格的买入限价单以期在买卖方都能得到快速成交，赚取买卖价差获取盈利。一些高频交易公司以这种策略为主要的交易策略，比如“Automated Trading Desk”于 2007 年 7 月被花旗集团收购，就是一家一直在市场中活跃的做市商策略的交易公司，其交易份额在纳斯达克和纽约证券交易所均占到总量的 6%。

磁带记录阅读交易：报价和交易量等市场数据中经常隐含着许多不易发现的信息；通过观察信息流，高频交易程序可提取出一些市场尚未公布的信息。由于所有的报价和交易量信息是公开的，所以这种策略并不违反目前所适用的法律规则。过滤器交易策略是此类交易策略中较为原始的一种高频交易策略，该策略涉及到大量的库存数据监测，观察重大或不寻常的价格变动或成交量活动。这种策略检测的条件包括公告、新闻或其他有影响事件，程序按照事件的性质生成买入或卖出信号。

事件套利：事件套利是指专门针对某类事件发生后市场反映所设计的交易策略。这种事件可以是经济事件也可以是某一行业的特殊事件，但其前提是事件发生后证券市场的价格将产生显著的变化。事件套利的目标是在事件发生前构建一种投资组合，以期在事件发生时点周围产生盈利。

统计套利：此类高频交易策略在各类资产市场中搜索资产价格之间的临时偏差。这种偏差通常具有可预测性，且价格暂时与历史数据的统计关系相背离。统计套利策略被广泛地应用于各类高流动性的证券产品市场，包括股票、期货、外汇等。据美国“TABB 集团”的估计，高频统计套利策略目前每年的总利润超过 210 亿美元。

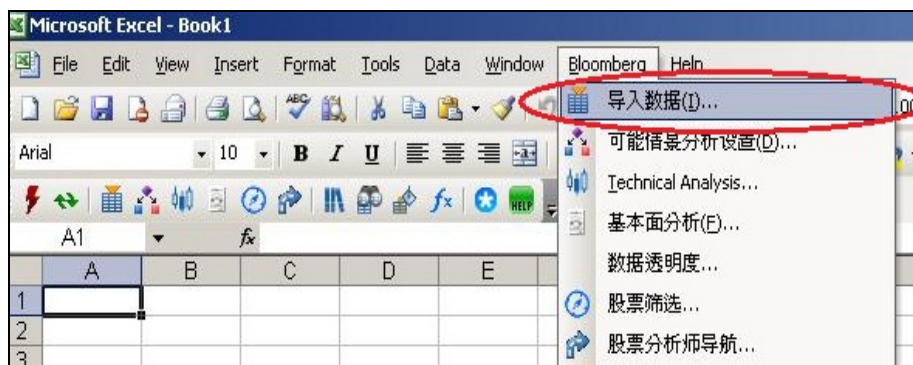
二、 沪深 300 股指期货高频数据来源与形式

1. 高频数据的来源

目前，中金所发布的股指期货高频数据可以通过各类证券行情软件和期货交易软件查看，但一般的软件不提供数据下载，而可提供下载的软件数据也往往止于分钟数据。为了更好地对这些高频数据进行下载，或得到实时的高频数据传输，一般需专业的数据经营商如 WIND、彭博或路透数据终端的支持。下面提供了从彭博下载高频数据的步骤：

- 1) 在 EXCEL 的 BLOOMBERG 插件内选择导入数据。

图 1：EXCEL 的 Bloomberg 插件数据导入步骤一



资料来源：光大证券研究所、Bloomberg

- 2) 进入彭博数据向导界面的历史日内跳价，输入合约代码。

图 2：EXCEL 的 Bloomberg 插件数据导入步骤二



资料来源：光大证券研究所、Bloomberg

3) 在参数界面选择需要的选项。

图 3: EXCEL 的 Bloomberg 插件数据导入步骤三



资料来源：光大证券研究所、Bloomberg

4) 下图为下载到 EXCEL 的数据格式

图 4: EXCEL 中 Bloomberg 的股指期货高频数据下载结果

Date	Type	Price	Size
4/21/2010 9:14:00	TRADE	3215	57
4/21/2010 9:14:00	BID	3214.8	3
4/21/2010 9:14:00	ASK	3215	1
4/21/2010 9:15:00	TRADE	3215	4
4/21/2010 9:15:00	BID	3215	41
4/21/2010 9:15:00	ASK	3215.8	1
4/21/2010 9:15:01	TRADE	3215	42
4/21/2010 9:15:01	BID	3215	4
4/21/2010 9:15:01	ASK	3217.8	1
4/21/2010 9:15:01	TRADE	3217.8	1

资料来源：光大证券研究所、Bloomberg

彭博终端免费提供滞后 15 分钟的高频数据, 若想获得实时数据需在终端上购买。购买流程为:

1) 在彭博终端上输入“EIS”，寻找中金所代码“CFFE”。

图 5: Bloomberg 股指期货实时高频数据下载订购步骤一



资料来源：光大证券研究所、Bloomberg

- 2) 选择申请产品，按照里面的申请步骤就可购买了，中国境内的订购费为每月 5 美元，中国以外的订购费为每月 24 美元。

图 6: Bloomberg 股指期货实时高频数据下载订购步骤二



资料来源：光大证券研究所、Bloomberg

下面是下载 WIND 高频数据的步骤：

- 1) 在 WIND 终端的工具栏选项下选择“高频行情数据终端”。

图 7: WIND 股指期货高频数据下载步骤一



资料来源：光大证券研究所、WIND

- 2) 在 WIND 资讯高频行情数据界面点击“新建方案”，在下载品种中选择“期货”。

图 8: WIND 股指期货高频数据下载步骤二



资料来源：光大证券研究所、WIND

- 3) 在范围内选择中金所品种，选中所要看的合约，在数据类型那选择盘口报价，在字段选择中选择所需要的指标。

图 9：WIND 股指期货高频数据下载步骤三



资料来源：光大证券研究所、WIND

- 4) 输出的 CSV 格式如图 10 所示。WIND 仅免费开放近 10 日的高频数据下载，如若需要更早数据，则需每年缴付额外的订购费。

图 10：WIND 股指期货高频数据下载结果

A	B	C	D	E	F	G	H	I	J	K	L	M
SecID	SecName	Date	Time	NewPrice	SellPrice	BuyPrice	SellVol	BuyVol	TotalVol	TotalAmount	PrePosition	Position
IF1011.CFE	IF1011	2010/11/3	9:14:00	3533	3534.4	3533	23	40	139	147326100	27851	27898
IF1011.CFE	IF1011	2010/11/3	9:15:00	3533	3533	3530.4	42	27	285	302062620	27851	27903
IF1011.CFE	IF1011	2010/11/3	9:15:01	3533	3532	3530.4	1	13	349	369883800	27851	27896
IF1011.CFE	IF1011	2010/11/3	9:15:01	3532	3532.6	3532	2	2	421	446190660	27851	27920
IF1011.CFE	IF1011	2010/11/3	9:15:02	3532	3530.4	3530	5	517	487	516125400	27851	27887
IF1011.CFE	IF1011	2010/11/3	9:15:02	3532.2	3533	3532.6	13	16	514	544730220	27851	27879

资料来源：光大证券研究所、WIND

2. 数据存储的形式

中金所目前发布的高频数据是经过其特殊算法处理后的数据，其最高频率达到每秒两次信息的传输，其间隔为等距的 500ms。每次信息的传输包括之前 500ms 内所有成交的加权成交价和成交量、最新时刻的买卖报价和报价量以及开盘后累计的持仓量。报价信息在同一时间公布的形式分为五档，即买一、买二至买五和卖一、卖二至卖五以及各档的报价量。

三、 针对沪深 300 股指期货高频数据微观结构的探索

1. 样本选取

在综合比较各个数据供应商所提供的沪深 300 股指期货高频数据的形式后，我们建立了以彭博数据下载为主要来源的高频数据库，并截取当月主力合约前四个交易月（2010/4/16 至 2010/8/20）的数据作为样本进行研究分析。需要指出的是，目前，由于彭博只提供一档报价且不提供持仓量变化，加上成交数据经过中金所处理，因此有相当的市场结构层面无法进行有效分析。在这里我们只对有限的信息加以研究，对样本内的数据进行微观结构的探访，以帮助读者初步了解股指期货市场微观方面的信息。日后，我们将争取整合更完整的高频数据库，对市场进行更完整的结构描述。

2. 样本内价格回报特性

高频数据在某些统计上的表现与传统低频数据有较大区别，它为投资者开拓了一个新的探索研究的方向。一般而言，价格回报率是研究分析的主要对象。回报率通常由连续两笔价格的变化除以第一笔价格获得，但学术研究学者则认为对数回报率更能体现价格回报的分布特性，因此我们这里的回报计算采取对数回报的方法。表 1 给出了样本内分笔数据以及以 1 分钟、5 分钟、10 分钟为观测间隔的对数回报率统计值。在统计过程中，我们删除了隔夜回报率和开盘集合竞价时的回报率以保证回报率所代表的价格变化都是由连续竞价造成的。

表 1：股指期货不同频率下的回报率统计

	最小值	中位数	平均值	最大值	标准差	偏度	峰度
分笔数据	-1.07%	0.00%	0.00%	1.07%	0.01%	-0.0002	1.788
1 分钟间隔	-2.62%	0.00%	0.00%	2.69%	0.22%	-0.039	1.739
5 分钟间隔	-3.46%	0.00%	0.00%	2.34%	0.32%	-0.057	1.725
10 分钟间隔	-3.25%	0.00%	-0.01%	2.50%	0.31%	-0.067	1.541

资料来源：光大证券研究所、Bloomberg

从表格中我们可以归纳出以下几点价格回报率特性：

- 1) 当观测频率较高时回报率分布平均值和中位数都趋向于零，也就是说期指价格在短期波动分布的均值可近似取零。
- 2) 当观测频率越高时回报率分布的偏度越小、正负极值的绝对数越相互接近，即回报的对称性在高频率数据中得到满足。
- 3) 当观测频率越高时回报率分布的峰度越大，即频率越高回报率出现厚尾的机率越大。
- 4) 方差作为价格波动率的估计，并不存在随观测时间间隔的线性增长。显示当观测间隔增加时、观测周期内存在价格变动的自相关性。

