

订单簿驱动策略与交易细节

资金交易策略报告之五

2013年12月11日

报告摘要：

● 一号程序化策略的跟踪情况与改进思想简述

系列报告的前几篇均立足沪深300股指期货的基本面——300只个股的成交明细数据进行研究，并形成较为成熟的股指期货程序化策略一号进行跟踪（6月26日开始），这一节将介绍一号策略近期的运行状况，以及策略在运行中逐渐改进的细节。

● 程序化策略交易细节设计：交易细节重于策略本身

我们在实践过程中逐步感受到高频交易的部署细节和交易环节上的细节处理甚至很大程度上比策略本身更为重要。我们将在这一节介绍我们在交易系统设计方面的一些核心理念和细节处理思想，这些细节也与我们将来开发的策略紧密相关。

● 关于时间久期、价格久期与成交量久期的区别与运用

这一节介绍策略运用到的久期的基本概念，包括时间久期、成交量久期和价格久期。并且综述我们一号策略在久期利用上的特点，介绍VNET和VPIN利用久期在风险管理方面的运用。

民生证券金融工程团队

分析师：王红兵

执业证号：S0100512090002

电话：(86755)22662129

邮箱：wanghongbing@mszq.com

分析师：夏钦

执业证号：S0100513100001

电话：(86755)22662124

邮箱：xiaojin@mszq.com

地址：深圳市福田区深南大道7888号
东海国际中心A座28层 518040

相关研究

《加权净成交量看大盘——资金交易策略低频版》 2012.11.19

《价格久期在大盘择时中的简单运用——资金交易策略高频版》 2012.11.21

《初探市场微观结构：指令单薄与指令单流》 2013.9.4

《交易执行细节，从模拟走向实战——CTA程序化交易实务研究之五》 2013.9

目录

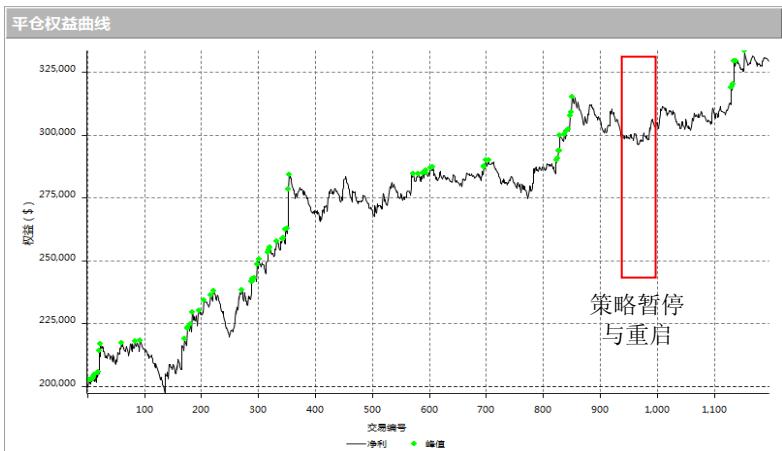
一、一号程序化策略的跟踪情况与改进思想简述.....	3
(一) 一号策略的属性逐渐发生变化.....	3
(二) 策略的核心逻辑及算法进行改进.....	4
二、新策略思想交流：久期的不同运用.....	7
(一) 久期的不同定义.....	7
(二) 不同久期对比.....	8
(三) 一些对久期的利用.....	8
三、下一阶段工作	10
四、风险提示	10
插图目录	11

