

## 因子方法论之一:基于日内模式的因子改进

### 以流动性因子为例

2018年2月22日

#### 研究结论

- 本篇报告为“因子方法论”系列报告第一篇。该系列将秉承“鱼不如渔”的基本思想,尝试透过现象看本质,总结出一些具有强普适性的规律结论或因子构造方法,为读者在各自的多因子研究工作中添砖加瓦。
- 传统的流动性因子忽略了A股市场中特定的日内交易结构信息,本文以此为出发点,给出了一个针对传统流动性因子的改进方案。
- 改进后的流动性因子,ICIR的绝对值达到4.44,IC月胜率高达88.9%,相比之下,改进之前的流动性因子ICIR绝对值只有2.44,月度胜率只有72.2%。
- 改进后的流动性因子在不同样本空间下均具有显著的选股能力,在股票5分组多空对冲情形下,其全样本年化超额收益达到32.4%,信息比率达3.98。在中证500成分股内选股,多空对冲年化收益达到25.2%,信息比率为2.46。在沪深300成分股内选股,多空对冲年化收益达到21.7%,信息比率为1.61。
- 本文所介绍的因子改进方案具有一定的普适性,同类型的因子都可以通过类似的步骤方法得到改造。
- **风险提示:**本报告所有的结论均使用历史数据回测得到,模型在未来存在失效风险。

证券分析师 高子剑

执业证书编号:S0600518010001

gaozj@dwzq.com.cn

证券分析师 魏建榕

执业证书编号:S0600517120003

weijr@dwzq.com.cn

## 目录

1. 前言 .....	4
2. 传统流动性因子介绍 .....	4
3. 传统流动性因子可能的改进方向 .....	5
4. A 股日内特定交易模式存在的证据 .....	6
5. 流动性因子遗漏的日内信息提取 .....	8
6. 改进后流动性因子的构造 .....	11
7. MIXLIQ 因子的绩效表现 .....	12
8. 总结 .....	17
9. 风险提示 .....	17

## 图表目录

图表 1: 原始流动性因子多空对冲净值曲线 .....	5
图表 2: 日内交易时段的分段示意图 .....	5
图表 3: 股票 A 与股票 B 的换手率分布 .....	6
图表 4: 不同时间段的“局部 R 方因子”均值之差的分布箱线图 .....	7
图表 5: 局部 R 方因子各个时段之间两两差值的 T 统计量 .....	8
图表 6: 5 个因子的对冲净值及 IC 表现统计 .....	9
图表 7: 提纯后的 5 个因子的对冲净值及 IC 表现统计 .....	11
图表 8: PURELIQ(2)、PURELIQ(4)、LIQ 的 IC 相关系数阵 .....	11
图表 9: MIXLIQ 与 LIQ 的 IC 序列对比 .....	13
图表 10: MIXLIQ 因子的行业内 IC 均值 .....	14
图表 11: MIXLIQ 与 LIQ 因子的市值分组下的 ICIR 对比 .....	14
图表 12: MIXLIQ 因子全样本选股等分 5 组多头净值曲线 .....	15
图表 13: MIXLIQ 与 LIQ 因子多空对冲表现 .....	16
图表 14: MIXLIQ 因子在三个样本空间内的选股对冲表现 .....	17

## 1. 前言

本着“鱼不如渔”这个理念，在全新的专题报告系列“因子方法论”中，我们将为读者倾情奉献多因子选股领域的研究心得，尝试透过现象看本质，总结出一些具有强普适性的规律结论或因子构造方法。

作为系列的开篇，本文以流动性因子为案例，详细阐述了一类对因子进行改进加强的方法流程，同样的方法可以适用于所有类似构造模式的其他因子。

## 2. 传统流动性因子介绍

流动性因子描述了资产的流动性，资产的流动性强弱可以定义为：在不引起市场剧烈波动的前提下，短期能快速买入或者卖出的资产数量的多少。对于股票而言，在微观层面上，资产的流动性取决于订单簿的厚度，以及盘口的价差高低，在宏观层面上，也有多种方法用以描述资产的流动性，如换手率(TR)，或 Amihud(2002)[2]提出的非流动性比率(ILLIQ)等指标较为常用。本文所探讨的流动性指标为基于换手率的流动性因子。

首先，本文将介绍传统的换手率因子LIQ的计算方法。令 $TR_{s,t}$ 为股票 $s$ 第 $t$ 日的换手率，若要计算 $t$ 时刻该股票的流动性因子，首先，将 $t-19$ 至 $t$ 日共20个交易日的换手率相加，得到最近20个交易日的总换手率，为了尽量保证因子在横截面上分布的正态性，对总换手率取对数，完成第一步，得到原始因子 $F_{s,t}^{raw}$ 。

$$F_{s,t}^{raw} = \ln \left( \sum_{i=0}^{19} TR_{s,t-i} \right)$$

实证表明，原始因子 $F_{s,t}^{raw}$ 与市值具有很强的负相关性，即市值越大的股票，换手率倾向于越低，而市值越小的股票，换手率倾向于越高。为了消除市值的影响，在横截面 $t$ 上将原始因子 $F_{s,t}^{raw}$ 作为因变量，以对数流通市值因子 $\ln FMV_{s,t}$ (下文简称为 $\ln FMV$ )为自变量，拟合回归方程：

$$F_{s,t}^{raw} = a_t + b_t \ln FMV_{s,t} + LIQ_{s,t}$$

将得到的残差项记为 $LIQ_{s,t}$ ， $LIQ_{s,t}$ 即为基于换手率得到的流动性因子，该因子越大，说明股票的流动性越强。为了简化表述，下文将 $LIQ_{s,t}$ 记为 $LIQ$ 因子。

为了测试 $LIQ$ 因子的选股能力，我们做了一个简单的回溯：

回溯时间段设置为 2013/04/26-2017/10/31；

样本空间为全体 A 股，在其中剔除上市不足 60 个交易日的次新股，并将 ST 个股，以及停牌个股一并剔除；

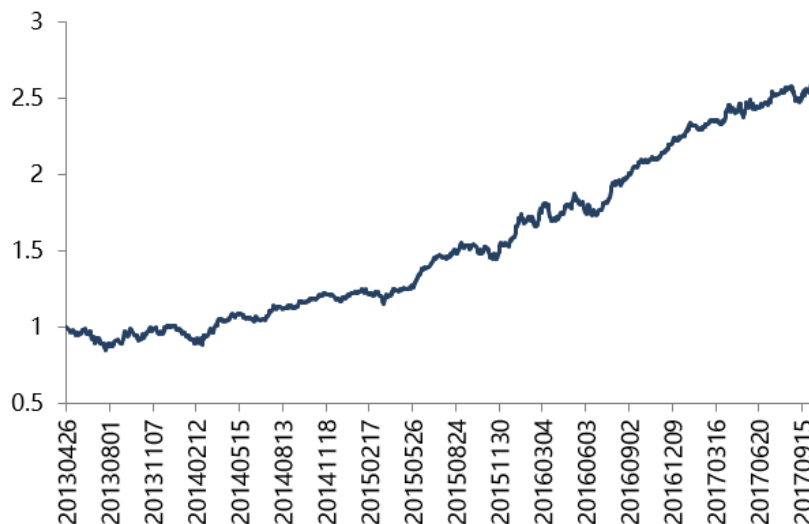
调仓频率为月度，每个月月底计算样本中个股的因子值，将个股按照因子值从小到大排序，等分 5 组，做多因子值最小的一组，做空因子值最大的一组，从而获得多空组合净值；

