



信息提纯，寻找高质量反转因子

——市场微观结构剖析之八

报告日期：2020-07-23

分析师：严佳炜

执业证书号：S0010520070001

邮箱：yanjw@hazq.com

分析师：朱定豪

执业证书号：S0010520060003

邮箱：zhudh@hazq.com

主要观点：

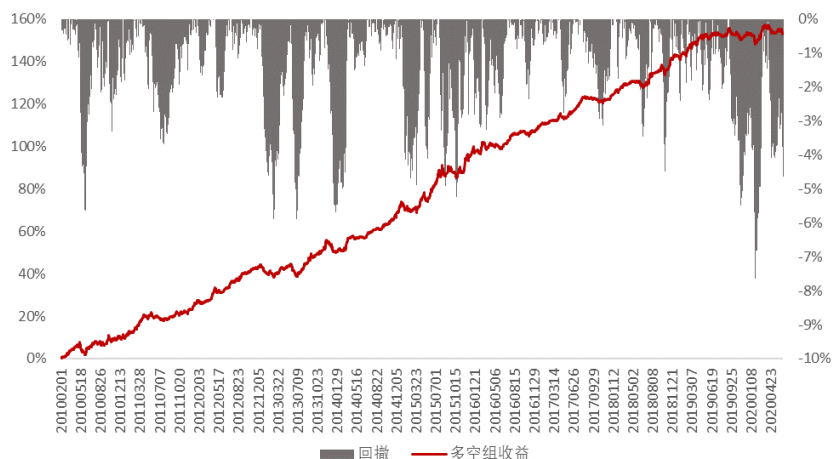
传统基于高频量价数据的因子，比如日内波动率、偏度、量价相关性等，往往将日内的价、量数据作为离散的点来进行使用，舍弃了价、量数据的时间属性。本篇报告中，我们另辟蹊径，将日内高频价格数据视为连续的时间序列，在此基础上进行统计分析，识别个股的日内趋势，以及叠加在趋势信号上的噪声。

具体的，我们运用经验模态分解模型（EMD），剥离股票高频价格的信号与噪音，构建信噪比因子（SNR）。信噪比因子本身即表现出较好的选股能力，RankIC 均值为 0.045，年化 ICIR 为 2.265，多空组的年化收益为 15.40%，年化波动为 6.90%，年化夏普为 2.233，月度胜率约 74%，最大回撤在 7.64%左右。具体而言，趋势性强，受到噪声扰动少的股票，下个月的表现更好。

进一步的，考虑到信噪比因子有着“信息提纯”的特性，是剥离噪声后的股价真实走势，因此可将其与反转因子进行结合。我们发现，信息提纯后，高信号强度的股票表现出更强的反转效果，因此构建了高质量反转因子组合。高质量反转组合相对原始反转在 IC 表现，多空组夏普以及多头端收益都有显著提升，RankIC 从-0.068 提升至-0.074，年化 ICIR 从-2.270 提升至-2.727，多头年化收益从 7.83%提升到 11.90%。

相关报告

1.《高频视角下成交额蕴藏的 Alpha——市场微观结构剖析之七》2020-06-10



资料来源：Wind，华安证券研究所

风险提示

本报告基于历史个股高频数据进行测试，历史回测结果不代表未来收益。未来市场风格可能切换，微观交易结构可能发生变化，Alpha 因子可能失效，本文内容仅供参考。

正文目录

1 引言.....	5
2 信噪比的计算.....	5
2.1 经验模态分解 (EMD)	6
2.2 信噪比 (SNR) 因子的特征与构建	8
2.3 信噪比 (SNR) 因子的表现	9
2.4 参数敏感性测试	11
2.5 新 SNR 因子的合成.....	13
2.6 信号与噪声的时间尺度	14
3 因子的深入剖析与应用	16
3.1 动量/反转因子的困境	16
3.2 信噪比与反转双分组解析	17
3.3 改进反转因子：信噪比增强反转.....	19
4 总结.....	20
5 风险提示	21

图表目录

图表 1 价格序列 1	5
图表 2 价格序列 2	5
图表 3 信号处理方法介绍 (傅里叶变换、小波分解、EMD)	6
图表 4 EMD 分离过程	7
图表 5 EMD 第一次分离结果	7
图表 6 2019/12/31 某股票的股价信噪分离	7
图表 7 分解到不同层数时, 信噪比因子截面中位数的 20 日移动平均	8
图表 8 SNR_layer2 因子分布	9
图表 9 SNR_layer3 因子分布	9
图表 10 SNR_layer2 月度 IC 序列	9
图表 11 SNR_layer2 多空组收益与回撤	9
图表 12 SNR_layer2 分 10 组年化收益	10
图表 13 SNR_layer2 分 10 组净值	10
图表 14 SNR_layer3 月度 IC 序列	10
图表 15 SNR_layer3 多空组收益与回撤	10
图表 16 SNR_layer3 分 10 组年化收益	10
图表 17 SNR_layer3 分 10 组净值	10
图表 18 多头/空头组股票的特征	11
图表 19 运用 EMD 分离至不同层数时的信号与噪声序列	11
图表 20 EMD 分离 2 层信噪比 MA 参数对比	12
图表 21 EMD 分离 2 层信噪比 EMA 参数对比	12
图表 22 EMD 分离 3 层信噪比 MA 参数对比	12
图表 23 EMD 分离 3 层信噪比 EMA 参数对比	12
图表 24 SNR_layer2 时间序列特征	13
图表 25 SNR_layer3 时间序列特征	13
图表 26 新 SNR 月度 IC 序列	14
图表 27 新 SNR 多空组收益与回撤	14
图表 28 新 SNR 分 10 组年化收益	14
图表 29 新 SNR 分 10 组净值	14
图表 30 60 日信噪比因子多空收益与回撤	14
图表 31 60 日信噪比因子分组年化收益	14
图表 32 20 日分钟连续价格信噪比多空表现	15
图表 33 20 日分钟连续价格信噪比分组年化收益	15
图表 34 日间价格断层 (2019/12/31, 000001.SZ)	15
图表 35 20 日反转因子月度 IC	16
图表 36 20 日反转因子多空收益与回撤	16
图表 37 20 日反转因子多头收益趋 0 (2014 年后)	16
图表 38 20 日反转因子逐年超额收益	17
图表 39 REVERSAL 和 SNR 分组收益示意图	18
图表 40 双分组多头收益对比 (2014 年以来)	18
图表 41 多头收益对比 (2014 年以来)	18