一、一号程序化策略的跟踪情况与改进思想简述

(一) 一号策略的属性逐渐发生变化

下图为调整后策略一号 2013 年 6 月 25 日至 2013 年 12 月 9 日的净值走势。

图 1：一号策略净值（6月 25 日——12月 9 日）

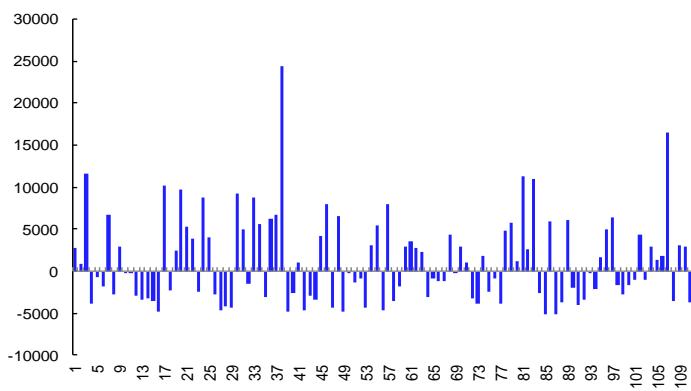


资料来源：民生证券研究院、Multicharts

不得不说 9 月与 10 月对策略的压力较大，在无法运用资金仓位管理和多策略平滑净值的情况下，策略净值一度出现了接近 10%（单手 70 点）的累计回撤。我们在 10 月一度停止了策略的实践交易，并总结之前的问题进行了修改，在之后的一周重新开始交易。

下图为跟踪期间（6月 26 日至 12 月 9 日）每日收益情况，策略平均每日交易 10.67 次。

图 2：一号策略日收益情况



资料来源：民生证券研究院

更重要的是，一号策略的策略属性正在发生变化，我们先来了解下高频交易大致的分类：

图 3: 高频策略大致分类

	高频策略	日内策略
底层数据采集频率	高频	高频
策略容量	较小	较大
持仓时间	短(分钟, tick)	长(分钟, 小时)
交易频率	高, 几十至上百次/日	低, 几次/日
隔夜	否	否
回撤	很小	较大
策略寿命	较短	较长
流动性要求	高	较低
速度要求	高	较低

资料来源：民生证券研究院

我们原先的一号策略属于日内策略范畴，交易频率较低，持仓期长，可承载资金量较大，但是相对的在震荡市里所需要的风险承载能力也较大。随着我们的交易系统不断改良，以及对订单簿研究的不断深入，我们逐渐将一号策略的交易频率提高，接近上表中左侧的高频交易策略。

目前一号策略每日发生交易次数在 15 笔左右（开平仓算一次），改进主要在两方面：

- 1、策略的核心逻辑及算法进行改进，原先对于成交明细表的大单算法及归类方法都被修改；
- 2、引入订单簿数据辅助成交明细表进行判断；
- 3、单位时间内成交量拆解赋予不同符号；
- 2、交易细节方面，进行多事件并发处理，改变原先粗糙的挂单下单方式，精细了追单过程，改变了出场逻辑，新系统基本不使用绝对点数止损。

我们在下文详细介绍。

(二) 策略的核心逻辑及算法进行改进

1. 大单算法进行改进

在首页的相关报告中已经较为详细地介绍了一号策略的核心思想，即通过沪深 300 指数 300 个成分股的成交明细数据实时计算指数大单净流向并藉此对沪深 300 指数或指数期货的短期走势进行判断，因此信号的核心在于筛选大单。我们在之前的报告中提过我们采取合并短时间内同价位成交单的方法，之所以这样做，是因为：一方面，短时间内同价位之间的订单簿可比性较强，一旦成交价格改变，整个订单簿排序变化，再进行时间序列上的比较则意义不大（或者需要复杂的数据处理）；另外，考虑到可能存在的拆单行为，以及 tick by tick 数据的噪音问题，我们合并一些性质相近的交易来过滤噪音，这种方法可以参考《Flow toxicity and liquidity in a high-frequency world》(O'Hara, David Easley, Marcos

M.Lopez de Prado, 2011)中的做法。

2、引入订单簿信息辅助判断

我们把订单簿看作还未发生的交易信息，并尝试挖掘其蕴含的信息，尤其是靠近成交价的前二、三档的信息尤其重要。订单簿与成交明细的许多规律都能给我们启示，例如我们可以统计观察小单、大单、市价单、限价单、撤单和新到挂单在时间轴上发出频率及频数，如下图所示：