模拟交易以第二个月首个交易日的开盘价作为股票交易价格，不买入涨

停股，不卖出跌停股。

执行上述回测，可以得到收益曲线，如图表 1 所示：

图表 1：原始流动性因子多空对冲净值曲线



资料来源：Wind 资讯，东吴证券研究所

2013 年 4 月 26 日至 2017 年 10 月 31 日，多空组合可以获得年化 24.4% 的对冲收益，多空对冲的信息比率为 1.73，月度胜率为 72.2%，最大回撤为 15.4%，该因子的年化 ICIR 达到了 -2.44。

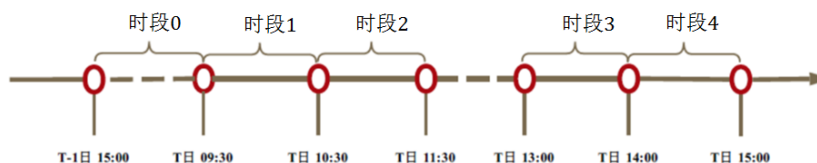
关于为何流动性较小的因子收益倾向高于流动性强的因子，在 Amihud 和 Mendelson (1986) [3] 提出流动性溢价理论有所解释，他们认为，为了补偿投资者因流动性而损失的交易成本，流动性较高的资产预期收益倾向于低于流动性较低的资产的预期收益。

### 3. 传统流动性因子可能的改进方向

虽然上文所介绍的基于每日换手率计算的流动性因子在回测结果上已经非常有效了，但其算法过于简单，所使用的数据也过于粗糙，频率过低，本篇报告将尝试在更加微观的层面上对其进行改进。

借鉴我们以往的研究，按照下图将每天的交易时间分为 5 个互不相交的时段：隔夜、9:30-10:30、10:30-11:30、13:00-14:00、14:00-15:00。

图表 2：日内交易时段的分段示意图



资料来源：Wind 资讯，东吴证券研究所

在下文中，使用  $TR_{s,t}^{(0)}$ 、 $TR_{s,t}^{(1)}$ 、 $TR_{s,t}^{(2)}$ 、 $TR_{s,t}^{(3)}$ 、 $TR_{s,t}^{(4)}$  分别代表第  $s$  只股票第

t日的隔夜时段、9:30-10:30、10:30-11:30、13:00-14:00、14:00-15:00 的换手率。  
则股票s在t日的换手率可以被写为

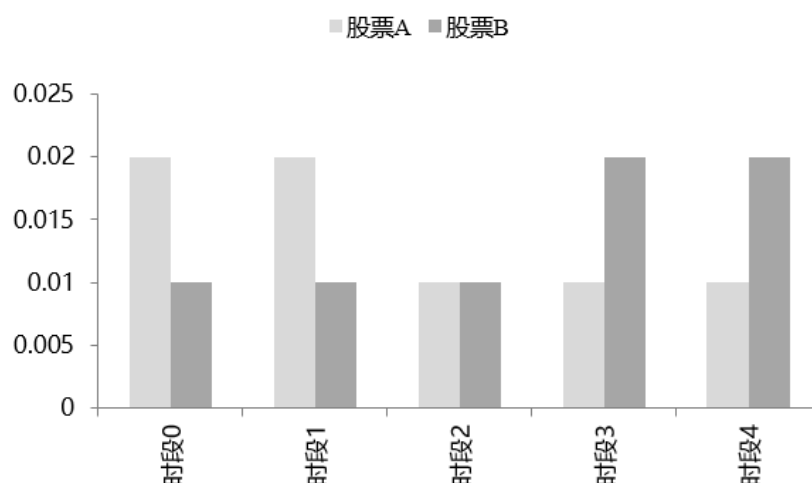
$$TR_{s,t} = TR_{s,t}^{(0)} + TR_{s,t}^{(1)} + TR_{s,t}^{(2)} + TR_{s,t}^{(3)} + TR_{s,t}^{(4)}$$

因此，第一部分中的累积对数换手率的对数因子  $F_{s,t}^{raw}$  可以写为如下形式：

$$F_{s,t}^{raw} = \ln \left( \sum_{i=0}^{19} TR_{s,t}^{(0)} + \sum_{i=0}^{19} TR_{s,t}^{(1)} + \sum_{i=0}^{19} TR_{s,t}^{(2)} + \sum_{i=0}^{19} TR_{s,t}^{(3)} + \sum_{i=0}^{19} TR_{s,t}^{(4)} \right)$$

在该计算中，每个时间段的换手率所贡献的权重以等权形式出现。假设两只流通市值相同的股票在过去 20 个交易日每天每个时段的换手率分布如图表 3 所示，由于两只股票全天累积换手率相同，在  $F_{s,t}^{raw}$  的计算公式中，两只股票的因子完全相同。因此，股票 A 在上午流动性强于股票 B，而股票 B 在下午的流动性强于股票 A，这种差异性信息，在日频的因子计算中，并没有被表现出来。

图表 3：股票 A 与股票 B 的换手率分布



资料来源：Wind 资讯，东吴证券研究所

#### 4. A 股日内特定交易模式存在的证据

从上一部分的分析可以看出，传统的流动性因子将每个交易日发生的所有交易量合并为一个整体，而并没有考虑日内交易模式的问题。

事实上，由于信息传递以及日内每个时段交易者结构不同等现象的存在，导致 A 股始终呈现出较为明显的特定的日内交易结构现象。这导致不同交易时段出现的同一种现象，其背后所包含的信息可能是完全不同的。

作为案例，我们从个股走势的关联性角度出发，给出一个不同时段的数据包含不同信息的实证证据。根据 CAPM 模型，个股的走势可以被分解为两个部分：与市场趋同的部分，以及市场收益所无法解释的特质收益部分。对 CAPM 模型进行拟合，所得到的回归方程 R 方越大，个股走势更加随波逐