图表 42 反转因子强化前后对比.....	19
图表 43 原始 20 日反转因子多空收益与回撤.....	19
图表 44 增强 20 日反转因子分组年化收益.....	19
图表 45 反转因子强化前后分组年化收益.....	20
图表 46 反转因子强化前后多头净值.....	20
图表 47 新SNR月度 IC 序列.....	20
图表 48 新SNR多空组收益与回撤.....	20
图表 49 新SNR分 10 组年化收益.....	20
图表 50 新SNR分 10 组净值.....	20
图表 51 反转因子改进前后多头对比.....	21

1 引言

本篇报告是“市场微观结构剖析”系列的第八篇，系列报告从行为金融学的视角出发，捕捉市场的异常价格震荡，挖掘隐藏在海量数据之下的 Alpha。本篇报告聚焦于分钟的价格序列，剥离时间序列中的信号与噪声，构建信噪比因子来度量高频交易数据中蕴藏的“信息含量”，研究发现，信噪比因子可以与反转因子有效结合。

最早提出“噪声交易者”(Noise Trader)概念的是 Kyle(1985), Black(1986) 的演讲指出噪声交易是金融市场存在的基础，同时也给金融市场带来了问题。噪声交易者指无法获得内部信息，非理性地把噪音当做信息进行交易的投资者。有效市场理论认为噪声交易者在市场的预期收益为负，当噪声交易者非理性地拉升或拉低了股票的价格，使其偏离其内在基本面价值时，理性投资者会反向套利，使价格回到内在价值水平，由于噪声交易者的经济效益为负，市场上的噪声交易者会渐渐消失。1990 年以来，随着行为金融学的兴起，噪声交易模型 (DSSW) 给出了噪声交易者长期存在于市场中的原因：1、噪声交易者获得的信息具有质量问题，可能是他们通过技术分析、做市商或是其他咨询机构获得的虚假信号，并以此作为投资的依据；2、理性投资者在价格偏离基本面时，试图通过反向交易策略获利，但由于噪声交易者行为具有随机性和不可预测性，降低了无风险套利者的套利意愿，难以使价格完全回归，产生了套利限制。

虽然噪声交易者在市场中是否能够长期存在和获利尚无定论，但目前的 A 股市场存在噪声的现象比其他成熟的资本市场更为严重，分析原因可能是：

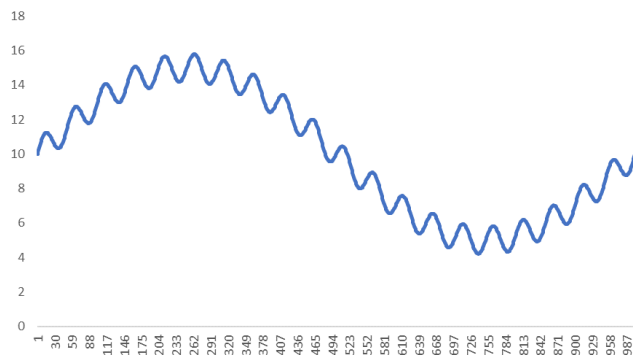
- 1、A 股个人投资者占比较高，存在许多的非理性交易，比如追涨杀跌、过度反应、反应不足等等，从上交所 2017 年给出的数据来看，个人投资者以约 20% 的持股市值贡献了超过 80% 的买卖净额。
- 2、投资者过度自信，将私人获得的信息作为交易的依据，使价格偏离内在股票价值。
- 3、流动性不足，交易对股价的流动性冲击尤其是小市值股票，使得股票价格更容易短暂地偏离价值。

本报告对日内股价序列进行信噪比的分析，集中挖掘潜在在分钟价格之下的趋势信息和噪声信息，并以此来构建信噪比因子，并且研讨信噪比因子的应用。

2 信噪比的计算

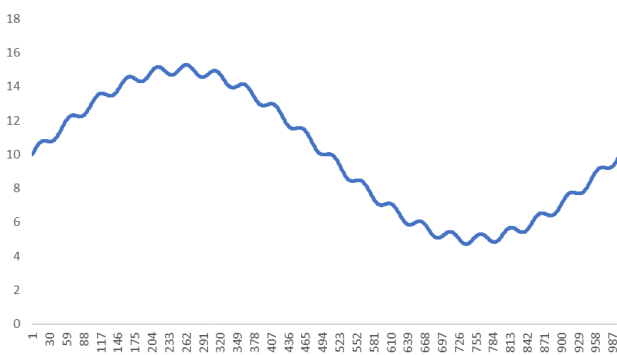
下图 1,2 模拟了两个简单价格序列，从整体趋势来看，价格都经历了一个先上涨，后下跌，最后又上涨的过程，但是序列 1 的短期价格震荡明显强于序列 2，序列 1 价格所包含的噪声占比高于序列 2。

图表 1 价格序列 1



资料来源：华安证券研究所整理

图表 2 价格序列 2



资料来源：华安证券研究所整理

在真实的金融价格序列中，虽然情况远远比此复杂，但整体上依然可以分为信号部分（趋势）和噪声部分（震荡）。金融序列的一个显著特点是非线性、非平稳，信号部分代表股票整体的价格趋势，是股价受到外部影响时呈现出的非平稳走势，噪声部分代表了股价的短期震荡，围绕着整体趋势上下波动，是股价自身进行短期调整的行为。

传统的日内统计指标，比如价格波动率，量价相关性，成交额占比等等，往往将日内的价、量数据作为离散的点来进行数学描述，舍弃了价、量数据的时间属性，例如，仅对价格序列计算波动率时，序列 1 的波动率为 3.58，序列 2 的波动率为 3.54，两者波动十分相近，难以分辨出序列 1 实质上噪声更多的特征。本文在剥离价格序列的信号与噪声时，基于日内价格整体在时间上的走势来分析，计算出价格序列的信噪比指标。