根据这些高频数据的特性，研究者可以选用相应的随机分布函数对股指期货市场进行仿真模拟，以达到各自测试研究的目的。

3. 成交价日内走势

沪深 300 股指期货于 2010 年 4 月 16 日开始交易，作为国内最大的金融衍生品市场，由于机构投资者尚未完全进入，市场内的交易以投机性占主导地位。通过观察主力合约的日内走势，可以明显发现期指在日内交易中时常出现急涨急跌的状态，其波动幅度大于现货的波动幅度，与 A 股个股的日内波动行情更为相似。

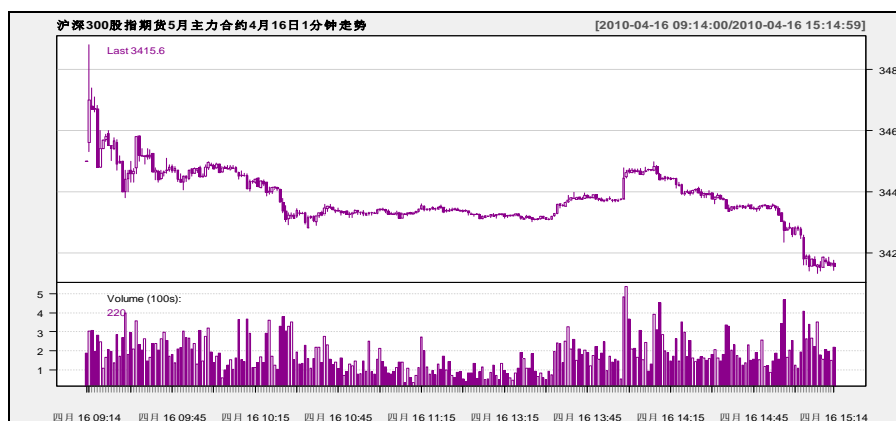
图 11 至图 16 给出了期指任意两个交易日的日内走势，从分笔数据或以 1 分钟、5 分钟为间隔的观测数据都可发现，期指日内波动以短期趋势为主，市场的涨跌往往在短短几分钟内完成。

图 11：5 月主力合约 4 月 16 日分笔数据价格走势

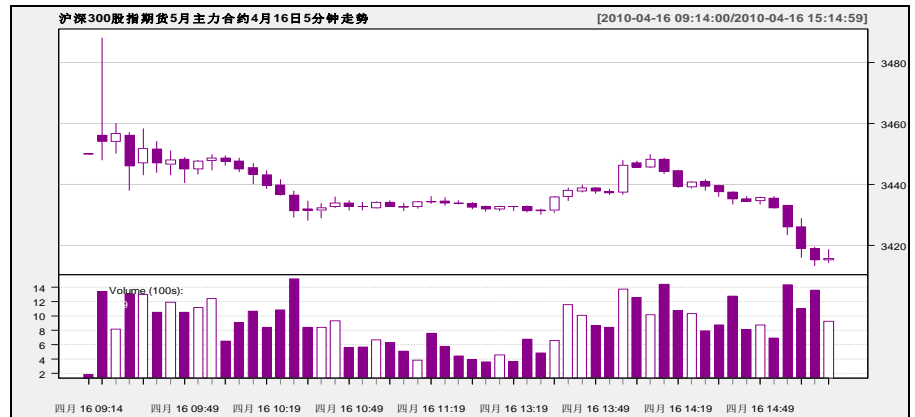


资料来源：光大证券研究所、Bloomberg

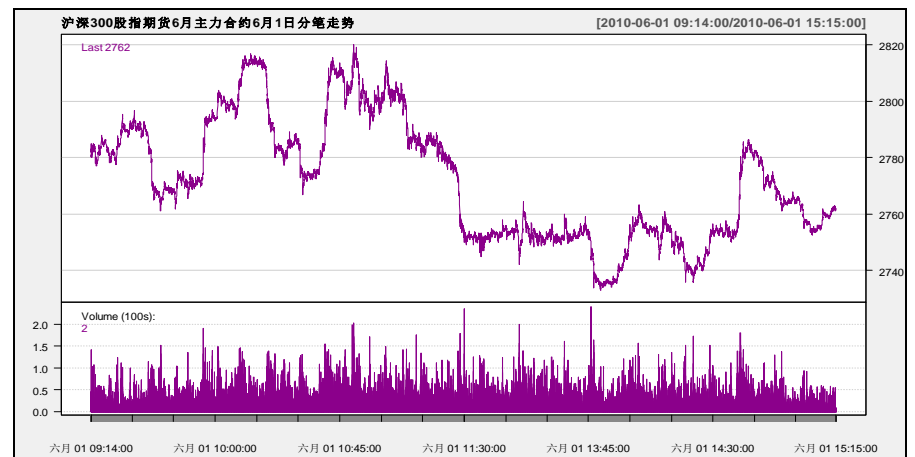
图 12：5 月主力合约 4 月 16 日一分钟间隔价格走势



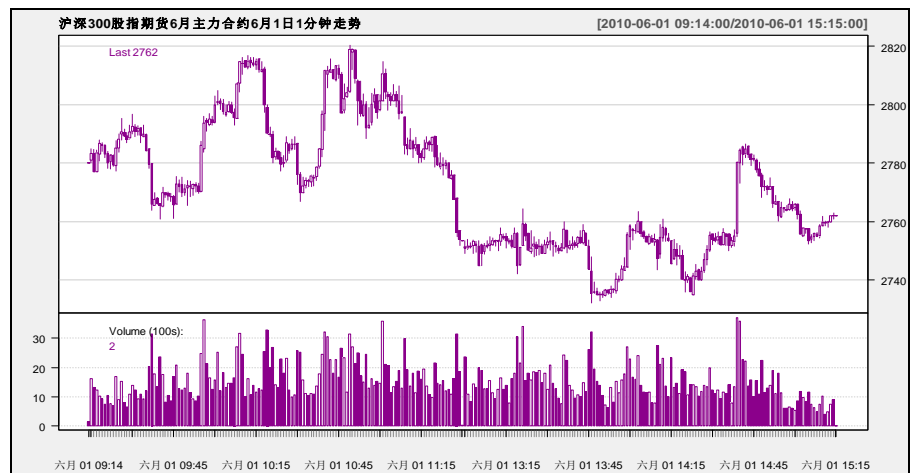
资料来源：光大证券研究所、Bloomberg

图 13：5 月主力合约 4 月 16 日五分钟间隔价格走势

资料来源：光大证券研究所、Bloomberg

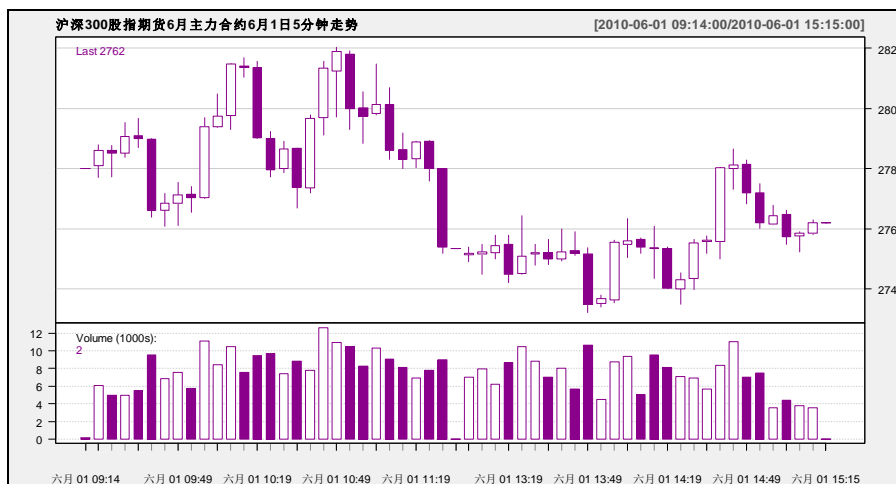
图 14：6 月主力合约 6 月 1 日分笔数据价格走势

资料来源：光大证券研究所、Bloomberg

图 15：6 月主力合约 6 月 1 日一分钟间隔价格走势

资料来源：光大证券研究所、Bloomberg

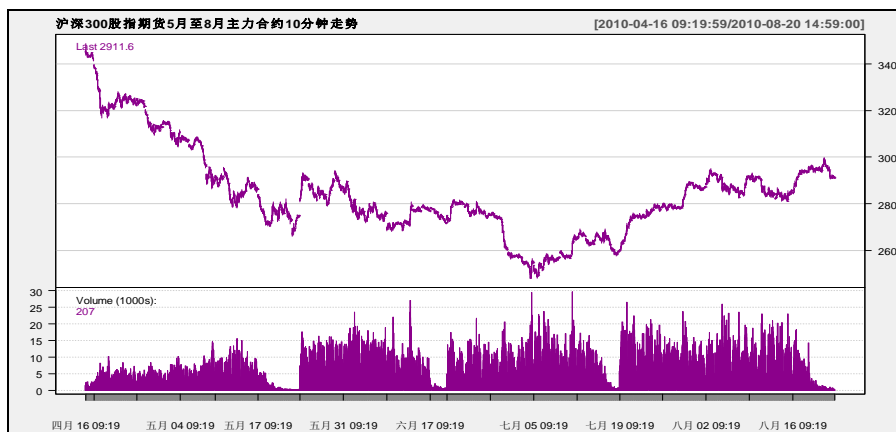
图 16：6 月主力合约 6 月 1 日五分钟间隔价格走势



资料来源：光大证券研究所、Bloomberg

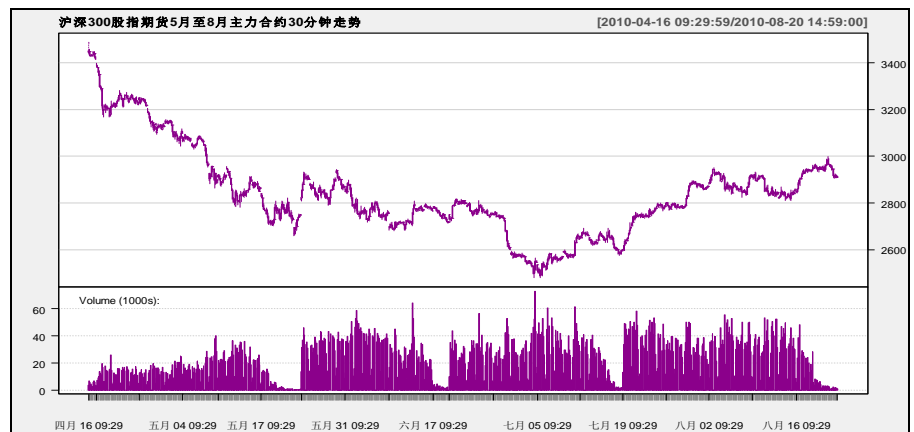
即使我们放大观测周期、扩展观测样本，我们仍能发现期指在短期内的疾速波动可以解释大部分的市场变化。图 17 至图 19 给出股指期货 5 月至 8 月主力合约以 10 分钟、30 分钟、60 分钟为观测周期的走势。我们认为股指期货这种短期剧烈波动的形态是市场非有效性的表现，市场信息掌握在少数投资者手中，一旦这些投资者获得有效信息将会快速做出投资决策，拉伸指数（无论是 A 股或是期指）。对于这类非有效市场，传统投资方法时常手足无措，但对于高频交易策略，其快进快出的方法却有大展身手的机会。