图 4：成交小单、成交大单、市价单的到达率

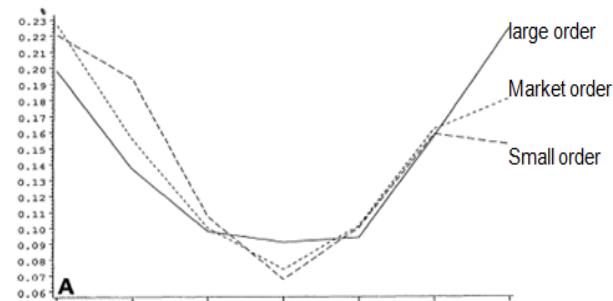
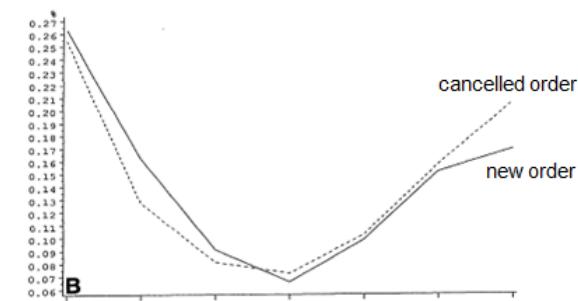


图 5：撤单与新挂单到达率



资料来源：民生证券研究院

大单、小单、市价单与撤单的到达率都呈现出 U 型特征，尽管这种特征在不同品种上的具体表现情况有区别，但大致形态类似，这让我们更倾向于在挂撤单最活跃的时候注重筛选真实突破和假突破（通常这些时段的波动较大），而在相对清淡的时点上选择不频繁交易；另一方面，撤单与新挂单的激进度给我们启发，我们可以通过衡量挂撤单各自的激进度来识别某一方向的趋势是否能够持续，并且是由撤单还是激进的挂单引起的。

我们再举一个例子，对于上面提到的不同类型的成交，我们可以尝试构建一个条件概率表，观察是否存在成交单类型方面的动量。这一研究在国外文献中多次被提及，例如在 NYSE 权重股方面就存在较为显著的“对角线”效应：

图 6：NYSE 权重股的成交单对角线效应

t-1	Appli- cation	Large Buy	Market Buy	Small Buy	New Bid Within	New Bid At	New Bid Below	Cancel Bid	Large Sell	Market Sell	Small Sell	New Ask Within	New Ask At	New Ask Above	Cancel Ask
Application	0.06	3.56	1.98	13.25	9.45	5.42	6.44	5.93	4.15	2.49	21.53	9.10	4.87	6.19	5.64
Large buy	2.94	7.29	3.41	15.34	13.05	5.01	5.90	5.21	1.80	1.11	14.39	7.26	4.65	6.83	5.81
Market buy	1.94	3.20	2.47	17.11	2.14	6.88	11.33	7.92	3.35	6.16	23.49	2.78	2.45	4.84	3.94
Small buy	2.43	3.71	2.89	20.49	9.67	5.27	6.33	4.02	2.35	1.75	19.77	5.88	4.54	6.20	4.72
New bid within	2.33	3.19	2.14	13.12	13.62	6.87	9.21	5.22	3.01	1.98	18.58	8.28	2.92	5.42	4.11
New bid at	1.92	2.23	2.40	14.01	16.52	6.88	7.36	4.93	1.61	0.90	22.64	6.05	3.84	5.30	3.40
New bid below	1.90	1.90	1.60	11.50	7.50	7.99	18.24	4.45	3.23	2.41	19.99	6.92	3.37	5.58	3.42
Cancel bid	2.67	2.27	1.71	10.56	12.45	5.21	8.05	9.70	3.80	2.26	19.54	8.70	3.57	5.79	3.72
Large sell	2.49	0.92	0.93	6.46	8.82	4.54	6.53	5.97	9.84	3.38	23.91	12.57	3.33	4.99	5.37
Market sell	1.27	2.89	6.33	16.16	2.28	2.92	4.28	3.60	4.82	3.04	29.35	2.21	4.98	9.50	6.37
Small sell	2.16	1.59	1.60	10.27	6.92	4.78	5.80	4.37	4.82	2.96	33.87	7.88	4.00	5.29	3.71
New ask within	2.15	2.01	1.87	9.86	9.43	4.03	6.04	4.06	4.44	2.60	22.67	12.61	5.53	7.76	4.94
New ask at	1.94	1.41	0.97	12.79	6.41	4.81	5.65	3.83	3.00	2.48	23.10	14.74	6.65	7.30	4.92
New ask above	2.08	2.24	2.15	10.81	7.24	4.12	5.86	3.36	2.81	2.10	21.35	5.85	7.07	19.36	3.60
Cancel ask	2.48	2.03	1.85	9.95	8.73	4.39	6.10	4.51	4.06	2.13	19.19	13.84	3.95	7.57	9.24
Unconditional	2.2	2.4	2.1	12.5	9.0	5.3	7.4	4.7	3.8	2.4	24.0	8.3	4.3	7.0	4.8