流，说明市场中知情交易所占比重越低。反之，个股与市场整体走势相关性越低，市场中知情交易所占比重越高。

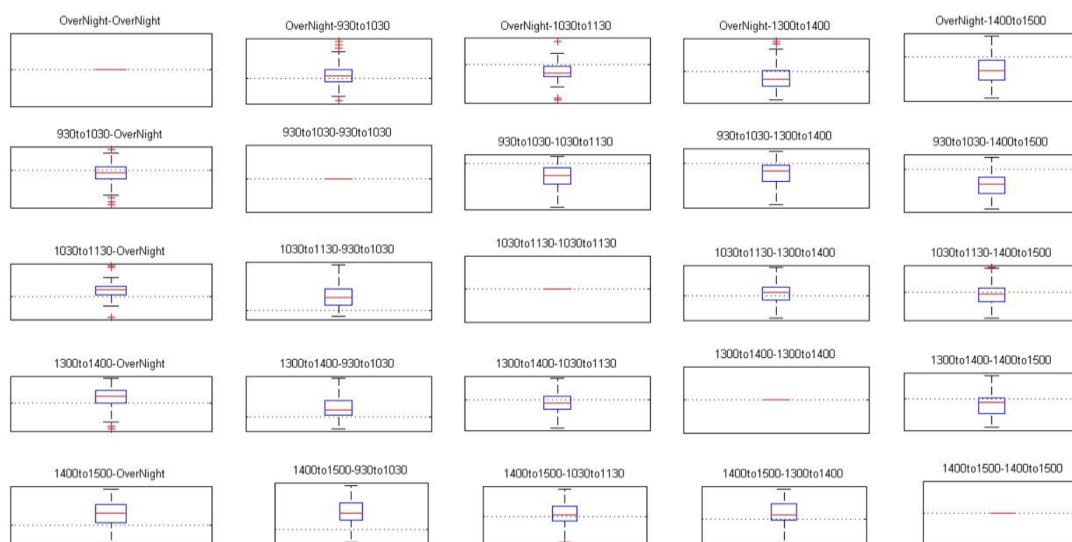
以  $R_s^{(j)}$  代表股票  $s$  在日内第  $j$  个时段的股票收益率，以  $R_I^{(j)}$  代表上证指数在日内第  $j$  个时段的收益率，仿照 CAPM 模型的形式，拟合如下回归方程：

$$R_s^{(j)} = a_s + b_s R_I^{(j)} + \varepsilon_s^{(j)}$$

其中， $a_s$  与  $b_s$  为回归系数，在  $t$  时刻，所使用的回归样本为股票  $s$  在  $t-19$  日到  $t$  日共 20 个交易日的个股每日分时段收益率数据，以及上证指数每日分时段收益率数据。回归所得到的  $R$  方衡量了该股票在特定时间段相对大盘收益的相关性高低，我们将分别使用 5 段行情信息计算得到的  $R$  方因子统称为“局部  $R$  方因子”。

我们计算每只股票的在 5 个交易时段的“局部  $R$  方因子”，然后，将其在横截面上取均值，得到 5 个“局部  $R$  方因子”的均值，在 2013 年 4 月 26 日至 2017 年 10 月底共 54 个月，每个月月底，对“局部  $R$  方因子”均值两两相减，得到如下图的箱线图矩阵，其中，每个子图中横虚线为 0 值所在的位置，箱线图中红色的实线为样本中位数，箱体的上沿与下沿分别代表样本的 75 分位数与 25 分位数。例如，箱线图矩阵的第一行，第五列的箱线图，代表隔夜时段的“局部  $R$  方因子”均值与 14:00-15:00 时段的“局部  $R$  方因子”的均值之差的箱线图。

图表 4：不同时段的“局部  $R$  方因子”均值之差的分布箱线图



资料来源：Wind 资讯，东吴证券研究所

为了更加严谨地说明每个时段股票走势的同度之间地差异，图表 5 给出了“局部  $R$  方因子”均值两两之差的 T 检验统计量。例如，第一行，第五列的元素 -6.59，代表了隔夜的“局部  $R$  方因子”的均值与 14:00-15:00 时段的“局部  $R$  方因子”的均值之差的 T 统计量为 -6.59 < -2，说明隔夜的“局部  $R$  方因子”的均值与 14:00-15:00 时段的“局部  $R$  方因子”的均值之差在 95% 置信水平小于 0。对于每个表中的 T 值，超过 2，则认为两者之差的均值在统计意义上显著大于 0，T 值小于 -2，则认为两者之差的均值在统计意义上显著小



于 0。

图表 5：局部 R 方因子各个时段之间两两差值的 T 统计量

	隔夜	9:30-10:30	10:30-11:30	13:00-14:00	14:00-15:00
隔夜	NaN	2.63	-5.00	-3.34	-6.59
9:30-10:30	-2.63	NaN	-8.72	-6.24	-8.84
10:30-11:30	5.00	8.72	NaN	2.11	-1.16
13:00-14:00	3.34	6.24	-2.11	NaN	-3.57
14:00-15:00	6.59	8.84	1.16	3.57	NaN

资料来源：Wind 资讯，东吴证券研究所

上表中几乎所有时段的两两配对之差的 T 统计量绝对值均大于 2，这说明个股在日内不同交易时段的趋同水平是始终存在显著差异的。

根据上文给出的箱线图矩阵，以及统计检验表格，可以大致得到如下股票趋同度水平的大小关系：[14:00-15:00]>[10:30-11:30]>[13:00-14:00]>隔夜>[9:30-10:30]。即每个交易日的最后一个小时，个股与市场走势的趋同度最高，而每日开盘后第一个连续交易的小时，个股的个性最足，这反映出市场中知情交易者更倾向于在上午第一个小时进行交易，而在下午收盘前一小时，市场中的信息基本被反映完毕，参与交易的更多是噪声交易者。这点发现与相关文献研究结论也不谋而合[4][5]。

这个案例也反映出，A 股市场的日内不同交易时段所包含的信息确实具有显著差异。这也就为接下来的进一步分析提供了数据支撑。

## 5. 流动性因子遗漏的日内信息提取

接下来，我们将挖掘日频的流动性因子 LIQ 所遗漏的信息。使用第一部分同样的计算方式，我们可以得到每只股票的“局部流动性因子”： $LIQ_{s,t}^{(0)}$ 、 $LIQ_{s,t}^{(1)}$ 、 $LIQ_{s,t}^{(2)}$ 、 $LIQ_{s,t}^{(3)}$ 、 $LIQ_{s,t}^{(4)}$ ，分别代表股票在过去 20 个交易日在隔夜时段、9:30-10:30、10:30-11:30、13:00-14:00、14:00-15:00 这 5 个时段的流动性，在下文中，这五个因子分别被简写为 LIQ(0)、LIQ(1)、LIQ(2)、LIQ(3)、LIQ(4)。