图 17：5 月至 8 月主力合约 10 分钟间隔价格走势



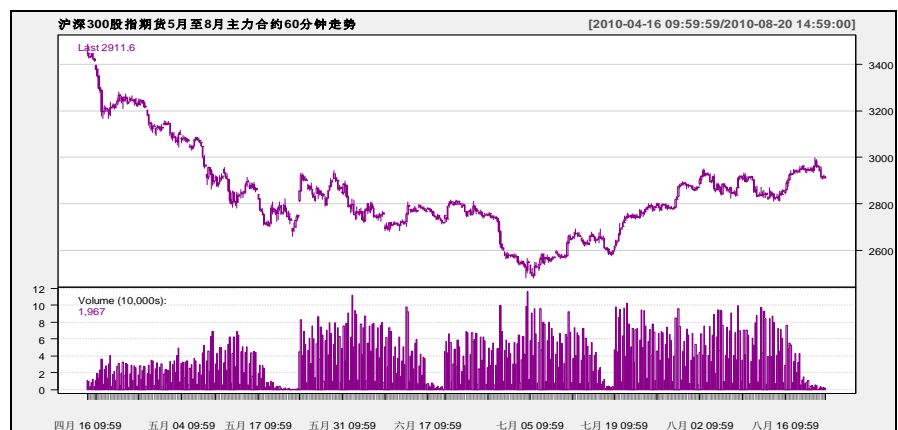
资料来源：光大证券研究所、Bloomberg

图 18：5 月至 8 月主力合约 30 分钟间隔价格走势



资料来源：光大证券研究所、Bloomberg

图 19：5 月至 8 月主力合约 60 分钟间隔价格走势

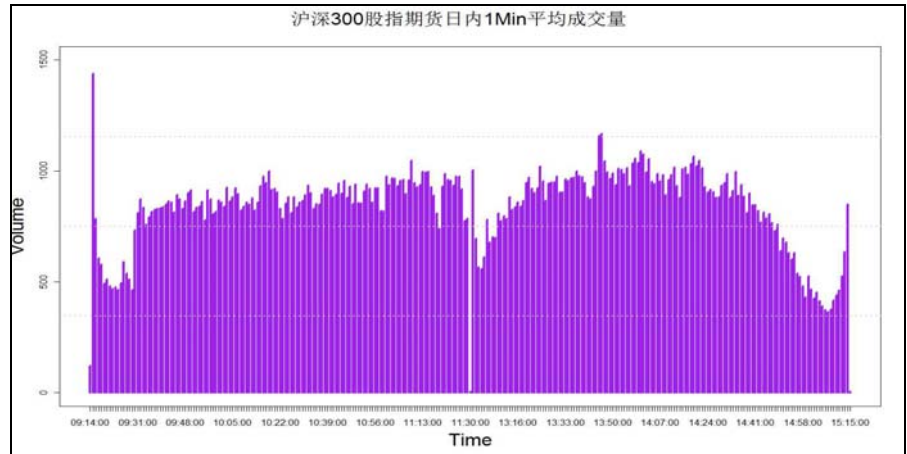


资料来源：光大证券研究所、Bloomberg

4. 成交量日内走势

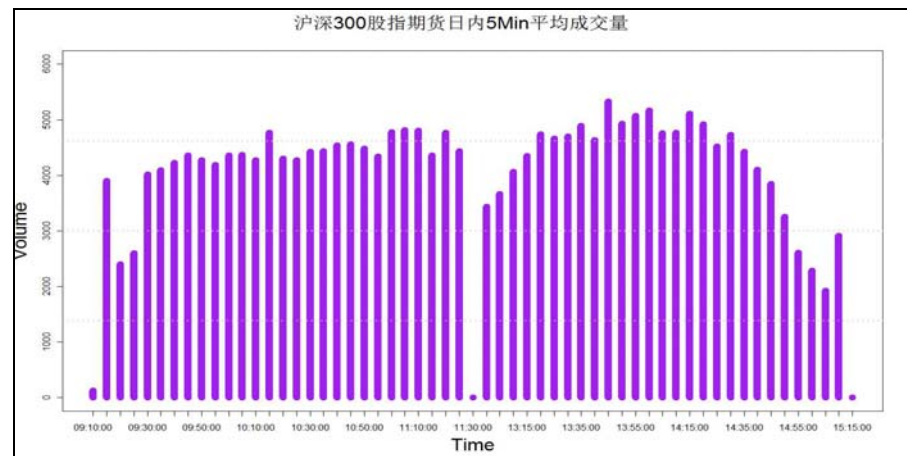
成交量对于高频交易策略起着举足轻重的作用，特别是超高频交易模型需要在短时间内拥有充足的成交量给与支持。沪深 300 股指期货单个交易日的成交量走势具有一定的周期性。由于期货开盘先于现货，收盘也晚于现货，在没有现货价格支撑的情况下，期指成交量在上午开盘和下午收盘时普遍较低。此外，在上午临近收盘和下午刚开盘的时刻其成交量也显著低于日内平均值，体现出投资者通常短线操作，在午间休市时期较为谨慎的投资心态。值得注意的是，成交量在下午收盘前的 45 分钟内呈现稳步减少的态势，而在最后 5 分钟内又有显著拉升。我们认为以投机为主的投资者大多选择在现货交易最后半小时平仓或者选择在期货交易的最后 5 分钟平仓，以避免持仓过夜的风险。图 21 至图 22 给出样本内主力合约的日内一分钟和五分钟成交量的平均值。

图 20：5 月至 8 月主力合约日内一分钟平均成交量



资料来源：光大证券研究所、Bloomberg

图 21：5 月至 8 月主力合约日内五分钟平均成交量

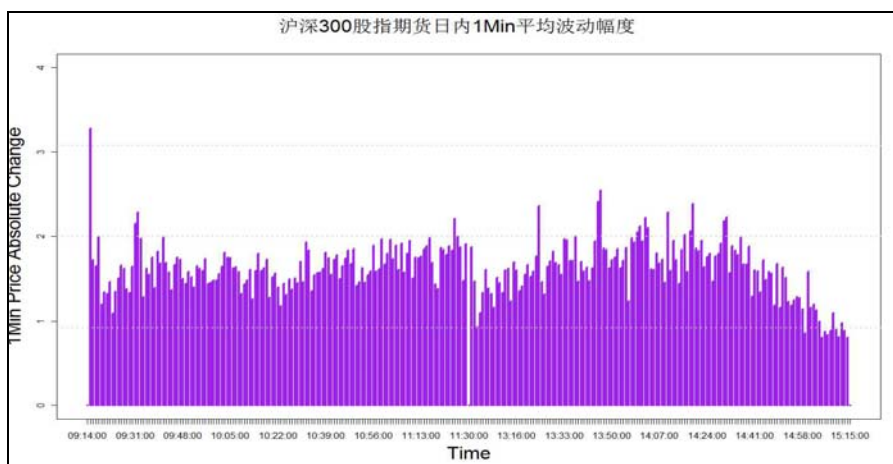


资料来源：光大证券研究所、Bloomberg

5. 成交价格日内波动幅度走势

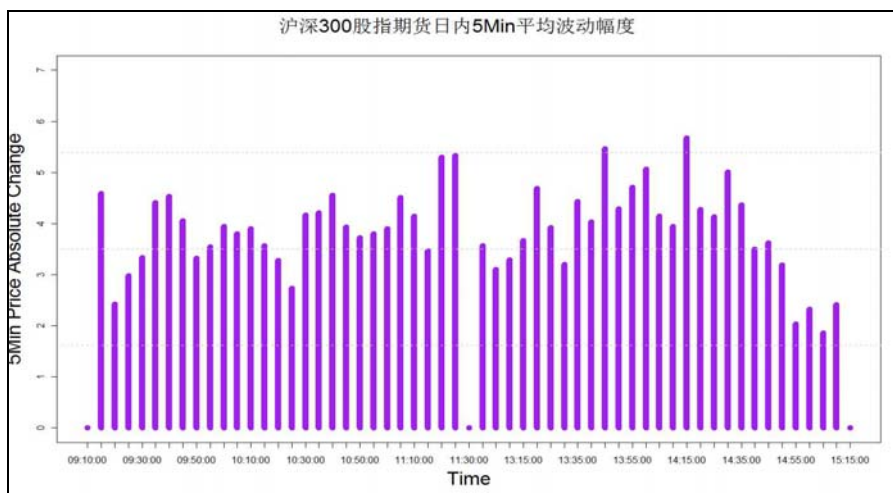
股指期货价格的日内波动幅度同样具有周期性。开盘时由于集合竞价的原因价格扰动通常较大，其后随着成交量的萎缩其波动幅度开始缩小，直到现货开盘后价格波动幅度维持在一个比较均衡的状态。下午开盘后受成交量的影响其波动幅度一开始也处于低位，随时间推移和成交量的逐步放大，波动幅度也返回正常状态。当成交量在最后 45 分钟逐渐缩小时，波动幅度也如预期的逐渐降低。从成交量和日内波动幅度走势来看，高频交易更适合在期货与现货同时交易的时刻进行。此段时刻不仅有充足的成交量，且波动幅度较大，盈利空间也更充足。图 22 至图 23 给出样本内主力合约的日内一分钟和五分钟波动幅度的平均值。