2.1 经验模态分解 (EMD)

经验模态分解 (Empirical Mode Decomposition, EMD) 是一种新型的、自适应的信号时频处理方法，特别适用于非线性非平稳的金融序列分析处理。该方法是依据数据自身的时间尺度特征来进行信号分解，无须预先设定任何小波基函数和尺度函数，能够根据局部特征自适应地进行信噪分离。

图表 3 信号处理方法介绍 (傅里叶变换、小波分解、EMD)

信号处理	分离方法	适用领域
傅里叶变换	将序列分解为直流分量和一组不同幅值、频率、相位的正弦波。	平稳信号，需要确定参数
小波变换	将序列分解到多尺度中,在每一尺度下把属于噪声的小波系数去除,保留并增强属于信号的小波系数,将无限长的三角函数基换成了有限长的会衰减的小波基。	非平稳信号，需要确定参数
经验模态分解	将序列分解成从高频到低频的若干个本征模函数和一个趋势函数。	非平稳信号，不需要参数

资料来源：华安证券研究所整理

经过经验模态分解处理，任意价格序列都是由若干个噪音函数（震荡部分）和一个信号函数（趋势部分）组成的。假设原始的价格序列为 $P(t)$ ，则

$$P(t) = \sum_{i=1}^n IMF_i(t) + r(t)$$

其中 $IMF_i(t)$ 被称为本征模函数，可以理解为价格序列分离出的各层数据震荡， $r(t)$ 是价格序列的整体趋势部分，代表价格在一天内的总体走势。我们可以将 $IMF_i(t)$ 近似理解为在不同时间维度下的波动，例如 10 分钟、5 分钟、2 分钟线上波动的叠加。通过自适应方法分离出的本征模函数 $IMF_i(t)$ 具有以下两个特征：函数在整个时间范围内，局部极值点和过零点的数目必须相等，或最多相差一个；在任意时刻点，局部最大值的包络（上包络线）和局部最小值的包络（下包络线）平均必须为零。

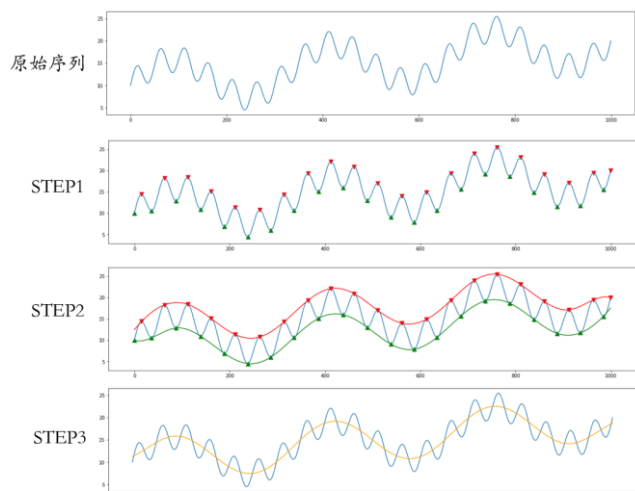
给出一个原始的价格序列 $P(t)$ ，EMD 分解算法可以按照以下步骤进行：

- 1、找出原始系列的局部极大值点和极小值点，如图表 4 的 STEP1，红色点是局部极大值点，绿色点是局部极小值点；
- 2、利用三次样条函数进行插值，得到极大值点组成的上包络线 (STEP2 中的红线) 和极小值点组成的下包络线 (STEP2 中的绿线)；
- 3、对上下包络线取平均，得到包络线的瞬时均值序列 $m(t)$ (STEP3 中的黄线)；
- 4、用原始的价格序列 $P(t)$ 减去均值序列 $m(t)$ 得到波动序列 $h(t)$ ， $h(t)$ 如果满足本

征模函数的特征，则为第一层本征模函数 $IMF_1(t)$ ，如果不满足，就将 $h(t)$ 作为序列按照 1~3 的步骤再次分解，直到分解出 $IMF_1(t)$ ；

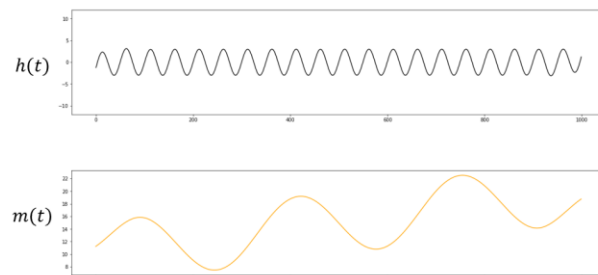
5、第一层分离趋势项 $r_1(t) = P(t) - IMF_1(t)$ ，接着将 $r_1(t)$ 作为新的原始序列，重复 1~4 的步骤，直到分解出 $IMF_2(t)$ 、 $IMF_3(t)$ 、 $\dots IMF_n(t)$ 和单调趋势项 $r_n(t)$ 。

