

订单簿事件研究：不平衡订单流 资金交易策略报告之六

核心观点：

● 基于订单簿事件的价格模型

绝大多数对于国内高频数据订单簿的分析都会遇到两个难题：一是难以观测到真实挂单、撤单与市价单的到达；二是订单簿每一档的挂单量是一个随机波动数字，而非一个封闭的系统。我们在本报告的分析旨在尽力克服这两类困难，并证明在价格预测方面，引入更多订单簿信息（例如挂单与撤单）比单纯使用成交量信息要更有效。

● 模型的变量与估计以及实盘中的约束

尽管我们提出了贴近真实订单簿事件的模型，但也仅仅是对真实报单事件的估计；此外，我们在实盘测试中发现由于观测期和预测期均较短，因此执行交易的持仓时间较短，导致每手预期收益不高，从而对滑点和执行价的要求较高，我们会在正文简单阐述我们对于这一部分的理解。

分析师

夏钦

☎：0755-83478797

✉：xiaqin @chinastock.com.cn

执业证书编号：S0130514050006

相关研究

- | | |
|------------------------------------|---------|
| 《初探市场微观结构：指令单簿与指令单流—
—资金交易策略之四》 | 2013.9 |
| 《订单簿驱动策略与交易细节》 | 2013.12 |

目 录

一、基于订单簿事件的价格模型	2
二、模型的变量估计	4
（一）变量的计算.....	4
（二）模型拟合结论.....	6
（三）一个简单的策略举例.....	6
三、风险提示	7

过去一年我们对程序化交易的研究总体来说是一个从低频数据向高频数据、从模拟到实战、从单纯重视策略到重视交易细节与交易系统的过程。显然，低频度的日内交易和高频交易之间并不存在孰优孰劣的区别，只要投资人能找到适用的、稳健的盈利模式，两种交易模式都是值得挖掘并实战的，只是高频交易在执行层面往往更加重视细节，看重交易成本和完成率。这一点即便是在低频度的日内交易上，也可以替投资人节省交易成本或者更好的理解市场。

在纷繁的高频数据里，我们对订单簿尤为关注，但是我们也曾困惑，花费大量精力去研究挂单、撤单和市价成交这类事件，是否真的会比单纯研究成交量来得更有效更有价值？这篇报告概括了我们在过去做的一些工作，以解答上面的疑问。

一、基于订单簿事件的价格模型

全文我们分析的对象都是股指期货的 level1 行情数据（尽管 level2 数据能提供更多深度，但是一样无法判定除了成交以外的订单簿事件），我们不妨先把订单簿事件分为下面六类：