资料来源：民生证券研究院

可以看到，上表对角线所示即为上一刻发生的成交类型，在下一刻发生同样成交类型

的概率。我们在改进一号策略时候充分考虑 IF 合约上的成交单类型的动量效应，事实上也只有这种动量特性才能帮助我们跟随趋势。

3、成交量拆解

给成交量赋符号是一项细致又艰苦的工作，在我们的早期报告里就提及过这一点。在成交量符号方面，目前并没有一个绝对权威的共识，大多使用的是 tick test 和 Lee-Ready，但是我们认为没有必要给一个时间周期的成交量赋统一的符号（事实上不管怎么分配符号都是主观的），我们采用的方法是讲成交量拆解到买卖双边上，我们更关注的是买卖净量。

图 7：成交量分解示意图

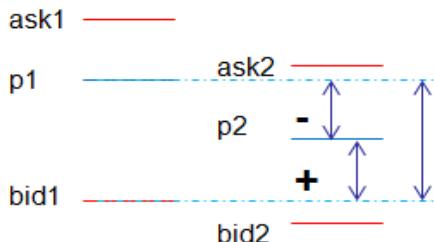


图 8：成交量分解伪代码

```

if close>=gofask[1] then
begin
    //成交量计为正向主动攻击
end
else if close<=gofbid[1] then
begin
    //成交量计为负向主动攻击
end
else if close<gofask[1] and close>gofbid[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gofask, gofbid, gofask[1], gofbid[1]关系进行分解
end;
end;

```

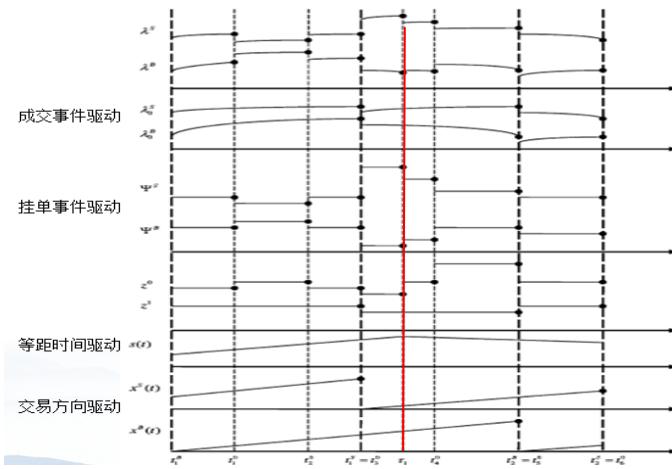
资料来源：民生证券研究院

如上图我们将成交量根据相邻成交价以及买一卖一价格分解赋予符号，求得的净量才是这个时间周期内我们关心的交易情况。

4、交易细节处理：事件并发——完整的逻辑闭环

对于事件并发我们可以用下图简单表述其思想

图 9：多事件并发控制



资料来源：民生证券研究院

上图想试图说明的是，在同一个时间坐标轴上，根据事件驱动因素的不同，可以把时间切割成不同的周期，不同周期之间又会存在并发。图中只有两个时间周期，其中红色线就是平分这两个时间周期的刻度，可以清晰看到等距时间驱动在红色线上有一个节点。又例如，交易方向驱动以买（卖）方向发生成交为触发节点，可以看到两个方向一共有 3 个事件节点，顺着节点向上可以看到成交事件驱动，同样是 3 个节点，但是不区分买卖方向。而最密集的是挂单事件驱动，盘口每次发生变动都会触发。我们大致可以把并发事件分为：