图 22：5 月至 8 月主力合约日内一分钟平均波动幅度



资料来源：光大证券研究所、Bloomberg

图 23：5 月至 8 月主力合约日内五分钟平均波动幅度

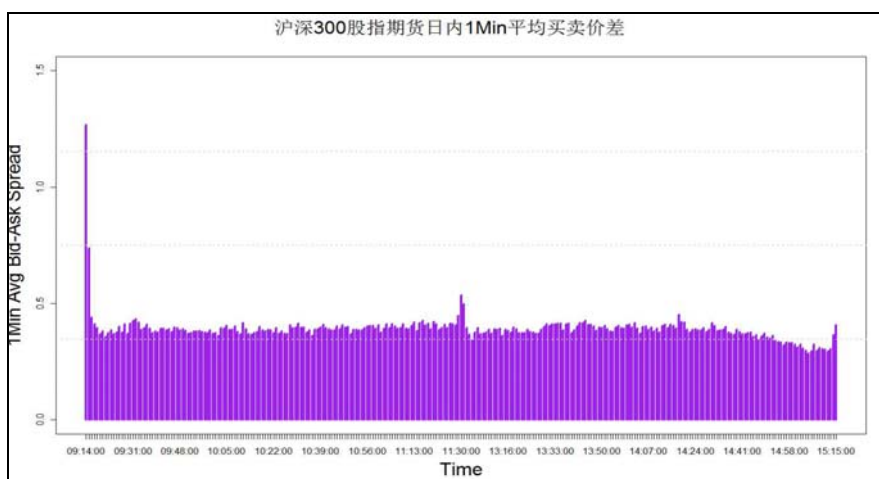


资料来源：光大证券研究所、Bloomberg

6. 买卖价差日内走势

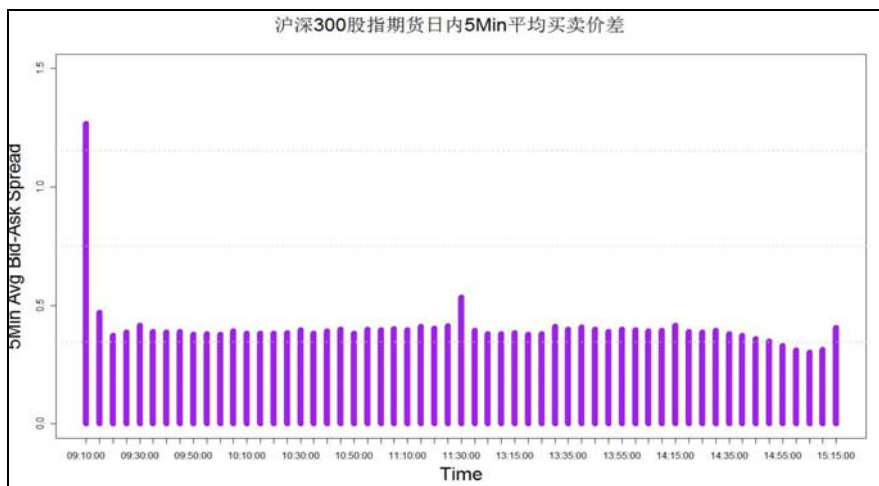
买卖价差通常被投资者当作市场流动性的指标。当价差较大时交易成本上升，交易机会将被交易成本削弱，因此价差的大小需要在交易执行的过程中被严格的监控。经初步分析，沪深 300 股指期货主力合约的买卖价差在日内的波动并不显著，除上、下午开盘和收盘时价差有显著不同，其余时间价差都较为稳定。图 25 至图 26 给出样本内主力合约的日内一分钟和五分钟平均买卖价差。

图 24：5 月至 8 月主力合约日内一分钟平均买卖价差



资料来源：光大证券研究所、Bloomberg

图 25：5 月至 8 月主力合约日内五分钟平均买卖价差

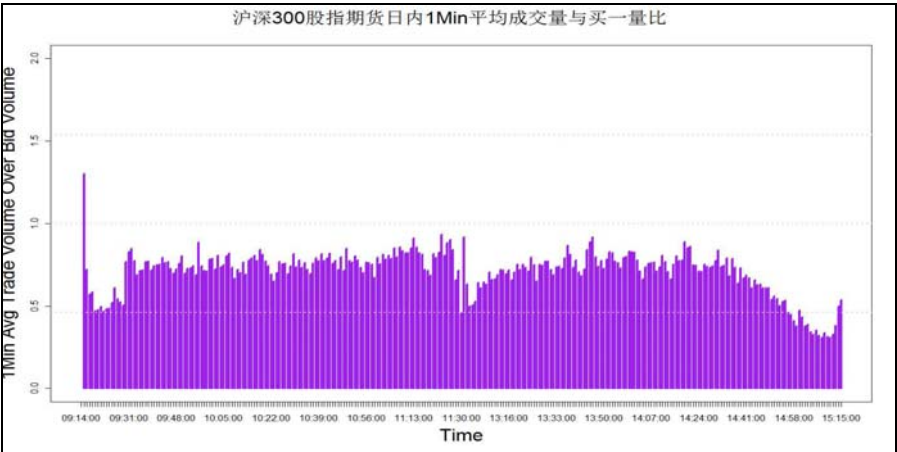


资料来源：光大证券研究所、Bloomberg

7. 成交量与报价量日内平均比例

在关注市场流动性的时候,我们通常需要比较成交量和报价量之间的关系,如成交量经常高于最优报价量,则代表市场容量不足,此时大额订单的成交需要通过与多档的报价配对来完成,交易的冲击成本较大,市场不支持快速的大额成交。反之,如成交量远小于最优报价量,则代表市场容量充足,冲击成本不阻碍短期的大额成交。图 26 至图 27 比较了一分钟内平均的成交量和最优买卖报价量之比。从走势来看,成交量与买卖报价量的比例基本复制了日内成交量的变化,可见日内的报价供应量基本维持在稳定的水平。从比例大小来看,日内成交量与买卖报价量除开盘集合竞价时,都小于但接近于 1,可见市场的供应量并不十分充足,短期的大额成交将面临冲击成本的挑战。

图 26：5 月至 8 月主力合约日内一分钟平均成交量与买一量之比



资料来源：光大证券研究所、Bloomberg

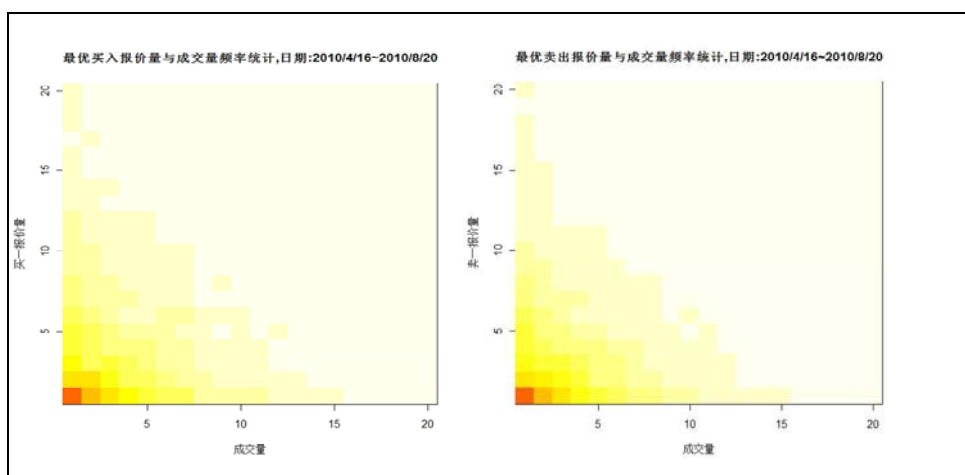
图 27：5 月至 8 月主力合约日内一分钟平均成交量与卖一量之比



资料来源：光大证券研究所、Bloomberg

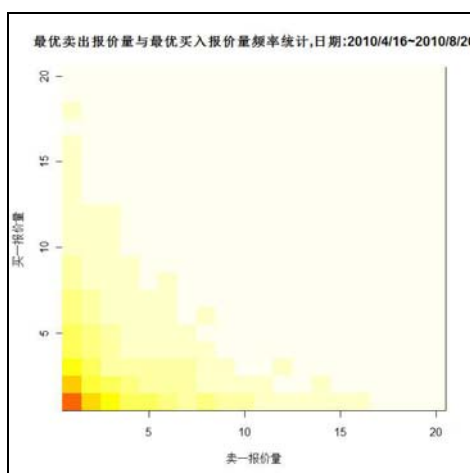
买卖报价量与成交量之间的平衡或失调状态通常蕴藏着市场潜在的方向性信息。当买方或卖方的报价量被成交量占据主动时，市场将出现短暂的单边走势，造成市场价格的移动。图 28 至图 29 对成交量与报价量之间的平衡关系做了频率统计，图中区域颜色越深代表发生的频率越高。由图可见成交量与报价量失衡的状态并不频繁，报价量高于成交量的频率更高，而买卖报价量之间也多处于平衡的状态。

图 28：5 月至 8 月主力合约成交量与最优买卖报价量频率统计



资料来源：光大证券研究所、Bloomberg

图 29：5 月至 8 月主力合约最优买入报价量与最优卖出报价量频率统计



资料来源：光大证券研究所、Bloomberg

8. 高频波动率计算

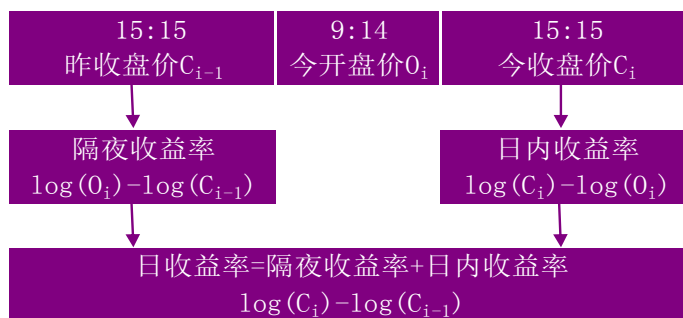
波动率作为传统投资中对投资风险的衡量，在量化投资中起着举足轻重的作用。在业界，人们通常利用标准差来测量波动率的大小，投资回报的标准差越大，所投资的证券的风险也就越大。然而传统的波动率计算方法通常只用到日收益率，且计算的结果为证券回报的历史波动，不具有前瞻性，这些不足阻碍了投资决策的制定。高频数据的利用为波动率的计算带来了新的方法，下面我们就传统波动率计算方法和利用高频数据计算方法对股指期货的波动率计算做一简单介绍和比较。（我们将于后续的深度报告中具体介绍基于高频数据开发的股指期货波动率以及其在高频交易中的应用。）