L_k^b ：买方挂单；	L_k^s ：卖方挂单；
C_k^b ：买方撤单；	C_k^s ：卖方撤单；
M_k^b ：买方市价；	M_k^s ：卖方市价。

则相对应的我们把买一和卖一价格的变动定义为：

$$\Delta P_k^{b,s} = (P_k^{b,s} - P_{k-1}^{b,s})$$

我们尝试说明订单簿本身是如何动态变化的，但是很快我们就发现第一个难题出现了：订单簿的深度不是一个封闭的系统，而是一个均值在 20 附近的随机波动的系统。

图 1：2013 年 IFHot 买一量概率分布图，均值=22

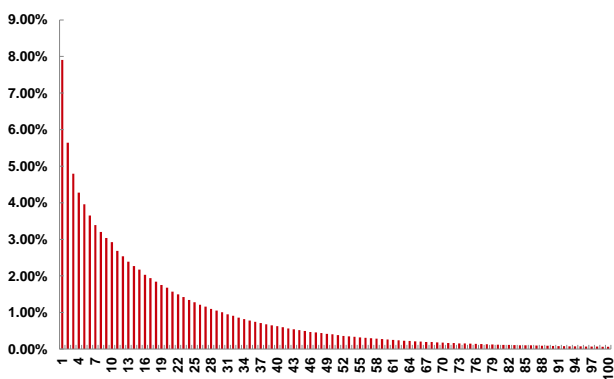
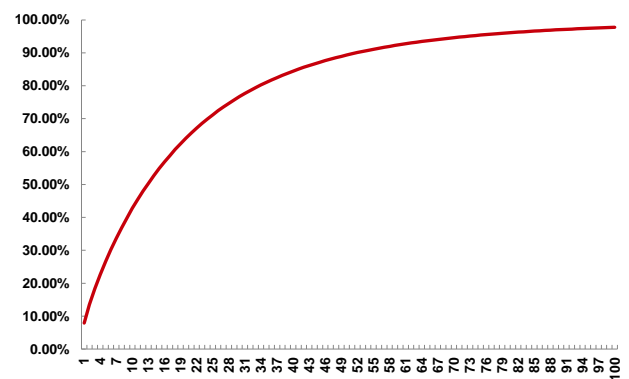


图 2：2013 年 IFHot 买一量帕累托分析



资料来源：中国银河证券研究部

为了更好地阐述订单簿的动态变化过程，我们参考 Rama Cont (2012) 提出的理想假设，假设订单簿每一个档位上的挂单上限为 D ，当某一价位上的挂单数量超过 D 的时候，就只能向下一个次优价位挂单。我们通过三种图来说明市价成交、挂单、撤单一些列事件对订单簿的影响，为了便于说明和观察，假设三种事件依次发生， $D=5$ 且最小价格变动单位为 δ 。

图 2 描述一个 $M=15$ 的市价卖单到达买盘，消除 $15/5=3$ 个队列，因此价格下跌 3δ ；图 3 则表示紧接着图 2 之后发生挂单 $L=7$ ，价格回升了两个最小单位；图 4 紧接着图 3 发生了 $C=4$ 的撤单事件。

图 3：市价单对订单簿影响

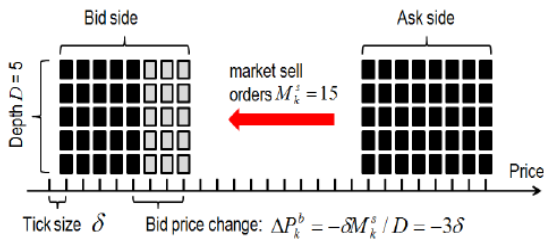
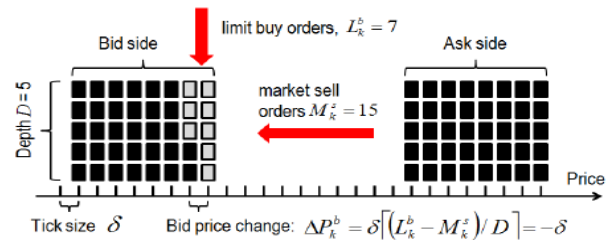
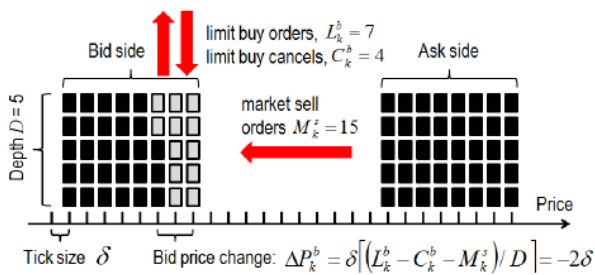


图 4：挂单对订单簿影响



资料来源：中国银河证券研究部, Rama Cont (2012)

图 5：撤单对订单簿影响



资料来源：中国银河证券研究部, Rama Cont (2012)