以  $LIQ_{s,t}^{(4)}$  为例，仿照  $LIQ_{s,t}$  的构造方法，首先，将  $t-19$  至  $t$  日共 20 个交易日的 14:00-15:00 时段的换手率相加，而后取对数，得到  $F_{s,t}^{Raw(4)}$

$$F_{s,t}^{Raw(4)} = \ln \left( \sum_{i=0}^{19} TR_{s,t-i}^{(4)} \right)$$

其中， $TR_{s,t-i}^{(4)}$  为股票  $s$  在  $t-i$  日 14:00-15:00 时段的换手率。而后，将  $F_{s,t}^{Raw(4)}$

因子在横截面上关于对数流通市值因子做回归：

$$F_{s,t}^{Raw(4)} = a_t + b_t \ln FMV_{s,t} + LIQ_{s,t}^{(4)}$$

其中， $a_t$ 、 $b_t$  为回归系数，截面回归的残差项即为我们所需要的得到  $LIQ_{s,t}^{(4)}$ ，即 LIQ(4) 因子。

从  $LIQ(0)$ 、 $LIQ(1)$ 、 $LIQ(2)$ 、 $LIQ(3)$ 、 $LIQ(4)$  的计算公式来看，这 5 个因子都可以被近似看作原始因子 LIQ 的一部分，与 LIQ 必然有较高的正相关性。但我们更加关心的是这 5 个因子所能提供的除 LIQ 之外的选股增量信息。



借鉴 Fama-Macbeth 回归, 可以在 LIQ(0)、LIQ(1)、LIQ(2)、LIQ(3)、LIQ(4) 因子中剔除 LIQ 所包含的线性信息, 从而得到 5 个纯净的“局部流动性因子”:  $\text{pureLIQ}_{s,t}^{(0)}$ 、 $\text{pureLIQ}_{s,t}^{(1)}$ 、 $\text{pureLIQ}_{s,t}^{(2)}$ 、 $\text{pureLIQ}_{s,t}^{(3)}$ 、 $\text{pureLIQ}_{s,t}^{(4)}$ , 在下文将其分别简记为  $\text{pureLIQ}(0)$ 、 $\text{pureLIQ}(1)$ 、 $\text{pureLIQ}(2)$ 、 $\text{pureLIQ}(3)$ 、 $\text{pureLIQ}(4)$ 。同样, 以  $\text{pureLIQ}_{s,t}^{(4)}$  为例, 我们详细介绍其计算方法, 使用的方法依然为简单的横截面回归:

$$\text{LIQ}_{s,t}^{(4)} = a_t + b_t \text{LIQ}_{s,t} + \text{pureLIQ}_{s,t}^{(4)}$$

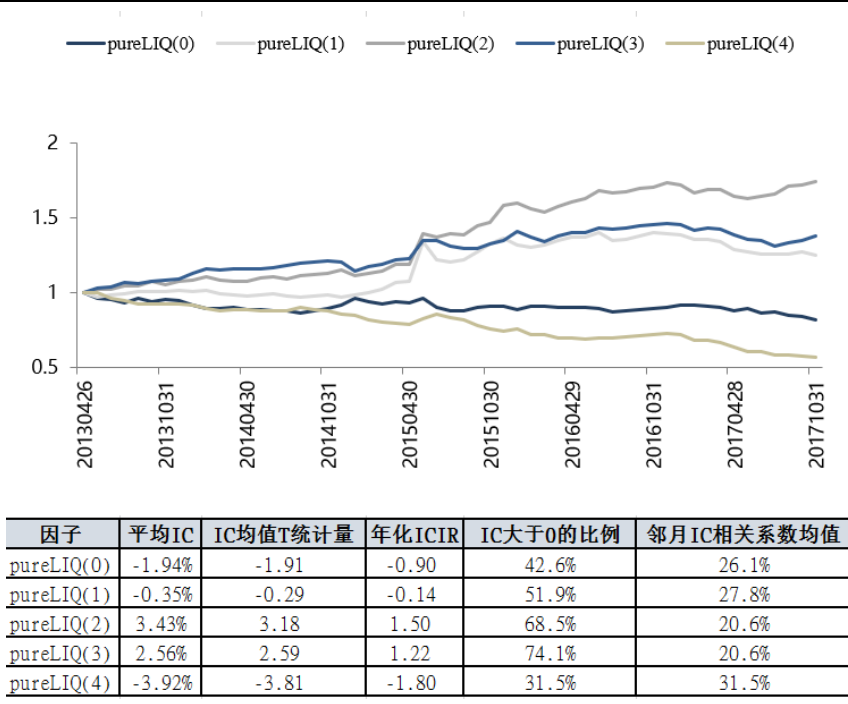
横截面回归得到的残差项即为所要最终得到的“纯净的局部流动性因子”  $\text{pureLIQ}_{s,t}^{(4)}$ , 即  $\text{pureLIQ}(4)$ 。使用同样的方法, 我们可以计算得到另外四个“纯净的局部流动性因子”  $\text{pureLIQ}(0)$ 、 $\text{pureLIQ}(1)$ 、 $\text{pureLIQ}(2)$ 、 $\text{pureLIQ}(3)$ 。

可以严格证明, 所得到的  $\text{pureLIQ}(0)$ 、 $\text{pureLIQ}(1)$ 、 $\text{pureLIQ}(2)$ 、 $\text{pureLIQ}(3)$ 、 $\text{pureLIQ}(4)$  因子在横截面上与原始的流动性因子 LIQ 以及对数流通市值因子 LnFMV 线性不相关, 若这 5 个“纯净的局部流动性因子”仍然能够贡献选股能力, 则可以将其看作原始的流动性因子 LIQ 所没有捕捉到的选股信息。

接下来, 我们测试这 5 个“纯净的局部流动性因子”的选股能力, 测试流程与第一部分中对原始流动性因子 LIQ 方法基本相同, 唯一的, 也是最重要的一点区别在于, 多空对冲收益是以因子值最大的一组为多头因子值最小的一组为空头。

图表 6 给出了  $\text{pureLIQ}(0)$ 、 $\text{pureLIQ}(1)$ 、 $\text{pureLIQ}(2)$ 、 $\text{pureLIQ}(3)$ 、 $\text{pureLIQ}(4)$  五个因子的多空对冲曲线, 以及各个因子的 IC 表现。