1) 传统波动率估计方法

传统的计算方法取出股指期货过去一段时间内（比如 22 天）的日收益率，

计算这段时间内收益率的标准差作为波动率的估计。一般情况下，日收益率是根据股指期货指数当天某个节点的价格与昨天对应节点的价格计算出来的。如果以对数价格变动计算日收益率的话，日收益率可以分解为隔夜收益率和日内收益率，即 R 日收益率 = R 隔夜 + R 日内。当隔夜信息较少时，股指当天开盘价与昨收盘价大致相同，隔夜收益率基本为 0，不会对当天日内波动率产生多大影响；但是隔夜如果发生重大利好或利空消息，股指当天会大幅高开或低开，隔夜收益率对日收益率的影响不能被忽略，这样导致估计出的波动率里包含了隔夜波动率、隔夜波动率对当天日内波动率的影响以及当天的日内波动率。传统波动率估计方法对这三部分构成的分解通常无能为力，因此在实际应用中常会捉襟见肘。

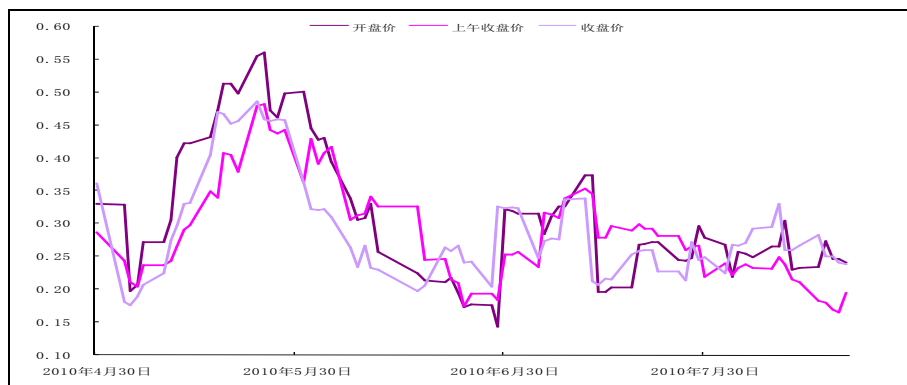
图 30：日收益率分解图



资料来源：光大证券研究所

在一般情况下，利用日收盘价来计算日收益率，同样也可以利用开盘价、上午收盘价来计算收益率，甚至可以用相隔两天对应的任意时间点来计算收益率，收益率计算方法的多样性为我们检验方差与标准差的稳健性提供了一个很好的手段。图 31 为利用开盘价、上午收盘价、收盘价计算的收益率所对应的光大股指期货主力合约连续指数的标准差。观察此图可以发现，不仅在波动率的大小水平上三种方法相差很大，而且在各自变化趋势上也不尽相同。对于一般投资者而言哪个方法所计算的波动率更真实、更符合实际情况也是值得商榷的问题。

图 31：光大股指期货主力合约连续指数历史波动率比较

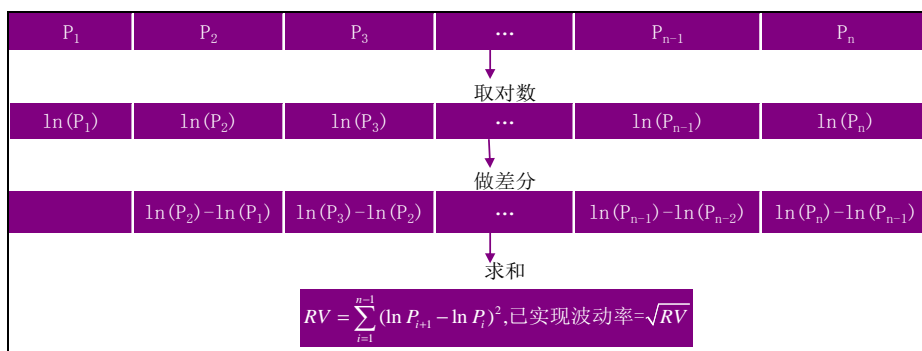


资料来源：光大证券研究所、Bloomberg

2) 传统日内波动率的估计方法：已实现波动率（Realized Volatility, RV）

金融高频数据包含的市场信息远远超过低频数据，由此基于高频数据估计的波动率比低频数据更为准确。已实现波动率（英文简称 RV）凭借其计算方法简便、信息量全面等优点，为多数市场实践者所使用。下图给出了已实现波动率的计算过程。

图 32：已实现波动率（RV）的计算过程



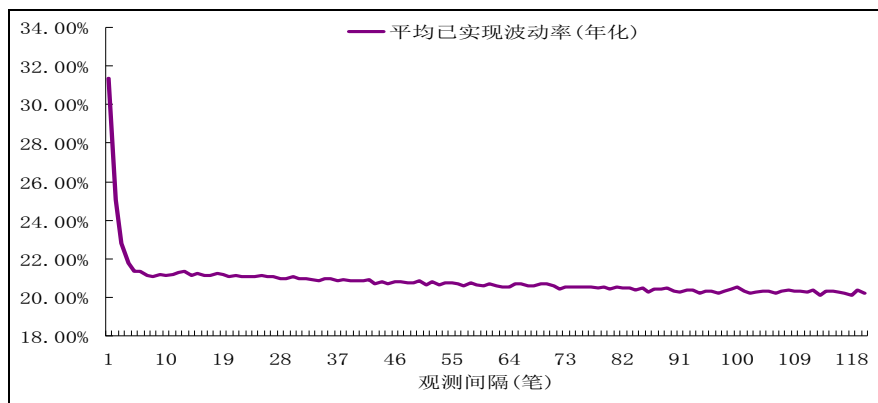
资料来源：光大证券研究所

虽然已实现波动率有效利用了高频数据的信息量，但其模型在假设中有一定的缺陷性。由于价格变化的非连续性、对信息反应的延迟性等市场微观结构噪音，我们在市场上观测到的价格可分解为真实的内在价格加上噪音项。当日内高频的数据量非常大、价格观测的间隔非常小时，市场噪音成为影响价格变化的主要因素，此时已实现波动率（RV）已经不是真实价格波动率的估计，反而成为噪音波动率的估计。根据理论研究，已实现波动率近似等于 2 倍样本量的噪音波动率，即 $RV \approx 2n\sigma_{\text{噪音}}^2$ 、

$\log(RV) \approx \ln(n) + \ln(2\sigma_{\text{噪音}}^2)$ ，其中 n 为样本量、 $\sigma_{\text{噪音}}^2$ 为噪音波动率。图

33 显示，当观测间隔很小时，已实现波动率（RV）非常大，成为噪音波动的估计；当观测间隔增长时，已实现波动率（RV）则迅速下降。

图 33：已实现波动率与样本量之间的关系



资料来源：光大证券研究所、Bloomberg

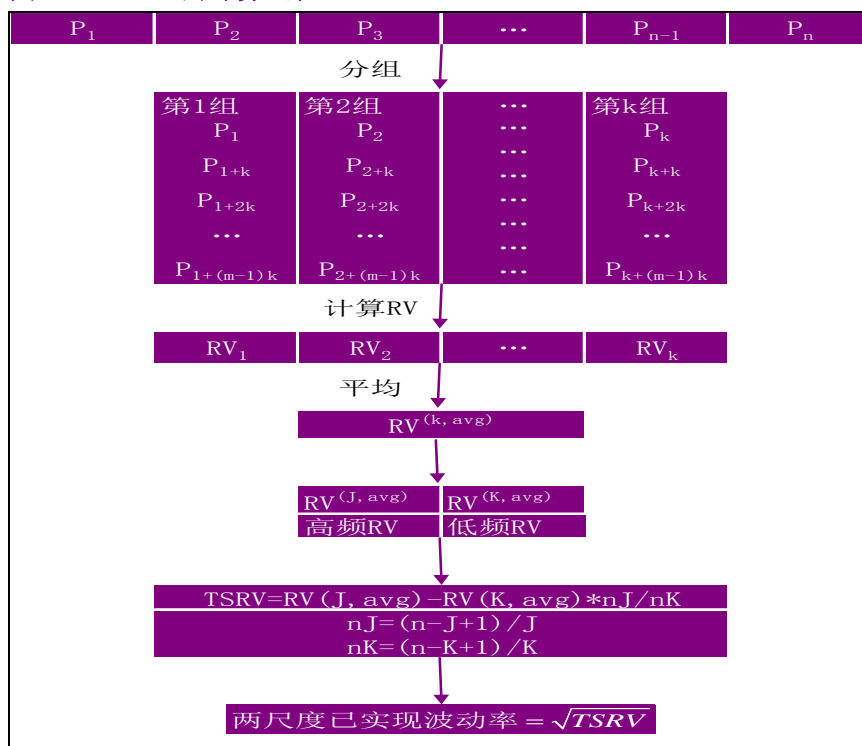
当数据量很大时，已实现波动率（RV）的估计偏差与样本量成正比，因此通过增加观测间隔、减少样本量为降低市场噪音对估计偏差影响的主要手段。但这种方法同样会丢失大量有用信息。对于股指期货的秒钟数据，如果每隔 30 秒钟取一个观测值，那么平均每 30 个观测值就会舍去 29 个数据，很明显，这种做法对市场信息的利用率是非常低的。

3) TSRV (Two Scale Realized Volatility) 日内波动率估计的方法

在超高频下, RV 对波动率估计的偏差非常大;在低频(相对)下, RV 对市场信息的利用率低下。针对 RV 的这些不足, Aït-Sahalia, Mykland, and Zhang (2004) 提出了 TSRV 估计方法。他们的估计方法首先将观测价格分为 K 组, 每组的子样本数量为 m , 组内连续观测值的时间间隔为 K , 然后计算每个子样本的已实现波动率 RV , 平均每个子样本的已实现波动率可得到平均已实现波动率 $RV(K, \text{avg})$; 类似地, 把观测价格再分为 J 组 ($J > K$), 重新计算得到平均已实现波动率 $RV(J, \text{avg})$, 最后通过构造 $RV(J, \text{avg})$ 与 $RV(K, \text{avg})$ 的线性组合得到 TSRV。在计算 $RV(J, \text{avg})$ 和 $RV(K, \text{avg})$ 的过程中, 不仅利用了全部样本信息, 还通过平均分组的已实现波动率保证了 TSRV 的稳健性。