这一过程非常容易理解，然而现实不会有这么理想，一方面各种订单簿事件会同步到达，另一方面每一个价位的深度也不会是固定的 D ，但是这并不影响我们给出一个准确的并且是具有普适性的表达式来说明买一与卖一价格的变动：

$$\Delta P_k^b = \delta \left[\frac{L_k^b - C_k^b - M_k^s}{D} \right]$$

$$\Delta P_k^s = \delta \left[\frac{L_k^s - C_k^s - M_k^b}{D} \right]$$

这两个表达式即为上面三种图所示的过程。我们在现实交易中可能更关注中间价的变动趋势，我们可以把中间价和中间价的变动表达为：

$$P_k = \frac{P_k^b + P_k^s}{2\delta}$$

$$\Delta P_k = \frac{\Delta P_k^b + \Delta P_k^s}{2\delta}$$

我们可以把 ΔP_k^b 与 ΔP_k^s 的定义代入上面中间价变动表达式：

$$\Delta P_k = \frac{L_k^b - C_k^b - M_k^s - L_k^s + C_k^s + M_k^b}{2D} + \varepsilon_k$$

我们定义分子即为 OFI (Order Flow Imbalance) :

$$OFI = L_k^b - C_k^b - M_k^s - L_k^s + C_k^s + M_k^b$$

则中间价变动表达式重写为:

$$\Delta P_k = \frac{OFI}{2D} + \varepsilon_k$$

可以看到, OFI 其实是揉杂了市价成交、挂单与撤单事件综合效果的因子。我们也可以单独研究市价成交事件的影响, 以此来对比说明我们付出这么多努力研究多事件是否有意义。我们把市价成交事件定义为:

$$TI_k = M_k^b - M_k^s$$

即买方成交减去卖方成交, 以说明买卖双方的能量差。我们只要对上文 ΔP_k 的定义稍作改动就能得到针对 TI 的表达式, 将市价单事件留在分子中, 而将其他部分统统放入噪音项:

$$\Delta P_k = \frac{TI_k}{2D} + \eta_k \quad \eta_k = \frac{L_k^b - C_k^b - L_k^s + C_k^s}{2D} + \varepsilon_k$$

这样一来, 如果 OFI 是显著的, 那么 TI 的拟合效果就会比较糟糕, 或者说噪音会很大, 因为大量信息被掺杂在扰动项里。

二、模型的变量估计

(一) 变量的计算

现在我们的工作集中在两个简单拟合的检验上, 具体如下图所示:

图 6: 模型的估计

$$\Delta P_k = \frac{OFI}{2D} + \varepsilon_k \quad \longrightarrow \quad \left\{ \begin{array}{l} \Delta P_{k,i} = \beta_i OFI_{k,i} + \varepsilon_{k,i} \\ \beta_i = \frac{C}{D_i^\lambda} + v_i \quad C=1/2, \lambda=1 \\ \Delta P_{k,i} = \beta_i^T TI_{k,i} + \eta_{k,i} \end{array} \right. \quad \text{侧重不同}$$

资料来源: 中国银河证券研究部

针对中间价变动的两个拟合式其实是一样的, 只是侧重不同; 易知 β_i 其实是一个关于盘

口深度倒数的表达式，我们可以对其进行具体建模，但是过程较为复杂在此报告中不做讨论，我们会在今后的报告中再行介绍。

尽管现在问题已经被大大简化，但是回想我们在文首提到的困难，OFI 依旧是无法根据最精确的直接定义来直接计算观测的，因为挂单、撤单和市价单成交事件并不会区分表现在 Level1 行情上（当然如果可以得到这样详细的订单簿数据，对提升模型效果应该大有帮助）。为此，我们需要一个代理算法来估算 OFI，再次我们介绍一种算法以做参考：

$$e_n = q_n^b \mathbb{I}_{p_n^b \geq p_{n-1}^b} - q_{n-1}^b \mathbb{I}_{p_n^b \leq p_{n-1}^b} - q_n^s \mathbb{I}_{p_n^s \leq p_{n-1}^s} + q_{n-1}^s \mathbb{I}_{p_n^s \geq p_{n-1}^s}$$