成交事件、挂单事件、挂单拒绝事件、撤单事件、撤单失败事件、系统异常警告事件。

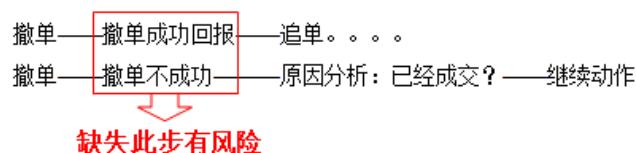
一号策略在最初并没有针对事件之间的并发做过多处理，但是随着我们实践交易推进，我们注意到并发事件控制是日内交易和高频交易都无法回避的问题，因此我们将很大一部分精力付诸于处理好事件并发。

处理并发事件的核心思想是：每一个事件都是完整的逻辑闭环。

（1）每个事件的发出和回报都要完整

我们可以通过下图形象的理解：

图 10：撤单与撤单成功回报



资料来源：民生证券研究院

撤单指令发出后，如果不等待撤单成功回报而盲目追单，可能造成实际仓位与意图不符。

（2）不能出现死循环

这种情况比较少见，例如如下情况：平仓委托发出——柜台发现没有持仓——平仓拒绝——没有分析平仓拒绝原因模块而继续报入平仓委托——柜台发现没有持仓——平仓拒绝——.....。

另外，在并发事件处理中，对于流程风险控制也很重要，例如不能在 1 秒中之内操作过多，反手开仓要做好安全保护机制避免上一持仓还没平掉导致开仓保证金不足；尽量取用交易所 tick 数据的时间戳，本地时间与交易所时间之间通常来说都有误差。

二、新策略思想交流：久期的不同运用

（一）久期的不同定义

久期最常见的定义是时间久期，我们平时所接触的切片行情数据都是按照时间久期进行划分的，比如 1 分钟 K 线，1 秒 K 线或 500 毫秒的股指期货 tick 数据，这种等时间间

距的划分方法被最广泛地接受，也成为无数策略指定信号周期的标准。

另一种常用久期是成交量久期，即当成交量的绝对值累计到一定阈值时定义为一个周期结束，例如某股票成交满 1000 手即为一个成交量久期，之后重新开始累计成交量进入下一个周期。这种久期定义方法最大的特点是刻画标的的成交活跃度，成交越活跃当然久期的时间间隔越短，反之亦然。

还有一种定义是价格久期，即当价格变动的绝对值达到一定阈值即定义为一个周期，例如某股票上涨或下跌 0.5 元即定为一个周期。这种定义方法最大的有点是捕捉价格的快速趋势性变化，这种情况下久期快速缩短；也能过滤价格长期窄幅震荡，这种情况下久期较长。

(二) 不同久期对比

事实上几种久期定义之间并没有好坏之分，不同属性的策略适用不同的久期概念。我们很容易想到：成交量久期和价格久期最明显的缺陷是他们需要事先定义一个阈值来定义触发久期周期，而这个阈值显然会比较主观，这使得样本外的使用或许较有风险。但是回头看时间久期其实也是一样有阈值的问题，我们已经习惯了秒、分这样的定义，但这些定义也有严格的始末点，假设现在定义 1 分 04 秒到 2 分 04 秒为一个分钟 K 线，很多信号效果一样会扭曲，又比如定义一分钟为 50 秒而非 60 秒，也会对策略产生较大影响。