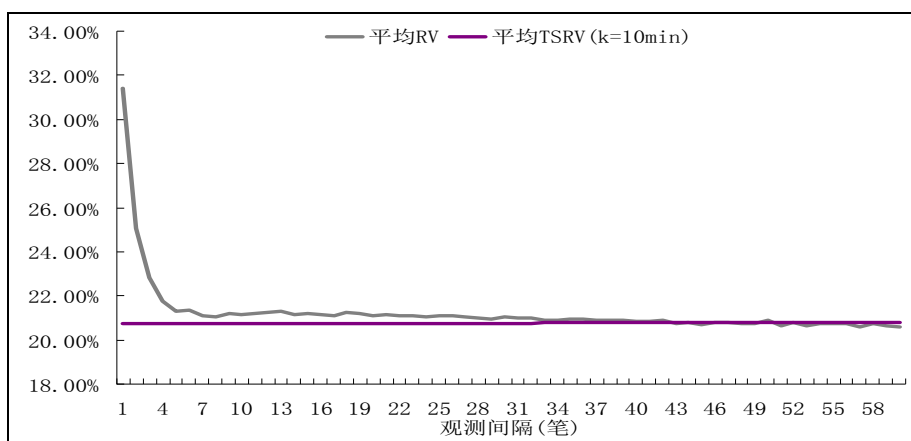
TSRV 利用两个不同观测间隔的平均已实现波动率的线性组合,恰好相互抵消了市场噪音波动率对已实现波动率的影响,使得 TSRV 成为波动率的渐近无偏估计。图 35 比较了 RV 与 TSRV 在不同观测间隔下的效果,在整体水平上,TSRV 有效的降低市场噪音对估计的影响;同时从稳健性来看,TSRV 的波动水平明显小于 RV。

图 34: TSRV 的计算过程



资料来源：光大证券研究所

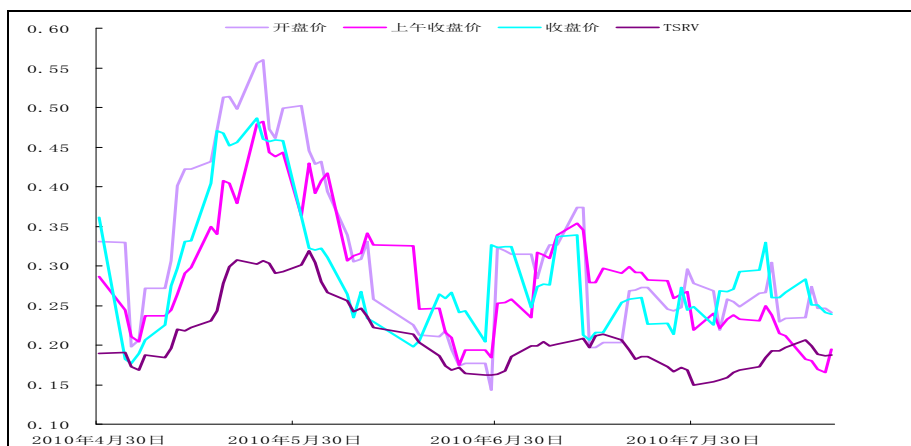
图 35: RV 与 TSRV (K=10min) 的比较



资料来源：光大证券研究所、Bloomberg

相对于传统的波动率计算方法，TSRV 所测量的波动率主要为日内交易时刻的波动率，由于剔除了隔夜信息，其大小通常小于前者。图 36 显示了测试样本内利用开盘价、上午收盘价、收盘价计算的 10 个交易日回报率的标准差与 TSRV 计算的波动率的 10 日平均移动序列。相对于标准差法，TSRV 的估值明显偏小且波动方向与前者也有明显的差异。

图 36: TSRV 与标准差法波动率的比较



资料来源：光大证券研究所、Bloomberg

9. 市场冲击成本模型介绍与应用

1) 冲击成本的形成

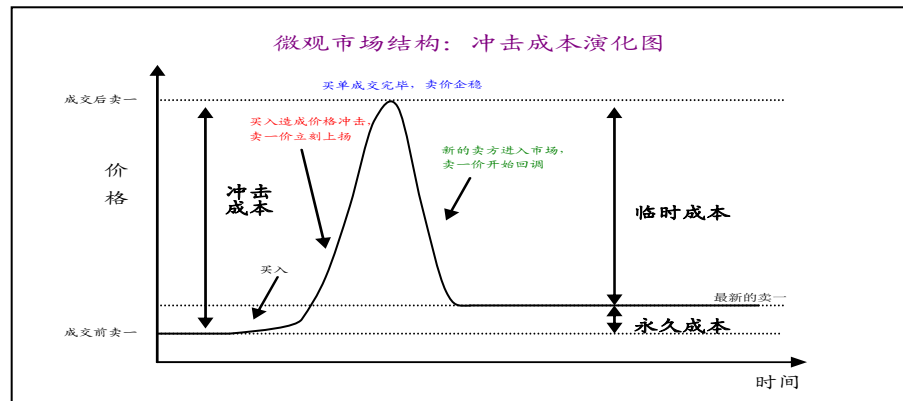
市场冲击成本也可称为流动性成本，是指一定数量的委托订单在迅速成交时对价格产生的影响。通常，冲击成本会对投资者造成负面影响，当投资者希望多头开仓时，冲击成本会造成最终成交价格的上涨。精确的冲击成本可通过真实的成交价格与理想中的市场价格（委托订单如果没有被执行）的比较来获得，但这一般较难以衡量；我们这里的研究则利用平均的成交价格与成交前的报价差作为替代。

冲击成本可以被分为临时性成本和永久性成本：

$$\text{冲击成本} = \text{临时性成本} + \text{永久性成本}$$

临时性成本反映的是由市场流动性所产生的成本。假如一个较大的订单需要在短时间内被执行的话，通常需要和多笔的报价进行配对，导致平均成交价格变高。永久性成本反映的则是订单所造成的长期效应，代表着订单因暴露于市场所携带的可交易信息内容。

图 37：买入成交造成冲击成本的演化图



资料来源：光大证券研究所

冲击成本的构成具有即时性，即冲击成本的发生与委托单成交发生同步。当一笔委托订单被传输进入市场时，需要和报价单进行配对，如最优报价单无法完全满足成交需求时，次优档次的报价单开始工作，以达到完全成交的目的。在成交结束的时刻，市场的报价也开始企稳，冲击成本可以这一时刻的报价与成交执行前的报价进行比较获得。假设在成交之后市场没有新的交易，而报价单则随投资者的意愿及时更新，最终根据之前所有交易信息会达成新的报价状态。新的报价与之前成交完毕时企稳的报价的差额称作为临时性成本，也就是说这部分成本虽然由投资者买单，但其对市场的效应是临时的。而剩余部分的冲击成本则被称作永久性成本，即市场由于成交而造成的真实价格变动。

$$\text{冲击成本计算公式：冲击成本} = \sum (x_j \times (p_j - p_b))$$

x_j 订单以 p_j 价格被执行的成交量

p_j 成交价格

p_b 成交前最优的买/卖报价

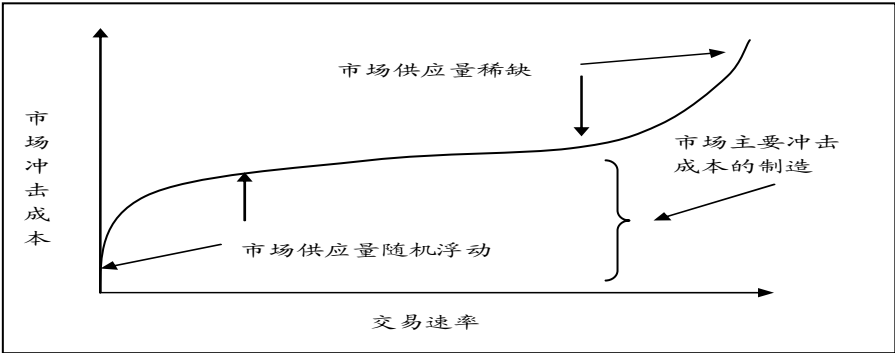
2) 影响冲击成本的因素

冲击成本的大小通常与单笔委托成交量、交易执行速度、市场的波动率成正比，与市场整体成交量成反比。单笔委托成交量越大，潜在需要与成交配对的报价档数越多，因此成本越大；对交易执行速度越高，在交易执行

中其他投资者提供的市场流动性越少，因此成本也越大；当市场波动率较大时，买卖价差和各档的报价差也将拉大，因此交易会受到更大的冲击成本；而市场的成交量作为市场的容量大小则体现了市场交易机会大小，交易机会越大，交易所受的冲击成本越小。

当成交速度不作为交易考虑的主要因素时，成交速度可以放缓，此时影响冲击成本的主要是永久性成本，即与成交量大小线性相关的成本。而当成交速度作为主要考虑因素时，冲击成本中的临时性成本开始扩大。当投资者对成交速度要求非常高时（如超高频交易策略），市场的流动性将会对临时性成本产生很大影响，临时性成本的上升速度开始加快。

图 38：冲击成本与成交速度的关系



资料来源：光大证券研究所

3) 冲击成本计算例证

下表列出了股指期货主力合约某一时刻的交易报价。报价分为卖出委托和买入委托，每个委托各分五档报价。假设，此时有买家进入市场想以市价购买 20 手合约，投资者会发现卖出委托里面的卖一报价量不足以完全配对成交，因此卖出委托的其他报价也将发生配对变化。

表 2：股指期货委托报价（成交前）

卖出委托			买入委托		
	价格	委托量		价格	委托量
卖五	3590	10	买一	3585.4	2
卖四	3589	10	买二	3585.2	64
卖三	3588.8	1	买三	3585	30
卖二	3588.6	4	买四	3584.2	16
卖一	3585.6	10	买五	3583.2	24