表示 if 条件，当括号内的条件成立时，括号外的表达式保留，当括号内的条件不成立时，括号外的表达式为 0。e 表示从 n-1 到 n 时刻的 OFI 值， q^b 分别为 n-1 及 n 时刻买一挂单量， q^s 分别为 n-1 及 n 时刻卖一挂单量。

由此我们可以针对各种不同的买一卖一变化情况得到不同的 e 的表达式，例如，当买一卖一价格都不变时，得到 e 如下：

$$e_n = q_n^b - q_{n-1}^b - q_n^s + q_{n-1}^s$$

它表示的是买一价位的净变化值减去卖一价位的净变化值，所谓净变化值即为一个时间段内挂单、撤单及市价单综合形成的效果。例如，e 表达式前两项表示买一挂单在价格没有发生变化下的净变化值，我们由上文已经可以知道，当价格不变时：

$$q_n^b = q_{n-1}^b + L_n^b - C_n^b - M_n^s \quad q_n^b - q_{n-1}^b = L_n^b - C_n^b - M_n^s$$

卖方挂单的净变化值同理。又例如当买一价格上涨，卖一价格也上涨时，e 的表达式为：

$$e_n = q_n^b + q_{n-1}^s$$

即为 n 时刻的买一挂单量与 n-1 时刻卖一挂单量之和来估计 OFI，其背后含义是：买方吃掉了所有上一刻挂在卖一的单 q_{n-1}^s 并且“溢出” q_n^b 成为新的买一量。其他多种情况都可以得到相应的表达式。通过在每一个 tick 滚动计算 e，可以最后得到某一时间段内的 OFI 值：

$$OFI_{k,i} = \sum_{n=N(t_{k-1,i})+1}^{N(t_{k,i})} e_n$$

拟合式中的其他两项较容易计算，中间价的位移可以表述为：

$$\Delta P_{k,i} = \frac{P_{N(t_{k,i})}^b + P_{N(t_{k,i})}^s}{2\delta} - \frac{P_{N(t_{k-1,i})}^b + P_{N(t_{k-1,i})}^s}{2\delta}$$

TI 的计算过程表述为:

$$TI_{k,i} = \sum_{n=N(t_{k-1,i})+1}^{N(t_{k,i})} (b_n - s_n)$$

其中, b_n 为买方发起成交量, s_n 为卖方发起成交量。具体甄别买卖方向的方法可以参考我们前期的报告《金融工程-交易技术: 订单簿驱动策略与交易细节》, 在此我们列出关键伪代码示意:

图 7: 成交量分解示意图

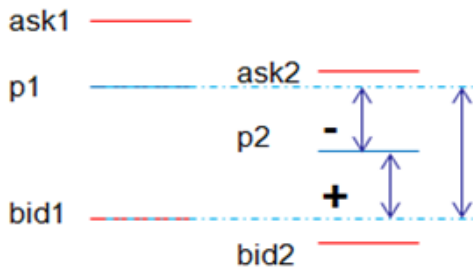


图 8: 成交量分解伪代码

```
if close >= qofask[1] then
begin
    //成交量计为正向主动攻击
end
else if close <= qofbid[1] then
begin
    //成交量计为负向主动攻击
end
else if close < qofask[1] and close > qofbid[1] then
begin
    if close > close[1] then
    begin
        //成交量按比例分解为主动正向攻击与被动正向攻击部分
    end
    else if close < close[1] then
    begin
        //成交量按比例分解为主动负向攻击与被动负向攻击部分
    end
    else if close = close[1] then
    begin
        //成交量按照 qofask, qofbid, qofask[1], qofbid[1] 关系进行分解
    end;
end;
end;
```