图表 6: 5 个因子的对冲净值及 IC 表现统计



资料来源: Wind 资讯, 东吴证券研究所

净值表现以及平均 IC 可以看出, 隔夜时段( $\text{pureLIQ}(0)$ )与 14:00-15:00 时

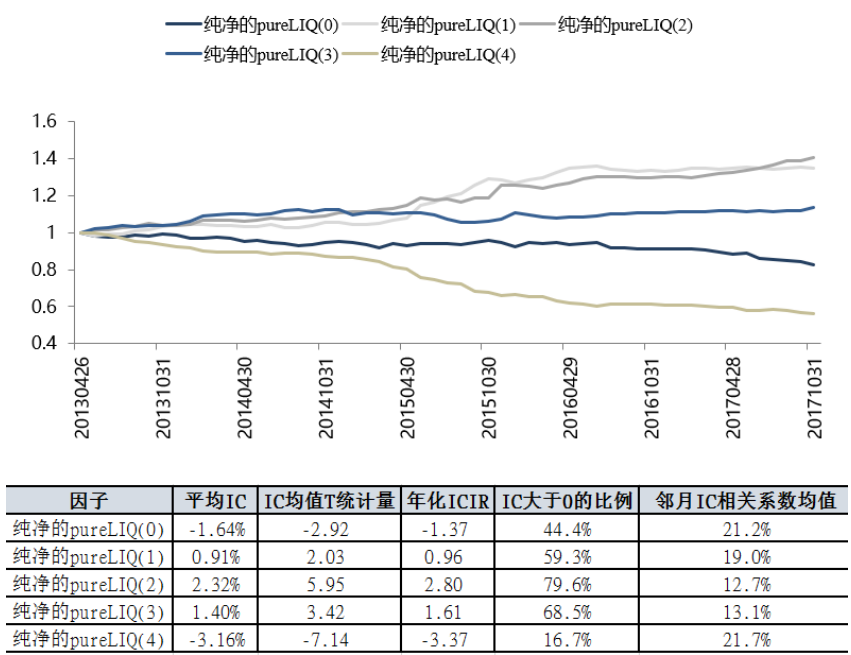
段(pureLIQ(4))所构造的“纯净的局部流动性因子”可以贡献负向的收益。其中，由 IC 均值的 T 统计量可得，14:00-15:00 时段的pureLIQ(4)因子贡献负向收益在 5%显著水平下显著。10:30-11:30( pureLIQ(2) )，13:00-14:00(pureLIQ(3))两个时段所构造因子可以贡献正向的收益。最后，9:30-10:30 的“纯净的局部流动性因子” pureLIQ(1)的 IC 均值与多空对冲净值曲线所表现的因子方向相互矛盾，这是由于使用该因子对个股 5 分组后，净值曲线单调性较差所导致，且该因子的 IC 均值 T 统计量仅为-0.29，远低于统计意义显著的临界值 2。

综合以上分析，5 个因子中，能提供显著选股能力的因子为pureLIQ(2)、pureLIQ(3)、pureLIQ(4)，分别对应 10:30-11:30、13:00-14:00、14:00-15:00 三个时间段。且从 IC 均值显著程度进行排序，有pureLIQ(4) > pureLIQ(2) > pureLIQ(3)，回忆上一部分，我们对个股相对于大盘的趋同程度的，最后的结论为 14:00-15:00、10:30-11:30、13:00-14:00 依次为个股与大盘趋同度最高的三个时段，其时间段恰好分别与pureLIQ(4)、pureLIQ(2)、pureLIQ(3)三因子所对应时间段相同，这说明在知情交易者不活跃的时段，我们所构造的“纯净的局部流动性因子”能提供更多传统流动性因子所无法挖掘的增量信息，而在知情交易者活跃的时段所构造的另外两个因子，其因子方向飘忽不定。这个现象也许可以从侧面说明，当市场中噪音交易占比较高时，使用换手率对个股的流动性的估计更加准确。

14:00-15:00 时段的因子 pureLIQ(4)收益方向为负，与传统的流动性因子的收益方向相同，表明在传统流动性因子的构造中，对于 14:00-15:00 这段时间所给予的权重不足，以至于其中的信息并没有被完全捕捉到，而对于 10:30-11:30、13:00-14:00 这两个时间段的 pureLIQ(2)、pureLIQ(3)因子收益为正，表明在原始流动性因子LIQ的构造中，这两个时间段的换手率被赋予了过高的权重。

为了验证pureLIQ(0)、pureLIQ(1)、pureLIQ(2)、pureLIQ(3)、pureLIQ(4)五个因子的收益能否被其他已知的风格因子或行业因子所解释，我们也分别检验了这五个因子剔除风格和行业影响后的表现，其中，风格因子共有 11 个，包括 Beta、短期动量、长期动量、价值、盈利、成长、杠杆、波动、流动性、规模、非线性规模，行业因子使用申万一级行业分类。下图中展示了剔除风格与行业后纯净因子的多空对冲净值以及 IC 表现。

图表 7：提纯后的 5 个因子的对冲净值及 IC 表现统计



资料来源：Wind 资讯，东吴证券研究所

提纯后的 5 个因子的 IC 均值 T 统计量的绝对值均有不同程度的提升，甚至原本并不显著的pureLIQ(1)因子在纯净化之后，T 统计量的绝对值也超过了 2。这表明，本文所提出的这几个“纯净的局部流动性因子”的选股收益并非来自于其他已知的风格因子，也就是说，这些因子的确贡献了额外的 alpha 收益。

## 6. 改进后流动性因子的构造

在这一部分中，我们将使用上文所得到的结果完成对已有的传统流动性因子LIQ的改造。

综合上一部分中所得到的统计结果，使用两个最强的“纯净的局部流动性因子”参与最后的因子改进，它们分别是：基于 10:30-11:30 的行情计算的因子pureLIQ(2)，以及基于 14:00-15:00 的行情计算的因子pureLIQ(4)。

上文已经提到，pureLIQ(2)与pureLIQ(4)两因子在横截面上与原始流动性因子LIQ正交，因此，不存由于因子在横截面的高相关性而导致的因子的 IC 序列之间的高度相关性，可以预期，pureLIQ(2)与pureLIQ(4)的 IC 序列与LIQ的 IC 序列之间的相关系数不会太高。图表 8 给出了三个因子的 IC 序列之间两两的相关系数。

图表 8：pureLIQ(2)、pureLIQ(4)、LIQ 的 IC 相关系数阵

	pureLIQ(2):IC	pureLIQ(4):IC	LIQ:IC
pureLIQ(2):IC	100.00%	43.94%	42.46%
pureLIQ(4):IC	43.94%	100.00%	-5.80%
LIQ:IC	42.46%	-5.80%	100.00%