资料来源：光大证券研究所

表 3：股指期货委托报价（成交时）

卖出委托			买入委托		
	价格	委托量		价格	委托量
卖五	3589	5	买一	3585.4	2
卖四	3589	5	买二	3585.2	64
卖三	3588.8	1	买三	3585	30
卖二	3588.6	4	买四	3584.2	16
卖一	3585.6	10	买五	3583.2	24

资料来源：光大证券研究所

当交易配对完毕，按照我们前面所介绍的假设没有新的成交发生，会有新的报价进入市场。一段时间后市场报价将会达到新的平衡。

表 4：股指期货委托报价（成交后）

卖出委托			买入委托		
	价格	委托量		价格	委托量
卖五	3586.8	3	买一	3585.4	28
卖四	3586.6	4	买二	3585.2	25
卖三	3586.2	3	买三	3585	42
卖二	3586	16	买四	3584.2	16
卖一	3585.8	2	买五	3583.2	24

资料来源：光大证券研究所

按照冲击成本的计算公式我们可以得到：

$$\text{冲击成本} = 300 \times ((3585.6 - 3585.6) \times 10 + (3588.6 - 3585.6) \times 4 + (3588.8 - 3585.6) \times 1 + (3589 - 3585.6) \times 5) = \text{¥}9660$$

利用 Kissell 和 Glantz 的冲击成本估计模型我们可以近似估算临时性成本和永久性成本（这里假设日均成交量为 23 万手合约）

$$\text{冲击成本} = 0.95 \cdot I_s \cdot \eta^{-1} + 0.05 I_s \Rightarrow I_s = \frac{MI_s}{(0.95\eta^{-1} + 0.05)}$$

其中 $\eta = (X + 0.5 \times Q) / X$ ， X 为成交量， Q 为日均成交量

$$\text{所以 } \eta = (20 + 0.5 \times 230000) / 20 = 5751$$

$$I_s = \frac{9660}{\left(\frac{0.95}{5751} + 0.05\right)} = 192563.8$$

$$\text{临时性成本} = (0.95 \times 192563.8) \div 5751 = 31.8$$

$$\text{永久性成本} = 0.05 \times 192563.8 = 9628.19$$

四、 沪深 300 股指期货高频交易机会

1. 市场有效性检验

证券市场的有效性一直是学术界争论不休的话题。我们利用沪深 300 股指期货的高频数据来研究股指期货市场是否存在有效性，即市场是否存在交易性机会可赚取稳定的利润。在此，我们选择一个较为简单直观的测试—游程检验（Runs Test），其测试方法是检测价格变化方向的持续性，即同一方向变化的价格是否连续。

下面我们简单介绍一下此方法的测试步骤

- 1) 首先我们对分笔成交价格的变化进行计数。当发生持续同方向变化时我们记这些持续的变化为一次连续，如下表。
- 2) 统计样本内正向连续和负向连续的累积次数，下表中共 $u = 8+5 = 13$ 次。统计样本内正向变化次数 $n_1 = 10$ 和负向变化的次数 $n_2 = 6$ 。
- 3) 如果价格变化是随机的，则发生连续的次数的期望值为

$$\bar{x} = \frac{2n_1n_2}{n_1 + n_2} + 1, \text{ 其标准差为 } s = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}.$$

- 4) 利用样本的观测值和其期望值及标准差，我们检测样本的随机性。我

$$\text{们的测试统计量为 } z = \frac{|u - \bar{x}| - 0.5}{s}.$$

表 5：股指期货主力合约 2010/4/16 开盘交易价格变化

日期	价格	时间	价格变化	变化方向	Runs
2010/4/16	3450	9:14:00			
2010/4/16	3456	9:15:00	0.075%	+	P1
2010/4/16	3456	9:15:01	0.000%		
2010/4/16	3453.8	9:15:01	-0.028%	-	N1
2010/4/16	3453	9:15:02	-0.010%	-	N1
2010/4/16	3455	9:15:02	0.025%	+	P2
2010/4/16	3455	9:15:03	0.000%		
2010/4/16	3458	9:15:04	0.038%	+	P3
2010/4/16	3458	9:15:04	0.000%		
2010/4/16	3460	9:15:05	0.025%	+	P4
2010/4/16	3460	9:15:05	0.000%		
2010/4/16	3462	9:15:06	0.025%	+	P5
2010/4/16	3465	9:15:07	0.038%	+	P5
2010/4/16	3465	9:15:08	0.000%		
2010/4/16	3463	9:15:08	-0.025%	-	N2
2010/4/16	3466	9:15:09	0.038%	+	P6

2010/4/16	3468	9:15:09	0.025%	+	P6
2010/4/16	3466	9:15:10	-0.025%	-	N3
2010/4/16	3466	9:15:10	0.000%		
2010/4/16	3468	9:15:11	0.025%	+	P7
2010/4/16	3466	9:15:11	-0.025%	-	N4
2010/4/16	3466	9:15:12	0.000%		
2010/4/16	3468	9:15:12	0.025%	+	P8
2010/4/16	3468	9:15:13	0.000%		
2010/4/16	3468	9:15:13	0.000%		
2010/4/16	3468	9:15:14	0.000%		
2010/4/16	3468	9:15:14	0.000%		
2010/4/16	3468	9:15:15	0.000%		
2010/4/16	3467.6	9:15:15	-0.005%	-	N5

资料来源：光大证券研究所、Bloomberg

在全样本内，我们检测的假设为 H_0 ：价格变化方向符合随机分布

H_1 ：价格变化方向有正相关性

经样本检测 $z = -235.6245$, $p\text{-value} < 2.2 \times 10^{-16}$, 由此我们可以拒绝零假设，认为分笔数据存在着一定的正相关性，期指市场在高频纬度的观测下不具备有效性。

2. 高频交易的利润机会

上文中提到高频交易带来的回报有可能达到上千倍的夏普比，在此我们用样本内的数据进行模拟检验。假设在每一个数据观测周期内交易策略所做的方向性判断都是正确的（暂不考虑交易成本），则在此频率下交易模型可能获得的最大夏普比为

$$SR = \frac{E[\text{观测周期内价格平均波动幅度}]}{\sigma[\text{观测周期内价格平均波动幅度}]} \times \sqrt{(\text{日内观测次数}) \times (\text{一年内的交易日})}$$

通过计算我们得到如下的结果：

表 6：不同频率下交易策略的最优表现比较

	平均最大获利	标准差	观测样本数量	最大年化夏普比
分笔数据	0.01%	0.01%	2,269,257	1555.4
1 分钟间隔	0.10%	0.22%	23,364	124.12
5 分钟间隔	0.19%	0.32%	4,958	70.9
10 分钟间隔	0.20%	0.31%	2,522	55.42

资料来源：光大证券研究所、Bloomberg

3. 高频交易机会的探索

基于对前文股指期货高频数据游程检验的测试结果，我们观察发现日内涨跌通常存在着急速连续上涨和下跌的现象，结合市场参与者大多以投机行为为主的特点，我们揣测短时间段的追涨杀跌策略可能提供盈利空间。那么怎样的信号可以作为这种急速价格变化的起始点呢？联想到上文提到的冲击成本模型，我们猜测超越买卖报价量的成交单也许能作为急速价格变化的催化剂。

1) 交易信号的产生

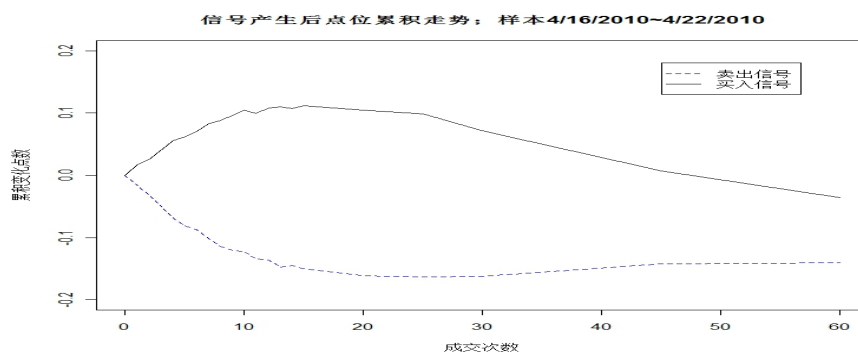
首先，我们根据每笔交易价格与交易前报价的关系确定此笔交易的触发方向，即买方触发还是卖方触发。如价格与买一接近，则此单被定为卖单触发；如价格与卖一接近，则被定为买单触发。如价格与买一和卖一距离相等，则不确认这单的触发方向。

其次，我们比较成交单量与报价单量的关系。如成交单量比买一和卖一的报价单量都大的时候，我们认为初步信号确立，期指市场有可能遵循着此笔成交的触发方向进行变动。

最后，我们比较当前成交价与近 10 笔成交价的算术平均值，如果成交价超平均值 0.2 点，且为买方触发并满足初步信号条件则买入信号确认；如果成交价低与平均值 0.2 点，且为卖方触发并满足初步信号条件则卖出信号确认。

利用上述规则，我们对测试样本内的主力合约进行逐日检验。我们的样本日期仍为 2010/4/16 至 2010/8/20，每日的样本截取区间为上午 9:30 至 11:15 和下午 13:15 至下午 15:00 以保证测试样本区间内市场有足够的流动性。下图给出了自买/卖信号发生后下一笔成交开始，合约成交价变化的累积走势。选取的子样本为 2010/4/16 至 2010/4/22。观察此图可以发现，当买入或卖出信号产生后，成交价格变化发生明显的差别走势。经过大约十几笔交易后，交易价格之间的差异达到最大值。