资料来源: 中国银河证券研究部, 《金融工程-交易技术: 订单簿驱动策略与交易细节》

(二) 模型拟合结论

针对下面三个简单拟合我们得到一些有趣的结论:

$$\Delta P_{k,i} = \beta_i OFI_{k,i} + \varepsilon_{k,i}, \quad \Delta P_{k,i} = \beta_i^T TI_{k,i} + \eta_{k,i}, \quad \Delta P_{k,i} = \alpha_i + \theta_i^o OFI_{k,i} + \theta_i^T TI_{k,i} + \varepsilon_{k,i}$$

(1) OFI 的参数 β_i 显著性明显优于 β_i^T , 事实上 β_i^T 勉强满足显著性。

(2) 将 OFI 与 TI 放到一个拟合式中不会带来任何增益。

我们初步认为考虑多种订单簿事件的因子 OFI 要比单纯考虑市价成交事件的 TI 因子更能解释中间价的移动, 换言之我们付出更多时间去研究挂单、撤单以及市单是有一定的意义的。

对于扰动项的进一步分析较为复杂, 我们不打算在本报告中展开而会在今后的报告中逐步分析。

(三) 一个简单的策略举例

策略展示并不是这篇报告的重点, 和以往的报告不同, 我们更多想说明一个思路而非一个现成的策略。事实上高频交易策略的实际部署和运行牵涉到诸多细节, 也无法通过一个简单的

报告来完整说明。

在我们开发实盘落地策略过程中，我们最关注的是期望收益与成本损耗之间的关系。过于频繁的交易往往伴随着较低的每手期望收益，这就导致策略净值对滑点和成本极为敏感，而回测的可信度就大大降低，而这很大程度上是因为回测过程很难准确模拟实际交易环境。举个最简单的例子，回测无法模拟抢单的过程，而单纯按照对价和对手盘的量来撮合，这种一个没有任何竞争者的市场无疑是理想化的，因此对于交易频率越高的策略，越是需要依靠真实的市场环境来测试回溯业绩的可靠性。

下图是我们举例的简单策略：

图 9：简单策略举例



资料来源：中国银河证券研究部

策略并没有涉及很复杂的信号计算，利用上文提到的 OFI 信号向前回溯 m 区间，每个区间 n 秒（即 $2n$ 个 tick），在每个区间核算区间 OFI，并最终预测下一个时间区间的价格走势概率。由于我们将观察区间的宽度 n 设得很长，因而信号较为稀疏，交易频率较低，日均不到 2 次，20 万一手以对价撮合扣手续费后收益率为 49%，累计最大回撤 6.9%，收益回撤比 6.8，本金回撤比 6.2，盈利天数 112，亏损天数 88。尽管效果一般，但是 OFI 概念的简单利用已经表现出一定的研究价值，我们会在之后的研究中继续挖掘订单簿信息并开发更贴近实际操作模型。

三、风险提示

本报告中的所有模型和结论均按历史数据测算，只供投资者参考，不必然保证未来有同样好的收益，亦不能完全排除未来的风险，特别提醒投资者注意。

插图目录

图 1: 2013 年 IFHot 买一量概率分布图, 均值=22	2
图 2: 2013 年 IFHot 买一量帕累托分析.....	2
图 3: 市价单对订单簿影响	3
图 4: 挂单对订单簿影响	3
图 5: 撤单对订单簿影响	3
图 6: 模型的估计.....	4
图 7: 成交量分解示意图	6
图 8: 成交量分解伪代码	6
图 9: 简单策略举例	7

评级标准

银河证券行业评级体系：推荐、谨慎推荐、中性、回避

推荐：是指未来 6-12 个月，行业指数（或分析师团队所覆盖公司组成的行业指数）超越交易所指数（或市场中主要的指数）平均回报 20% 及以上。该评级由分析师给出。