图表 4 EMD 分离过程



资料来源：华安证券研究所整理

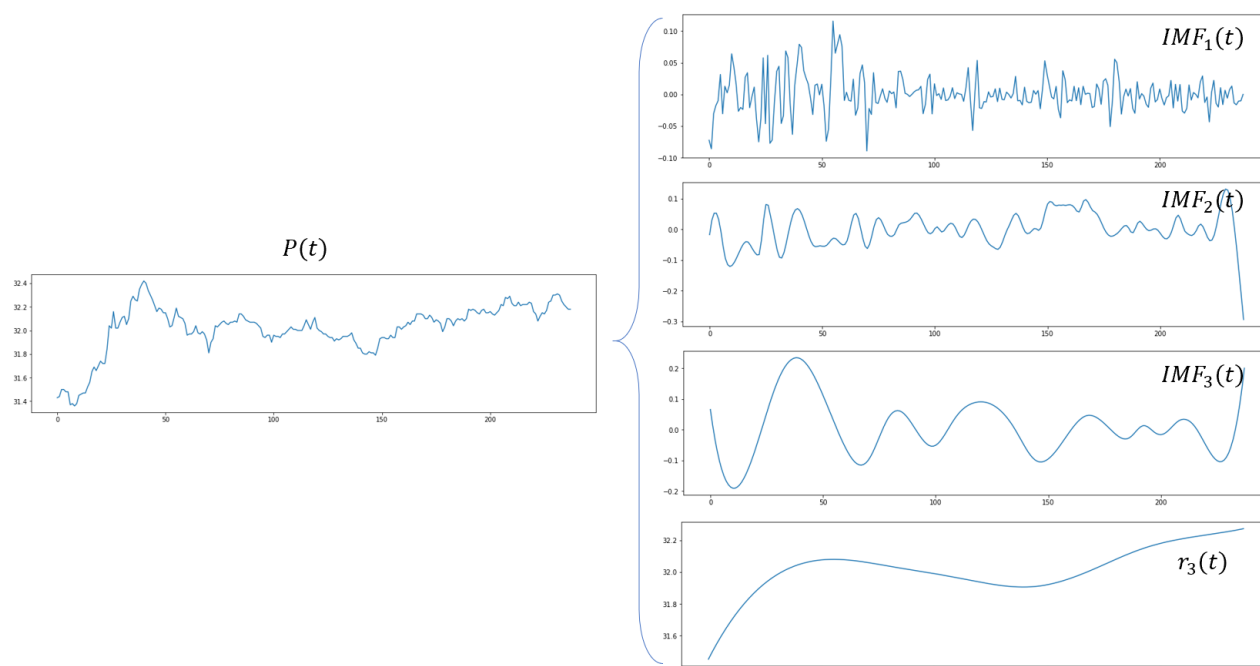
图表 5 EMD 第一次分离结果



资料来源：华安证券研究所整理

在实际操作中，为了计算简便，本文省去了判定波动序列 $h(t)$ 是否为本征模函数的过程，将每一次分离得到的 $h(t)$ 近似地作为 $IMF(t)$ ，对日内价格序列进行信噪分离时发现，当分解至二或三层时，信号序列平滑特性明显，能够接近地描述原始序列。

图表 6 2019/12/31 某股票的股价信噪分离



资料来源：wind，华安证券研究所

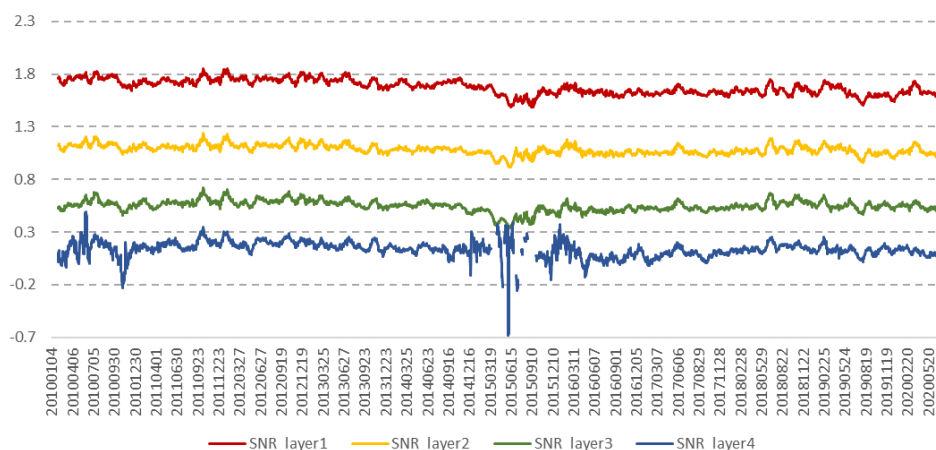
如图，是2019/12/31某股票的日内价格序列分离至三层时的效果，三层趋势项 $r_3(t)$ 基本反映了当日股价的整体走势，是价格序列的信号部分，而原始序列减去趋势序列，即所有IMF的和，为价格序列的噪声部分。将日内信噪比（Signal to Noise Ratio, SNR）定义为信号序列的标准差与噪声序列标准差的比值的对数。

$$SNR = \log \left(\frac{\text{std}(r_n(t))}{\text{std}(P(t) - r_n(t))} \right)$$

2.2 信噪比（SNR）因子的特征与构建

运用上一节介绍的EMD信噪分离方法，可以计算出每只股票每日价格的信噪比，每多分离一层，价格序列中识别为噪声的部分就越多，信号序列就越平滑。以 SNR_layer_k 代表信噪分离了k层，绘制全市场信噪比的中位数序列。所有股票日内价格序列都可以进行一层和二层的信噪分离，约99.65%的样本可以分离至三层，约44.50%的样本可以分离至四层，分离的次数越多，噪声部分的波动越大，信噪比因子越小，如图7所示，1~3层信噪比在时间序列上分化明显，4层信噪比开始失真，在2015年中的极端行情下，绝大多数股票表现为急涨急跌时，分离四层无法实现。

图7 分解到不同层数时，信噪比因子截面中位数的20日移动平均



资料来源：wind，华安证券研究所

本文中，将日内价格序列分离至第二或三层时，趋势表现出剥离噪声的特性且与原股价保持高相关。我们首先测试是否能将信噪比系数直接作为选股因子，描述股票价格波动中趋势、噪声含量的高低。月度因子值以月底信噪比因子值做指数衰减加权平均合成。（参数调节见第3节）

$$factor_month = \sum_{i=1}^d \frac{\theta^i}{\sum_{j=1}^d \theta^j} \times factor_day_j$$