资料来源：Wind 资讯，东吴证券研究所

相关性结果的分析甚至超出了我们的预期，对于因子收益方向相反的 pureLIQ(1)与 LIQ 因子对，以及 pureLIQ(2)与 pureLIQ(4)因子对，其 IC 序列之间两两的相关系数不但不为负，反而均为较大的正值，分别为 42.5%和 43.9%，将它们线性叠加，可在保证期望收益的前提下，大幅降低收益曲线的波动，从而大幅提高收益的稳定性，并最终获得更高的收益风险比。同样，对于因子收益方向相同的两因子 pureLIQ(4)以及 LIQ，其 IC 序列的相关系数为 -5.8%，说明两因子间的线性叠加同样可以在期望收益保证的前提下，大量抵消收益曲线的波动性。

将多个因子进行加权线性叠加的方式有很多种，例如，以往研究中，我们经常使用最优 ICIR 的加权方式。另外，还有其他相对更加稳健的加权方式，例如 IR 加权，最大化复合因子的 IC 加权等等。但这不是本篇报告的重点，本着稳健且不过度优化的原则，我们选择了最原始的等权加权方式，按照因子的收益方向，对 pureLIQ(2)、pureLIQ(4)、LIQ 分别赋予 -1、1、1 的权重。

具体而言，就是通过横截面标准化后的三个因子按照 (-1,1,1) 的权重进行线性叠加，得到混合流动性因子 MixLIQ<sub>s,t</sub>，下文简记为 **MixLIQ** 因子，该因子的计算公式表达如下：

$$\text{MixLIQ}_{s,t} = -\frac{\text{pureLIQ}_{s,t}^{(2)} - \text{mean}(\text{pureLIQ}_{s,t}^{(2)})}{\text{std}(\text{pureLIQ}_{s,t}^{(2)})} + \frac{\text{pureLIQ}_{s,t}^{(4)} - \text{mean}(\text{pureLIQ}_{s,t}^{(4)})}{\text{std}(\text{pureLIQ}_{s,t}^{(4)})} + \frac{\text{LIQ}_{s,t} - \text{mean}(\text{LIQ}_{s,t})}{\text{std}(\text{LIQ}_{s,t})}$$

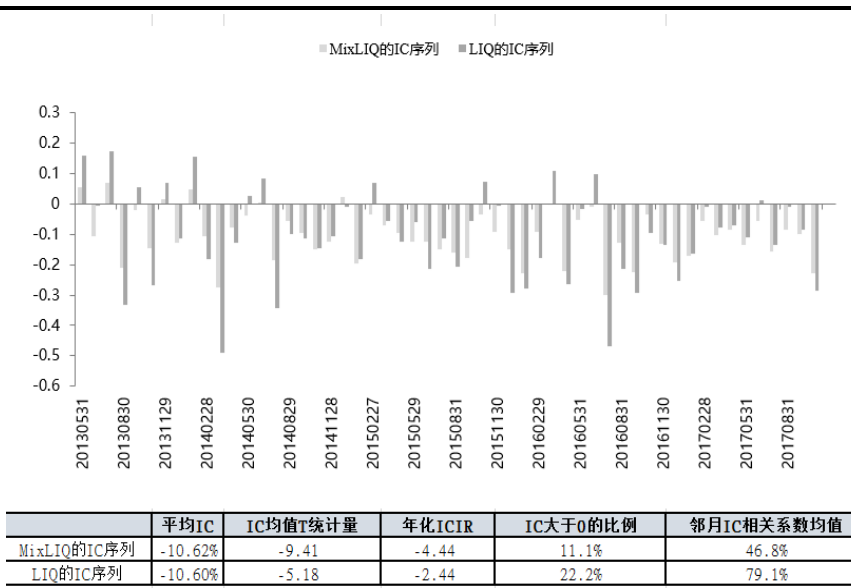
所得到的混合流动性因子 MixLIQ 与原始的流动性因子 LIQ 的收益方向相同，因子值越小，个股的预期收益率越高，因子值越高，个股的预期收益越低。

## 7. MixLIQ 因子的绩效表现

按照第一部分所给出的回测方案，我们对 MixLIQ 因子的选股能力进行了回测。

图表 9 展示了 MixLIQ 因子与 LIQ 因子各自的 IC 序列对比。

图表 9: MixLIQ 与 LIQ 的 IC 序列对比



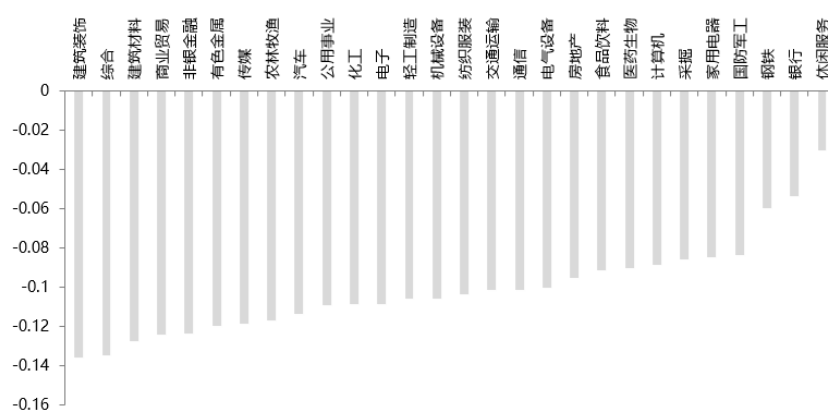
资料来源: Wind 资讯, 东吴证券研究所

我们发现, MixLIQ因子的 IC 序列在均值上并没有显著提升, 但其在 IC 的波动上做到了极大的降低, 这与我们在前面做因子间 IC 序列相关性分析时所预期的结果完全一致, IC 波动的下降致使其均值 T 统计量以及年化 ICIR 的绝对值获得了巨大的提升, 其中, ICIR 原始因子的-2.44, 提升了将近一倍, 达到了-4.44。预测失败的月份(IC 值大于 0)的比例也从 22.2%降低到 11.1%。

MixLIQ因子相对于原始因子LIQ唯一的缺陷在于, 因子邻月 IC 相关系数的均值从原始的 79.1%下降到 46.8%, 这意味着换仓时换手率会有一定程度的提高。若以换手率正比于 $\sqrt{1-\rho}$ 的准则衡量, 新因子相对于原始因子的换手率提高了 59.5%。但 46.8%的邻月因子相关性, 在同类型收益风险比的已知因子中, 仍然处较高的水平。

为了测试MixLIQ因子在每个行业中的选股能力, 图表 10 展示了该因子在每个行业下的月度 IC 均值, 在所有 28 个申万一级行业内, 该因子的均能产生不错的选股效果, 其中, 最强的建筑装饰行业内, 因子的 IC 均值达到近 -14%, 即便在最弱的休闲服务行业, IC 均值也能达到-3%。以上结果表明, 该因子具有较强的行业普适性。