所以三种久期定义都有阈值敏感方面的问题，我们认为只要能够找到适应某一个(量、价、时间)周期的久期参数，确保其稳健性，就有使用价值。

(三) 一些对久期的利用

1. 一号策略对价格久期与时间久期的综合利用

一号策略主要利用的是 300 个股的成交明细表如下图：

图 11：成交明细表举例

代码	名称	时间	价格	成交量	时差
SZ000001	平安银行	2012-11-01 09:30:00	A 13.18	900	
SZ000001	平安银行	2012-11-01 09:30:09	13.18	10	9
SZ000001	平安银行	2012-11-01 09:30:12	13.19	500	3
SZ000001	平安银行	2012-11-01 09:30:30	13.18	400	18
SZ000001	平安银行	2012-11-01 09:30:39	B 13.18	500	9
SZ000001	平安银行	2012-11-01 09:30:45	13.19	1000	6
SZ000001	平安银行	2012-11-01 09:30:51	13.19	100	6
SZ000001	平安银行	2012-11-01 09:31:00	13.19	1200	9
SZ000001	平安银行	2012-11-01 09:31:09	13.19	4500	9
SZ000001	平安银行	2012-11-01 09:31:12	C 13.18	1300	3
SZ000001	平安银行	2012-11-01 09:31:18	13.18	3700	6
SZ000001	平安银行	2012-11-01 09:31:24	13.19	200	6
SZ000001	平安银行	2012-11-01 09:31:30	13.19	4790	6

资料来源：民生证券研究院

我们合并同价位且相邻时间较近的成交单，一方面是考虑价格久期的因素，即价格发生单位变动（例如 0.01 元）视为一个重要的事件周期；另一方面是考虑时间久期因素，

时间距离过长则视为两个同价位之间并没有显著联系。我们合并同价位短时间内的交易单，并再对成交量方向进行拆分求得净值，对沪深 300 个股均做同样处理，从而得到沪深 300 指数的总体净成交量情况。这一过程中牵涉到的时间阀值与价格久期阀值当然都是主观定下的，通过样本期的测试，并尽可能采用自适应参数，因此在样本外实际交易中需要不断维护稳定性。

2、VPIN 的提出以及其在流动性风险控制方面的运用

对传统 PIN 模型的修正非常多，VPIN 修正的角度主要是成交量久期，它将标的日内成交量等分为 N 份，每次达到 $1/N$ 成交量计为一个成交量久期。VPIN 的优势在于不用估计 PIN 模型中的不可观测的待估参数 (α 信息事件发生概率, δ 坏消息发生概率, ε 不知情交易者订单到达率, μ 知情交易上订单到达率)，以及克服波动率聚集效应带来的干扰(以成交量久期来代理波动率聚集效应)。但另一方面，VPIN 与 PIN 对市场的交易环境有较强假设，在中国市场的实证目前没有特别显著的结论。感兴趣的读者可以参阅《Flow toxicity and liquidity in a high-frequency world》(O'Hara, David Easley, Marcos M.Lopez de Prado, 2011)。

3、VNET 的提出及其与价格久期的结合

VNET 首先需要将标的的日内价格区间均分为 N 份，这首先是一个主观选参数过程，其次在每一个价格久期中寻找变量来解释市场深度，例如：久期间隔时间 ptime，交易笔数，交易量，真实价格移动，平均成交单规模等等。而 VNET 的定义为：

$$VNET = \log |\sum (d_i vol_i)|$$

其中 d 是符号变量 (1 为买入, -1 为卖出), vol 即为成交量。一个常用的根据历史估计久期时间间隔的方法是 ACD 模型 (类似 ARMA, 一个条件点过程)：

$$\psi_t = \omega + \alpha_1 X_{t-1} + \beta_1 \psi_{t-1}$$

上图 ACD(1,1) 表示条件久期是过去条件久期与实际久期的函数，我们可以利用这个公式去估算上文提到的久期之间的时间间隔 ptime:

$$EPTIME_t = \omega + \alpha_1 PTIME_{t-1} + \beta_1 EPTIME_{t-1}$$

类似的我们可以引入各种变量来对 VNET (方向性的净成交量) 进行拟合，比如：

Spread: $\log(\text{ask}/\text{bid})$;

Num: $\log(\text{trades during the price-duration})$;

Vol: $\log(\text{total vol during the price-duration})$;

Num_spr: $\log(\text{trades occurring at large spreads})$ 即 spread 大于最小价格单位时的成交笔数；

Vol_spr: $\log(\text{vol occurring at large spreads})$ 即 spread 大于最小价格单位时的成交量；

Pjump: $\log(\text{absolute price change over duration})$;