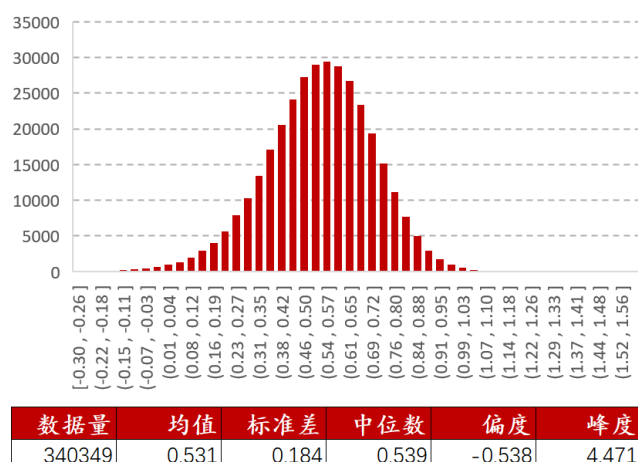
$$\text{衰减系数} \theta = 1 - \frac{2}{1+d}$$

图表8 SNR_layer2因子分布



资料来源: wind, 华安证券研究所

图表9 SNR_layer3因子分布



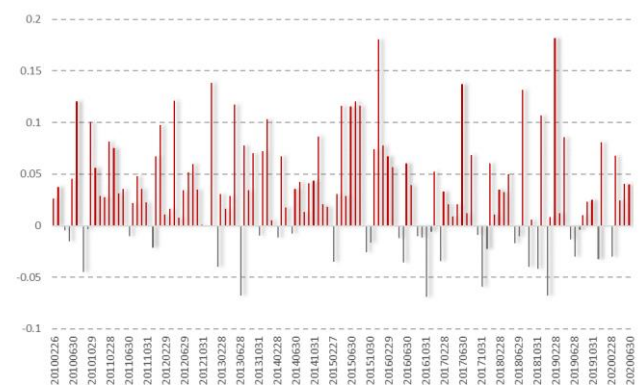
资料来源: wind, 华安证券研究所

观察因子值的直方图, SNR_layer2、SNR_layer3 因子的均值分别在在 1.064、0.531 左右, 因子分布都呈现略微的左偏以及尖峰的特征, 左尾说明信噪比低的股票从样本中分化了出来, 即因子较好地识别出了价格嘈杂度高的股票。另外, 由于 SNR_layer3 相较于 SNR_layer2 剥离了更多的噪声, 因此信噪比值相对更小也是合理的。

2.3 信噪比 (SNR) 因子的表现

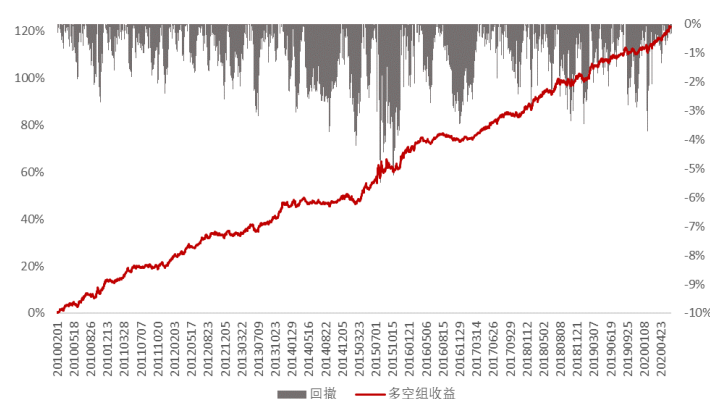
行业市值中性化后, SNR_layer1 因子没有显著的选股效果, SNR_layer2 因子的月度 rankIC 均值为 0.032, 年化 ICIR 为 2.145, IC 值为正的概率为 72%。按照因子值从小到大依次分成十组, 构建等权选股组合, 统计各组收益, 1~10 组的年化收益呈现单调变化。多空组的年化收益为 11.60%, 年化波动为 6.41%, 年化夏普为 1.80, 最大回撤在 5.48%左右。因子与 20 日反转、20 日均换手、20 日波动等传统因子基本无相关性, SNR_layer2 因子贡献的 Alpha 是全新的。