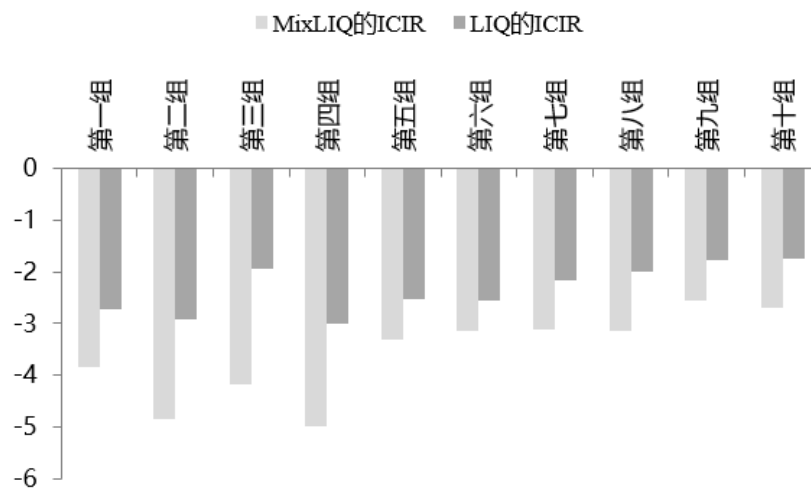
图表 10: MixLIQ 因子的行业内 IC 均值



资料来源: Wind 资讯, 东吴证券研究所

在实证研究中, 我们发现, 原始的换手率因子 LIQ 具有较强的市值衰减特性, 即在大市值个股中, 该因子的表现没有小市值的股票中好, 这导致该因子在沪深 300 或中证 500 内的选股能力远不及中证 1000 内的选股能力强, 从而导致其真实实用性大打折扣。为了检验新因子的市值衰减特性, 我们将样本内的个股按照流通市值的高低将样本内个股排序, 而后将股票等分十组, 第一组为流通市值最小的一组, 第十组为流通市值最大的一组, 在每一组内分别计算因子与下期收益率的 IC 值, 进而得到测试时间段内每个市值区间内 MixLIQ 因子的 ICIR 值, 图表 11 给出了原始流动性因子 LIQ 与新因子 MixLIQ 在不同流通市值区间内的 ICIR 对比。

图表 11: MixLIQ 与 LIQ 因子的市值分组下的 ICIR 对比



资料来源: Wind 资讯, 东吴证券研究所

结果显示, 新因子 MixLIQ 也存在一定的市值衰减特征, 但在每个市值区间内, 改进后的 MixLIQ 因子的 ICIR 均大幅优于原始因子 LIQ 的 ICIR 值。新因子在每个市值区间内的 ICIR 表现均至少达到年化-2.70 的水平。

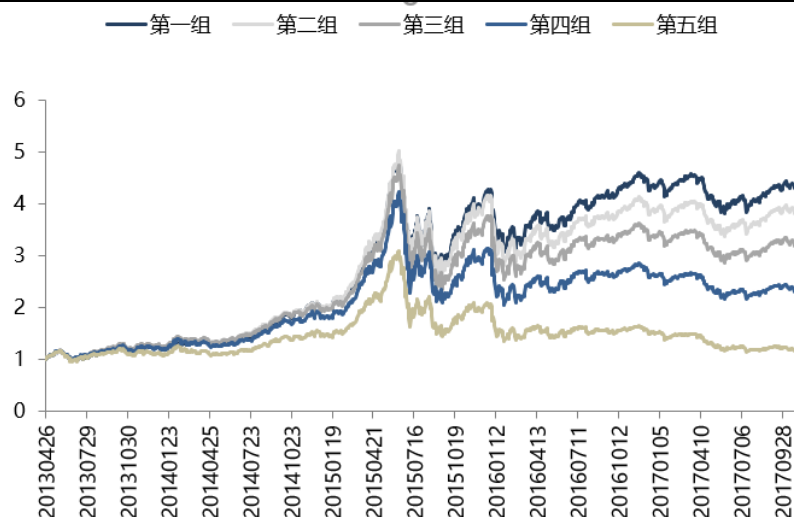
为了验证新因子实际的收益水平, 以及因子选股能力的单调性, 我们按照第一部分所给出的回测框架, 根据股票 5 分组, 组内等权配置的分组方法,



获得 5 条多头净值曲线。其中，因子最小的一组记为第一组，因子最大的一组记为第五组。

图表 12 展示了 5 分组的中每组的净值表现。5 组的收益表现具有较好的单调性，但值得注意的是，新因子继承了原始的流动性因子的一项缺点——因子的收益多数来源于空头。具体在图中的表现为，偏多头的第一组与第二组的收益差距小于偏空头的第四组与第五组的收益差距。

图表 12: MixLIQ 因子全样本选股等分 5 组多头净值曲线

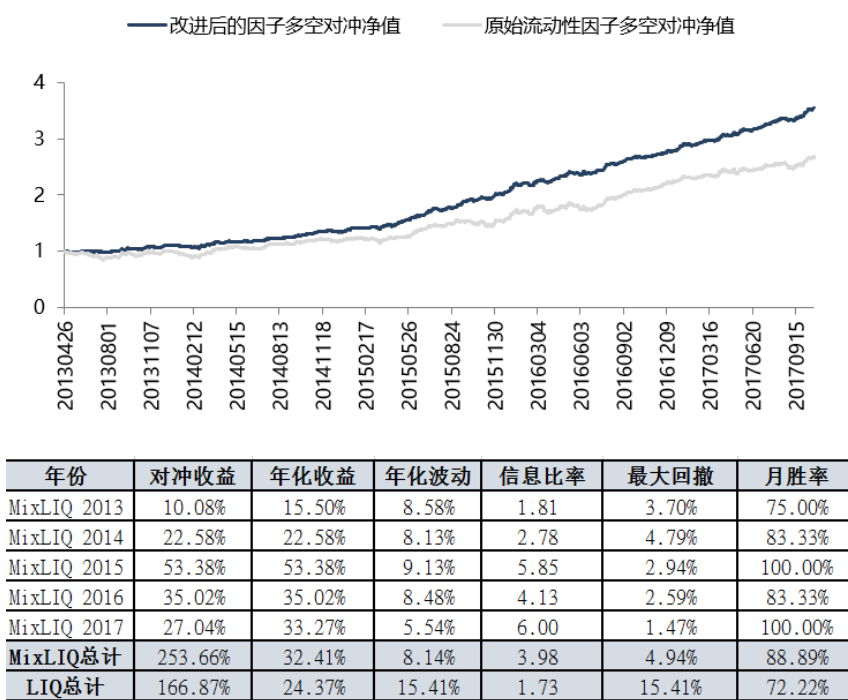


资料来源：Wind 资讯，东吴证券研究所

将第一组股票组合作为多头，将第五组股票组合作为空头，多空对冲后得到该因子在全样本下的多空对冲收益曲线，图表 13 展示了新因子对冲组合净值与原始因子对冲组合净值的对比，该图表同时展示了新因子分年的多空对冲收益统计，以及整个测试时间段的多空对冲收益统计，并同时附以原始因子的多空对冲组合的收益统计作为对照。