Logetime: log(epetime);

Ptime_err: log(ptime/epetime);

从而能够逐步验证显著性得到一个解释性较好的拟合（以 IF 合约为例）：

$$VNET_t = \beta_0 + \beta_1 spread(-1) + \beta_2 vol(-1) + \beta_3 num(-1) + \beta_4 logetime + \beta_5 ptime_err$$

其中 num 表示一个久期周期中发生的交易次数，用于刻画交易密度，类似的还有 vol，也可用于刻画交易密集程度，并且可以描述给定 VNET 环境下买卖双方的不平衡程度。

三、下一阶段工作

我们下一步将做平台转移，我们正在开发能完整控制多事件并发的交易平台，策略也将移植到新的交易平台进行测试。

策略方面，随着我们对高频交易、交易细节理解的加深，我们将偏向于开发交易频率更高的策略，引入类 market maker 思想，后期我们会在这一方面继续与大家进行沟通交流。

四、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: 一号策略净值 (6月 25 日——12月 9 日)	3
图 2: 一号策略日收益情况.....	3
图 3: 高频策略大致分类.....	4
图 4: 成交小单、成交大单、市价单的到达率.....	5
图 5: 撤单与新挂单到达率.....	5
图 6: NYSE 权重股的成交单对角线效应.....	5
图 7: 成交量分解示意图.....	6
图 8: 成交量分解伪代码.....	6
图 9: 多事件并发控制.....	6
图 10: 撤单与撤单成功回报.....	7
图 11: 成交明细表举例.....	8

分析师与联系人简介

王红兵，西安交通大学管理科学与工程硕士，2004年进入证券行业从事金融工程研究，研究领域包括衍生品、量化选股、量化行业配置、事件投资等，坚持简单、实用的研究风格，曾开发了考虑分析师预期的权证定价模型、期现套利现货优化模型、行业选股量化框架、基于宏观、行业基本面以及市场情绪的行业配置模型、基于分析师的系列事件选股模型等。

夏钦，2011年毕业于美国 Brandeis University 并获得金融与经济学硕士，本科毕业于西南财经大学经济信息工程学院。2011年开始从事金融工程研究工作。主要研究领域包括：量化选股、指数增强、行业配置及衍生品交易模型。坚持实用、创新的研究风格，曾开发改进型市盈率指标CAPE、现金流择时体系、反转型行业配置模型及指数增强系列模型。历任华泰联合证券研究所金融工程研究员。

分析师承诺

作者具有中国证券业协会授予的证券投资咨询执业资格和相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

民生证券研究院：

北京：北京市东城区建国门内大街28号民生金融中心A座17层； 100005

上海：浦东新区银城中路488号太平金融大厦3903室； 200120

深圳：深圳市福田区深南大道7888号东海国际中心A座； 518040

免责声明

本报告仅供民生证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。

本报告是基于本公司认为可靠的已公开信息，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，在不同时期，本公司可发出与本报告所刊载的意见、推测不一致的报告，但本公司没有义务和责任及时更新本报告所涉及的内容并通知客户。

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。客户不应单纯依靠本报告所载的内容而取代个人的独立判断。本公司也不对因客户使用本报告而导致的任何可能的损失负任何责任。

本公司未确保本报告充分考虑到个别客户特殊的投資目标、财务状况或需要。本公司建议客户应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。

本公司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参与本报告所提及的公司的金融交易，亦可向有关公司提供或获取服务。本公司的一位或多位董事、高级职员或/和员工可能担任本报告所提及的公司的董事。

本公司及公司员工在当地法律允许的条件下可以向本报告涉及的公司提供或争取提供包括投资银行业务以及顾问、咨询业务在内的服务或业务支持。本公司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。

若本公司以外的金融机构发送本报告，则由该金融机构独自为此发送行为负责。该机构的客户应联系该机构以交易本报告提及的证券或要求获悉更详细的信息。

未经本公司事先书面授权许可，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。本公司版权所有并保留一切权利。

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，均为本公司的商标、服务标识及标记。