图表10 SNR_layer2月度 IC 序列



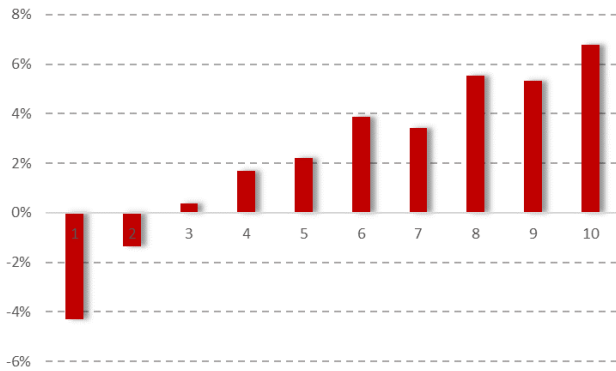
资料来源: wind, 华安证券研究所

图表11 SNR_layer2多空组收益与回撤



资料来源: wind, 华安证券研究所

图表 12 SNR_layer2分 10 组年化收益



资料来源: wind, 华安证券研究所

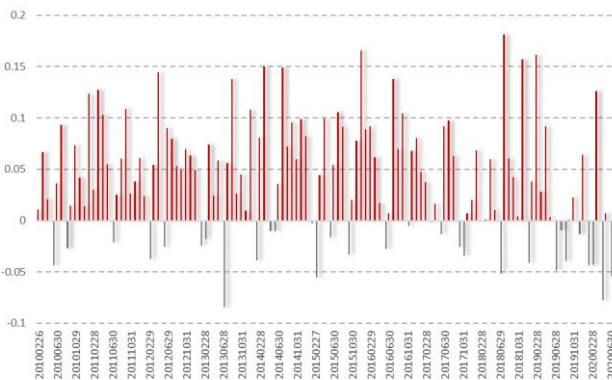
图表 13 SNR_layer2分 10 组净值



资料来源: wind, 华安证券研究所

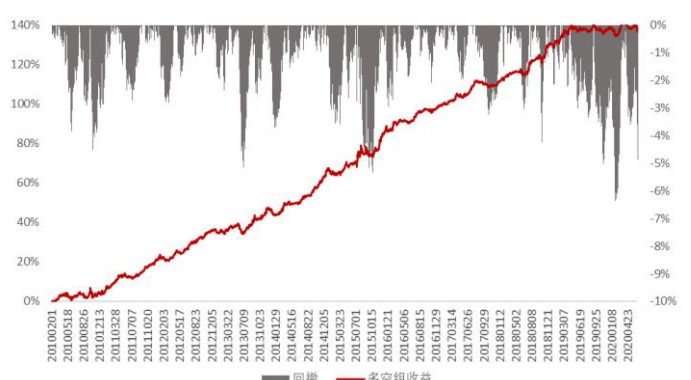
行业市值中性化后, SNR_layer3 因子的月度 rankIC 均值为 0.040, 年化 ICIR 为 2.465, IC 值为正的概率为 74%。多空组的年化收益为 13.76%, 年化波动为 6.16%, 年化夏普为 2.233, 最大回撤在 6.38%左右。整体表现略微优于 SNR_layer2, 但 2019 年后半年多空曲线趋平, SNR_layer3 与传统因子 20 日均换手、20 日波动的相关性在-0.2 左右。主要是由于分离的层数越多, 噪声序列与原始序列的相关性越大, 因而噪声序列的标准差会越来越靠近原始序列的标准差。因此, 从分析增量 Alpha 的角度来说 SNR_layer2 更佳。

图表 14 SNR_layer3月度 IC 序列



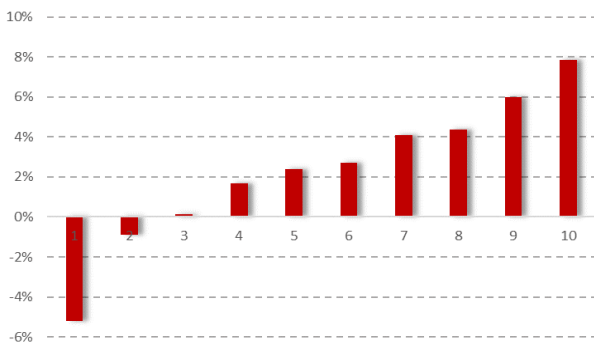
资料来源: wind, 华安证券研究所

图表 15 SNR_layer3多空组收益与回撤



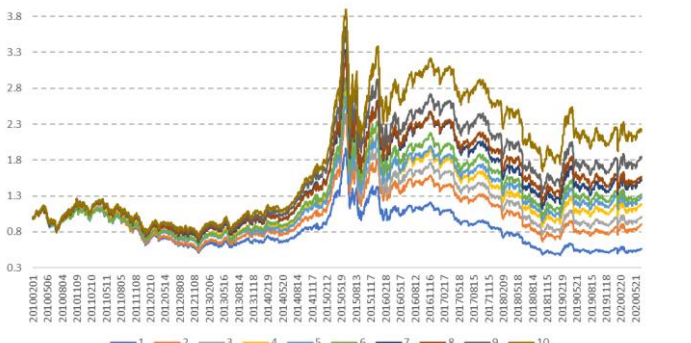
资料来源: wind, 华安证券研究所

图表 16 SNR_layer3分 10 组年化收益



资料来源: wind, 华安证券研究所

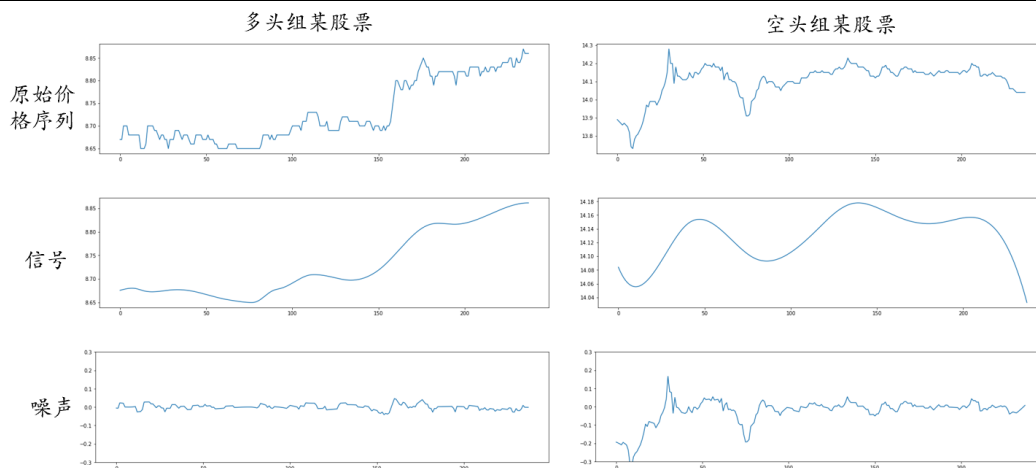
图表 17 SNR_layer3分 10 组净值



资料来源: wind, 华安证券研究所

从IC符号,分组收益来看,该因子为正向因子,也即:股票的信噪比越大,股票的下个月收益越高,反之,信噪比越小,股票的下个月收益越低。从行为金融学的角度出发,某只股票存在的虚假信息或是噪声交易者越多,造成价格偏离内在价值越频繁或是偏离度越高,该股的下月表现越差。