图 39：信号产生后价格点位的累积走势（子样本 4/16/2010-4/22/2010）



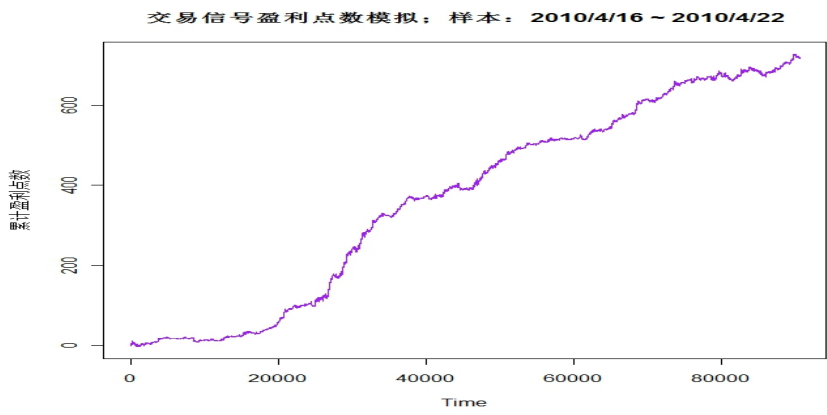
资料来源：光大证券研究所、Bloomberg

在对整体样本进行检验后发现，这种差异性的走势具有稳定性和持续性，因此我们认为这种信号可以作为模拟的交易信号。

2) 交易模型测试

在对信号的产生进行统计后发现，此类信号发生的频率过高，平均每日的换手率高达上千次。过多的信号将会造成频繁的交易，因此必须选用一些手段来降低信号产生的频率。由前述的测试发现，样本内信号发生后大约十笔交易时，买卖信号的差异最为明显。因此我们认为信号合理的持续时间应该为十笔交易，十笔交易之内如无新的信号产生，交易应该结束。十笔内如有新的信号产生，则新的信号将取代旧的信号，信号持续时间也将重新开始计算。根据修改后的信号，我们对样本再次进行模拟测试，查看作为交易信号时，策略的获利情况。

图 40：交易信号盈利点数模拟走势（子样本 4/16/2010-4/22/2010）



资料来源：光大证券研究所、Bloomberg

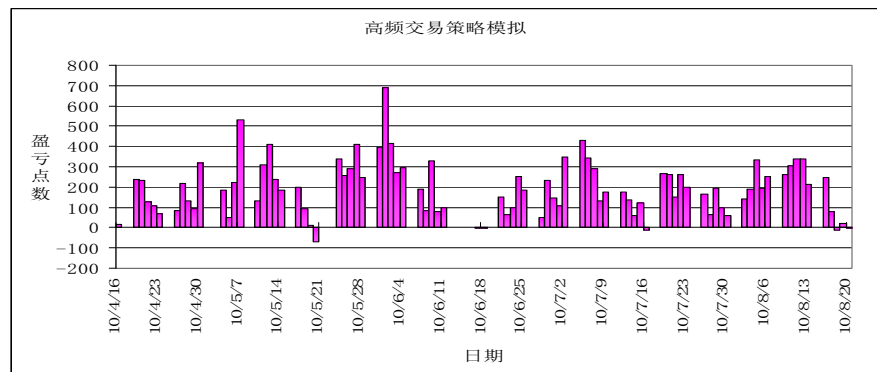
从上图可以发现，作为交易信号时，交易策略的盈利能力惊人！在选取的子样本区间内，信号所产生的盈利点数达到 717 点，回报率达到 20%以上。在对全样本进行信号检验后我们统计得到的结果于下表：

表 7：交易信号全样本盈利统计

全样本模拟（共 87 天）		剔除交割日前 4 天的样本（共 71 天）	
平均	190.73	平均	225.30
盈利天数概率	93%	盈利天数概率	100%
最大盈利	687.60	最大盈利	687.60
最大亏损	-71.80	最小盈利	14.2

资料来源：光大证券研究所、Bloomberg

图 41：交易信号全样本盈利统计



资料来源：光大证券研究所、Bloomberg

样本内的模拟结果极有说服力。全样本内共计交易日 87 天，此策略获得正收益的天数达到 81 天，而且所有的亏损交易日均为交割周所发生。如剔除交割周最后 4 天的交易日，策略的平均赢利点位可达到 225 点/天。

4. 交易策略的实战可行性分析

以上的模拟交易策略在实战应用中忽略了三个主要因素

1) 交易成本

高频交易获利的最大障碍是交易费用，由于交易频率可达到几秒钟一次，累加起来的高额手续费会使原本的利润丧失甚至变为亏损。对于沪深 300 股指期货市场，中金所和期货公司所收取的手续费和佣金累计为合约价值的万分之零点八，一般折合为 0.2-0.3 个点，而我们的信号产生后平均累积收益却不到 0.2 个点。因此我们需要提取更强烈的信号，以保证在扣除手续费和佣金之后仍能稳定地获取正向收益。

2) 交易执行

交易信号产生后，交易的执行过程是高频交易成功与否的另一重要因素。通常在信号产生后，需要在最短的时间内发送交易单，并以设定好的价格成交。在上述的策略中，我们假设在信号产生后，以下一笔成交价作为模拟的成交价格。但在实际情况中，由于外部因素限制我们也许无法以那么快的速度执行交易。此外，模拟的交易具有方向性，在无法判断信号产生后下笔交易是由买方还是卖方触发的前提下，我们假设的交易价格会有偏差。

3) 中金所限制

股指期货自 2010 年 4 月 16 日推出以来仅半年多的时间，作为股票市场最主要的衍生品市场，其发展备受关注。中金所为限制市场的过度投机，限制投资者单边的持仓量和交易次数，因此在交易制度改革前，高频交易的

实现会受到较大阻力。

五、 股指期货高频交易的未来展望

1. 市场容量的增长

股指期货的成交量自其推出以来呈波动的增长态势，单日最大的成交量可达到47万手，而近期低点却有不足15万手。对于流动性要求极高的高频交易策略来说，日内成交量和报价量所提供的交易机会尤为关键，目前尚不十分稳定的市场容量仍无法很好支撑此种交易策略的市场拓展。对于市场容量的增长，我们认为有以下几个关键点：

第一，单边持仓量限制的放松有望加大市场的参与度。无论是对个人投资者或是机构投资者，目前持仓量的限制都约束了他们对市场的参与程度。如果随着市场的逐步成熟，此方面的限制或有望放松，市场容量随之也会得到有效的提升。

第二，自动化高频交易的参与者增多。投资者通常都是利益的追随者，当越来越多的投资者认识到高频交易可带来的潜在的利润时，市场将出现会越来越多的高频交易者，这些追求短期利润的参与者也将是市场上不可忽视的流动性提供者。

第三，公募基金套期保值交易的开始。套期保值的功能是股指期货在公募基金投资中最闪亮的热点，它是机构投资者最有效的抗风险工具之一，同时也可以增加股指期货市场的流动性。

总体而言，我们对股指期货市场未来的容量增长抱着较为乐观的态度，相信随着市场的进一步成熟，市场流动性的提升会为高频交易者带来很好的投资机会。

2. 机构参与度的增加

由于高频交易对于研发人员、设备投入的要求都非常高，相应的投入费用也会较为昂贵，目前国内市场仅有数十家私募机构参与，且这些机构的主要成员均有海外市场高频交易的经验。对于投研能力较强的国内公募基金，这同样是一个较好的投资挖掘领域。值得注意的是，由于公募基金的监管较严，且中金所以对每日的交易次数有一定的限制，这在一定程度上降低了高频交易的获利空间。据我们了解，目前已有数家基金公司希望在其专户理财帐户上运用此种交易技术为客户理财。如果这一尝试能够成功，将会加速高频交易在市场的受欢迎度，监管层也会适度放宽对高频交易的限制措施。

3. 交易机会的挖掘

高频交易的出现为投资者提供了无限的想象空间，基于各种研究方向的投资策略都有机会在此市场中赢得一席之地。对于交易机会的挖掘将主要集中在量化策略的研究以及交易策略执行的优化上。目前国外最优秀的高频交易团队如“GETCO”、“Two Sigma”、“Renaissance Technology”等都有上百人的量化和 IT 研究队伍，且多数研究人员为国际知名学府的教授或博士。作为这一新领域的跟随者，我们将更多借鉴国外先进的量化投资理念，在将来的研究中介绍更多更全面的高频投资策略和技巧。

光大证券股份有限公司

上海市新闻路 1508 号静安国际广场 邮编：200040
总机：021-22169999

光大证券研究所

上海市新闻路 1508 号静安国际广场 3 楼 邮编：200040
总机：021-22169999 传真：021-22169114

销售小组	姓名	办公电话	手机	电子邮件
北京				
	李大志	010-68561122-1717	13810794466	lidz@ebscn.com
	郝辉	010-68561722	13511017986	haohui@ebscn.com
	孙威	010-68561122-1779	13701026120	sunwei@ebscn.com
	吴江	010-68561122-1732	13718402651	wujiang@ebscn.com
上海				
	戴茂戎	021-22169999-9086	13585757755	daimr@ebscn.com
	王莉本	021-22169999-9083	13641659577	wanglb@ebscn.com
	杨日昕	021-22169999-9082	13817003122	yangrx@ebscn.com
深圳				
	王汗青	0755-83024403	13501136670	wanghq@ebscn.com
	黎晓宇	0755-83024434	13823771340	lix1@ebscn.com
	黄鹂华	0755-83024396	13802266623	huanglh@ebscn.com
QFII				
	濮维娜	021-62152373	13301619955	puwn@ebscn.com
	陶奕	021-62152393	13788947019	taoyi@ebscn.com
	满国强	021-62152393	15821755866	mangq@ebscn.com

免责声明

本研究报告仅供光大证券股份有限公司（以下简称“本公司”）的客户使用。

本报告由光大证券股份有限公司研究所编写，以合法地获得尽可能可靠、准确、完整的信息为基础，但不保证所载信息之精确性和完整性。光大证券研究所将随时补充、修订或更新有关信息，但未必发布。本报告根据中华人民共和国法律在中华人民共和国境内分发，供投资者参考。

本报告中的资料、意见、预测均反映报告初次公开发布时的判断，可能会随时调整。报告中的信息或所表达的意见不构成任何投资、法律、会计或税务方面的最终操作建议，本公司不就报告中的内容对最终操作建议作出任何担保。

在任何情况下，本报告中的信息或所表达的建议并不构成对任何投资人的投资建议，光大证券股份有限公司及其附属机构（包括研究所）不对投资者买卖有关公司股份而产生的盈亏承担责任。

本公司的销售人员，交易人员和其他专业人员可能会向客户提供与本报告中观点不同的口头或书面评论或交易策略。本公司的资产管理部和投资业务部可能会作出与本报告的推荐不相一致的投资决策。投资者应明白并理解投资证券及投资产品的目的和当中的风险。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎抉择。

本报告的版权仅归本公司所有，任何机构和个人未经书面许可不得以任何形式翻版、复制，刊登，发表，篡改或者引用。