谨慎推荐：行业指数（或分析师团队所覆盖公司组成的行业指数）超越交易所指数（或市场中主要的指数）平均回报。该评级由分析师给出。

中性：行业指数（或分析师团队所覆盖公司组成的行业指数）与交易所指数（或市场中主要的指数）平均回报相当。该评级由分析师给出。

回避：行业指数（或分析师团队所覆盖公司组成的行业指数）低于交易所指数（或市场中主要的指数）平均回报 10% 及以上。该评级由分析师给出。

银河证券公司评级体系：推荐、谨慎推荐、中性、回避

推荐：是指未来 6-12 个月，公司股价超越分析师（或分析师团队）所覆盖股票平均回报 20% 及以上。该评级由分析师给出。

谨慎推荐：是指未来 6-12 个月，公司股价超越分析师（或分析师团队）所覆盖股票平均回报 10% - 20%。该评级由分析师给出。

中性：是指未来 6-12 个月，公司股价与分析师（或分析师团队）所覆盖股票平均回报相当。该评级由分析师给出。

回避：是指未来 6-12 个月，公司股价低于分析师（或分析师团队）所覆盖股票平均回报 10% 及以上。该评级由分析师给出。

夏钦，证券分析师。本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，本人承诺，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接接受到任何形式的补偿。（本人承诺不利用自己的身份、地位和执业过程中所掌握的信息为自己或他人谋取私利）。

免责声明

本报告由中国银河证券股份有限公司（以下简称银河证券，银河证券已具备中国证监会批复的证券投资咨询业务资格）向其机构或个人客户（以下简称客户）提供，无意针对或打算违反任何地区、国家、城市或其它法律管辖区域内的法律法规。除非另有说明，所有本报告的版权属于银河证券。未经银河证券事先书面授权许可，任何机构或个人不得更改或以任何方式发送、传播或复印本报告。

本报告所载的全部内容只提供给客户做参考之用，并不构成对客户的投资建议，并非作为买卖、认购证券或其它金融工具的邀请或保证。银河证券认为本报告所载内容及观点客观公正，但不担保其内容的准确性或完整性。客户不应单纯依靠本报告而取代个人的独立判断。本报告所载内容反映的是银河证券在最初发表本报告日期当日的判断，银河证券可发出其它与本报告所载内容不一致或有不同结论的报告，但银河证券没有义务和责任去及时更新本报告涉及的内容并通知客户。银河证券不对因客户使用本报告而导致的损失负任何责任。

银河证券不需要采取任何行动以确保本报告涉及的内容适合于客户。银河证券建议客户如有任何疑问应当咨询证券投资顾问并独自进行投资判断。本报告并不构成投资、法律、会计或税务建议或担保任何内容适合客户，本报告不构成给予客户个人咨询建议。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的银河证券网站以外的地址或超级链接，银河证券不对其内容负责。本报告提供这些地址或超级链接的目的纯粹是为了客户使用方便，链接网站的内容不构成本报告的任何部份，客户需自行承担浏览这些网站的费用或风险。

银河证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。银河证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

银河证券无需因接收人收到本报告而视其为客户。本报告是发送给银河证券客户的，属于机密材料，只有银河证券客户才能参考或使用，如接收人并非银河证券客户，请及时退回并删除。

所有在本报告中使用的商标、服务标识及标记，除非另有说明，均为银河证券的商标、服务标识及标记。

银河证券版权所有并保留一切权利。

联系

中国银河证券股份有限公司 研究部

北京市西城区金融街 35 号国际企业大厦 C 座
上海浦东新区富城路 99 号震旦大厦 26 楼
深圳市福田区福华一路中心商务大厦 26 层
北京市西城区金融街 35 号国际企业大厦 C 座
北京市西城区金融街 35 号国际企业大厦 C 座
公司网址：www.chinastock.com.cn

机构请致电：

北京地区：傅楚雄 010-83574171 fuchuxiong@chinastock.com.cn
上海地区：何婷婷 021-20252612 hetingting@chinastock.com.cn
深广地区：詹璐 0755-83453719 zhanlu@chinastock.com.cn
海外机构：高兴 83574017 gaoxing_jg@chinastock.com.cn
海外机构：李笑裕 83571359 lixiaoyu@chinastock.com.cn