图表 18 多头/空头组股票的特征



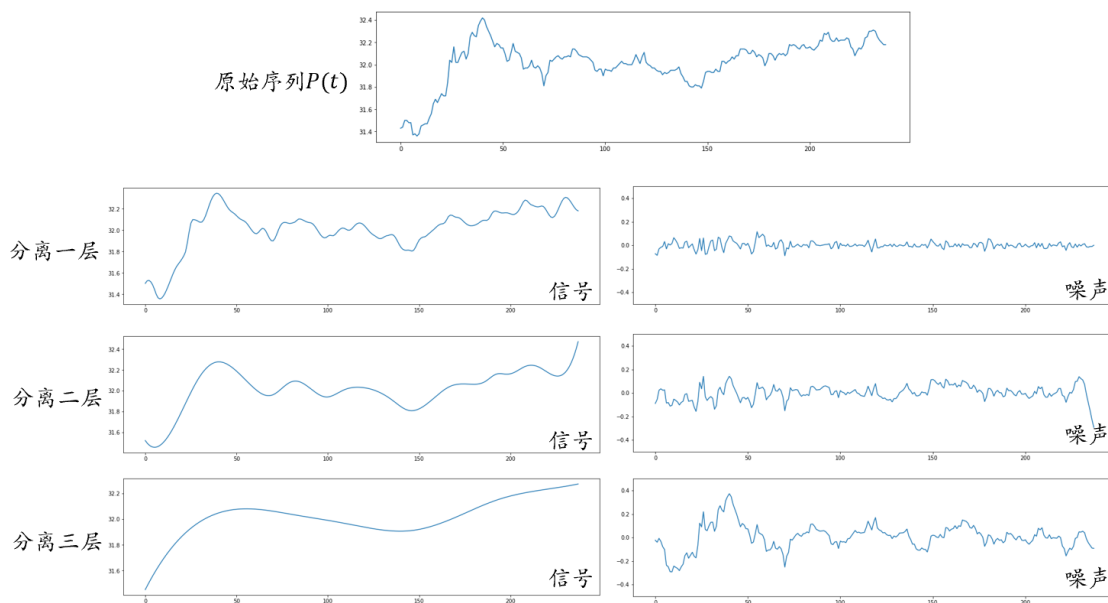
资料来源: wind, 华安证券研究所

分别从多头组/空头组抽取了两只股票,信噪比分别为 1.57 和-0.70,绘制原始价格序列与信号噪声序列如图 13 所示,多头组股票的显著特征是噪声震荡非常小,而空头组股票的噪声振幅显著大一些。

2.4 参数敏感性测试

在对价格序列分离信噪时,选择适当的分离层数对因子的构建事关重要,依然以 2019/12/31 某股票的日内价格序列为例,图展示了分离至不同层数时,信号序列以及噪声序列的走势,随着分离层数的变多,股票的信号部分的波动越来越低,走势越来越平滑,而噪声部分的波幅渐渐变大,每分离一次就有更多的波动被识别为噪声。

图表 19 运用 EMD 分离至不同层数时的信号与噪声序列



资料来源: wind, 华安证券研究所

经过统计分析，市场上大部分的股票在分离一层时噪声含量仍然很高，至二或三层时能够较好展现日内趋势，因此 SNR_layer2 和 SNR_layer3 选股效果较好，继续往下分离时，由于股票价格极值点的限制，超过一半的股票样本无法分离至四层。

我们从三个维度进行参数敏感性测试：1. 信噪比因子分离层数（2/3 层）；2. 构建月度因子时，以月末为基日，向前回溯天数（5~20 日）；3. 加权方式（一般移动平均（MA）和指数加权移动平均（EMA））。考察不同参数组合下，因子的表现。

图表 20 EMD 分离 2 层信噪比 MA 参数对比

SNR_layer2	多空组				
MA 天数	rankIC	年化ICIR	年化收益	最大回撤	年化IR
5	0.025	1.952	9.46%	6.41%	1.719
6	0.027	2.044	9.26%	6.16%	1.574
7	0.027	2.176	9.94%	5.93%	1.692
8	0.027	2.067	9.86%	5.74%	1.609
9	0.028	2.057	10.32%	4.58%	1.706
10	0.027	1.921	10.70%	4.54%	1.728
11	0.029	2.011	11.57%	5.29%	1.822
12	0.031	2.114	11.18%	4.83%	1.742
13	0.032	2.179	11.26%	5.40%	1.754
14	0.032	2.206	10.72%	5.42%	1.668
15	0.032	2.209	10.55%	5.64%	1.620
16	0.031	2.162	10.34%	4.55%	1.607
17	0.030	2.039	9.99%	5.55%	1.529
18	0.029	2.024	9.27%	5.48%	1.415
19	0.029	1.985	10.43%	6.15%	1.568
20	0.028	1.950	9.91%	5.74%	1.497

资料来源：wind，华安证券研究所

图表 21 EMD 分离 2 层信噪比 EMA 参数对比

SNR_layer2	多空组				
EMA 天数	rankIC	年化ICIR	年化收益	最大回撤	年化IR
5	0.024	1.688	9.51%	6.48%	1.778
6	0.025	1.790	10.80%	6.13%	1.967
7	0.026	1.893	11.15%	6.03%	1.966
8	0.027	1.930	10.33%	7.01%	1.773
9	0.028	1.974	10.71%	5.81%	1.792
10	0.028	1.979	10.44%	5.94%	1.702
11	0.029	2.022	11.54%	6.01%	1.864
12	0.030	2.074	11.79%	5.98%	1.883
13	0.031	2.124	11.50%	5.66%	1.831
14	0.031	2.154	11.17%	5.42%	1.768
15	0.032	2.180	11.60%	5.48%	1.808
16	0.032	2.185	11.65%	5.35%	1.814
17	0.032	2.164	11.24%	5.55%	1.741
18	0.032	2.166	11.18%	5.35%	1.733
19	0.032	2.164	11.27%	5.18%	1.728
20	0.032	2.154	10.91%	5.32%	1.662

资料来源：wind，华安证券研究所

图表 22 EMD 分离 3 层信噪比 MA 参数对比

SNR_layer3	多空组				
MA 天数	rankIC	年化ICIR	年化收益	最大回撤	年化IR
5	0.032	2.407	11.75%	5.07%	2.275
6	0.035	2.419	11.47%	6.40%	2.136
7	0.035	2.388	13.22%	5.80%	2.363
8	0.036	2.443	11.29%	6.75%	1.979
9	0.037	2.444	12.66%	7.03%	2.157
10	0.036	2.311	12.35%	6.60%	2.074
11	0.039	2.408	12.35%	6.83%	2.035
12	0.040	2.446	12.92%	7.96%	2.082
13	0.041	2.429	13.62%	6.65%	2.149
14	0.041	2.433	12.44%	7.27%	2.006
15	0.041	2.359	12.54%	6.37%	1.994
16	0.040	2.294	12.80%	6.39%	2.003
17	0.039	2.239	12.01%	6.89%	1.882
18	0.039	2.214	11.54%	7.02%	1.810
19	0.039	2.189	12.23%	8.56%	1.888
20	0.039	2.163	11.57%	8.46%	1.745