图表 13: MixLIQ 与 LIQ 因子多空对冲表现



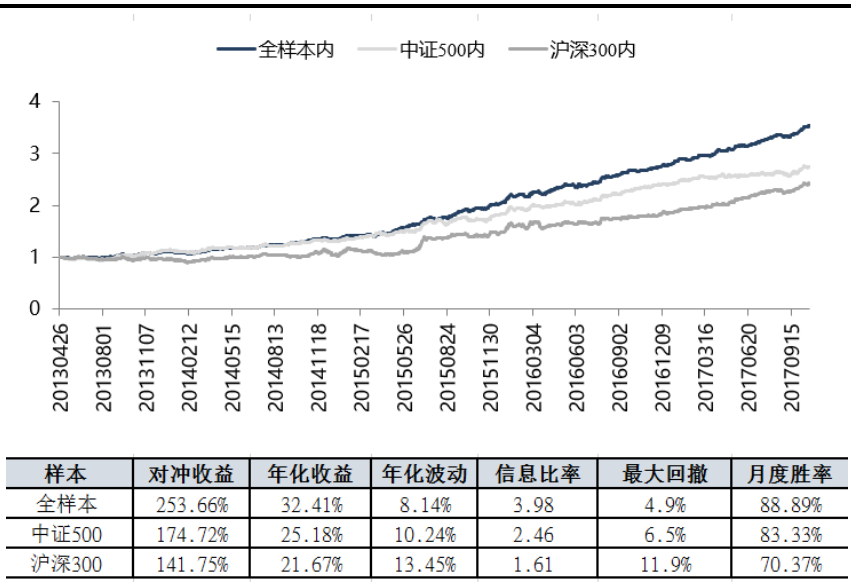
资料来源: Wind 资讯, 东吴证券研究所

结果显示, 从 2013 年 5 月开始到 2017 年 10 月底, 新因子 MixLIQ 的多空对冲年胜率达到 100%, 月胜率达到 88.9%, 其中, 2015 年与 2017 年两年的月胜率达到 100%。

新因子在全测试时段内的多空对冲收益由原来的年化 24.4% 提升至 32.4%, 最大回撤由 15.4% 下降至 4.9%, 信息比率由 1.73 提高至 3.98, 月胜率由 72.2% 提升至 88.9%。年化收益的提高与年化波动的下降, 提高了因子的 Alpha 性质, 降低了其 Beta 性质, 使得 MixLIQ 由 Beta 因子蜕变成成为优秀的 Alpha 因子。

为了验证该因子在不同的样本空间内的表现, 我们同样检验了 MixLIQ 因子在中证 500 与沪深 300 内做多对冲的净值表现。图表 14 展示了 MixLIQ 因子在全样本、中证 500 成分股、沪深 300 成分股内多空对冲的净值曲线和收益统计。

图表 14: MixLIQ 因子在三个样本空间内的选股对冲表现



资料来源: Wind 资讯, 东吴证券研究所

与之前 IC 表现出市值衰减的特性一致, 多空对冲净值表现的排序为全样本>中证 500>沪深 300。在中证 500 成分股内, 多空对冲的年化收益为 25.2%, 最大回撤为 6.5%, 信息比率为 2.46, 月度胜率为 83.3%。而在沪深 300 成分股内, 多空对冲的年化收益降低为 21.7%, 最大回撤增加至 11.9%, 信息比率降低至 1.61, 月胜率降低至 70.4%, 但仍然处于还不错的水平。

## 8. 总结

在本篇报告中, 我们利用更加精细化的日内模式对原始的以日频数据计算的流动性因子进行了结构上的改进。实证结果表明, 因子的改进结果令人振奋, 新因子相对于原始因子, 不但降低了波动, 同时提高了收益, 从而使得新因子由原来的 Beta 因子进化成为了优秀的 Alpha 因子。

在报告的最后, 我们再次重申“鱼不如渔”这个本系列报告将始终秉承的研究理念, 本篇报告的实证结果固然令人振奋, 但我们更想为读者展示的是对因子的改造过程。事实上, 同类型的因子都可以通过类似的步骤方法得到改造, 因此, 本篇报告更重要的核心思想在于, 提供了一个改进因子的普适方案。在这里我们期待能够抛砖引玉, 让该方法在不同的读者手里结出不一样的果实。

## 9. 风险提示

本报告所有的结论均使用历史数据回测得到, 模型在未来存在失效风险。

附注:

[1] 东吴金工成员陈实亦为本报告重要贡献者

[2] Yakov Amihud. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets, 2002(5):31-56.

[3] Yakov Amihud, Haim Mndelson. Asset pricing and the bid-ask spread. Journal of Financial Economics, 1986(17):223-249.

[4] 周明, 于渤. 信息传播的日内特征与释放过程——基于深圳股市的实证研究. 《证券市场导报》, 2008(6): 16-21.

[5] 攀登, 施东晖. 知情交易概率的测度模型及其影响因素分析. 《管理世界》, 2006(6): 18-26.

## 免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司（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。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司不对任何人因使用本报告中的内容所导致的损失负任何责任。在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。如引用、刊发、转载，需征得东吴证券研究所同意，并注明出处为东吴证券研究所，且不得对本报告进行有悖原意的引用、删节和修改。

## 东吴证券投资评级标准：

### 公司投资评级：

买入：预期未来 6 个月个股涨跌幅相对大盘在 15% 以上；

增持：预期未来 6 个月个股涨跌幅相对大盘介于 5% 与 15% 之间；

中性：预期未来 6 个月个股涨跌幅相对大盘介于-5% 与 5% 之间；

减持：预期未来 6 个月个股涨跌幅相对大盘介于-15% 与-5% 之间；

卖出：预期未来 6 个月个股涨跌幅相对大盘在-15% 以下。

### 行业投资评级：

增持：预期未来 6 个月内，行业指数相对强于大盘 5% 以上；

中性：预期未来 6 个月内，行业指数相对大盘-5% 与 5%；

减持：预期未来 6 个月内，行业指数相对弱于大盘 5% 以上。

东吴证券研究所

苏州工业园区星阳街 5 号

邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>