资料来源：wind，华安证券研究所

图表 23 EMD 分离 3 层信噪比 EMA 参数对比

SNR_layer3	多空组				
EMA 天数	rankIC	年化ICIR	年化收益	最大回撤	年化IR
5	0.029	2.157	10.02%	6.93%	1.988
6	0.031	2.251	11.63%	6.78%	2.241
7	0.033	2.310	12.56%	6.74%	2.333
8	0.034	2.374	12.88%	6.36%	2.320
9	0.036	2.408	13.20%	5.70%	2.295
10	0.036	2.398	12.45%	6.53%	2.134
11	0.038	2.441	13.49%	6.86%	2.287
12	0.039	2.458	13.71%	7.20%	2.288
13	0.040	2.471	13.85%	7.21%	2.300
14	0.040	2.477	14.30%	7.12%	2.352
15	0.041	2.465	13.76%	6.38%	2.233
16	0.041	2.453	13.63%	6.85%	2.194
17	0.041	2.433	13.67%	6.56%	2.175
18	0.041	2.423	14.00%	7.23%	2.209
19	0.042	2.415	13.20%	7.23%	2.059
20	0.042	2.402	13.06%	7.16%	2.012

资料来源：wind，华安证券研究所

从天数来看，因子的计算以 15~20 天最佳，加权方式是指数加权移动平均更好，越靠近月末的日期价格表现，对下月的影响更大，这与前几个系列报告中提到的其他日内高频因子的表现十分类似。

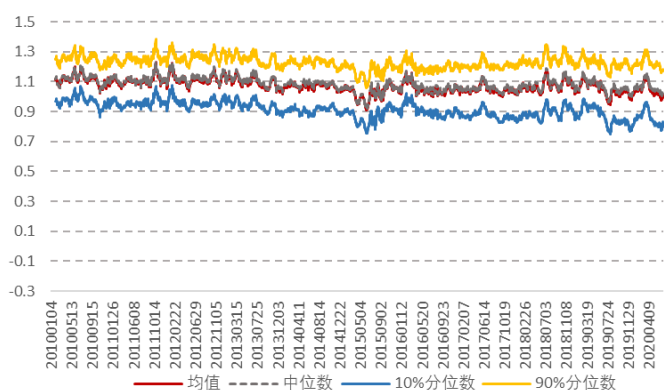
在计算合适的信噪分离层数时，我们也考虑了对于不同波动状况的股票，是否可以采取不同的分离层数来识别趋势与噪声，主要有以下方向改进：1、利用股票的不同波幅或是流动性状况给予合适的分离层数；2、利用分离出的信号序列与原始序列的相关性或是噪声序列与原始序列的相关性来判定合适层数；3、EMD 方法也存在一些技

术手段上的问题，比如模态混合问题，起点和终点的极值判定出现误差问题，可以采取更加精细的分离方法，比如集合经验模态分解（EEMD）、互补集合经验模态分解（CEEMD）等。

2.5 新 SNR 因子的合成

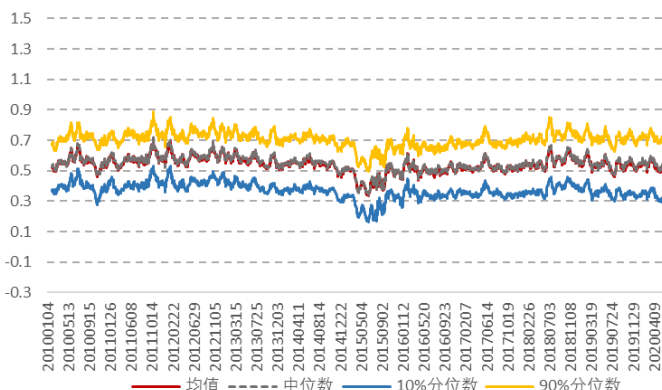
虽然信噪分离层数对于股票来说至关重要，但不同层数信噪比差距极大，如图 24、25 所示，两层信噪比在历史上的均值大约在 1.1 左右，三层信噪比在历史上的均值大约在 0.5 左右。如果对不同的股票，有的使用二层信噪比，有的使用三层信噪比作为因子值，在横截面上对股票分组时，按照因子值大小排序的结果是不可信的，更多地受到人为分离层数的干扰，调整效果是不佳的。

图表 24 SNR_layer2 时间序列特征



资料来源：wind，华安证券研究所

图表 25 SNR_layer3 时间序列特征



资料来源：wind，华安证券研究所

我们推测“合适的”信噪分离层数在二至三层之间，运用 SNR_layer2 和 SNR_layer3 合成新 SNR 因子，当股票日内价格波动较大时，剥离更多的噪声，因子值向 SNR_layer3 偏离，赋予 SNR_layer3 更大的权重，波动较小时，因子值赋予 SNR_layer2 更大的权重。首先在横截面上对股票日内波动率进行极值处理，0-1 标准化，其次根据标准化波动率缩至合适的权重区间（当 δ 等于 1 时，权重即为波动率分位数， δ 趋向无穷时，为因子等权合成），最后对二层、三层信噪比因子加权合成每只股票的新 SNR 因子。

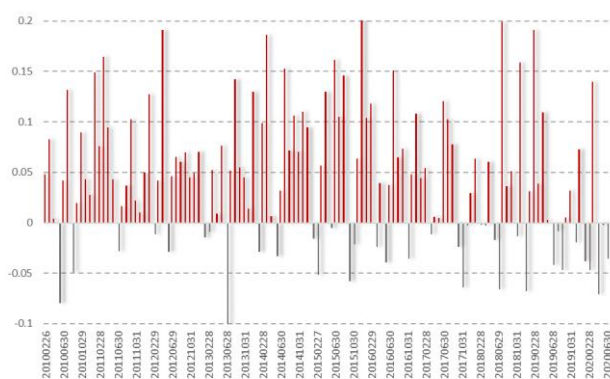
$$Vol_{std} = \frac{Vol - \min(Vol)}{\max(Vol) - \min(Vol)}$$

$$Weight_{Vol} = 0.5 + \frac{Vol_{std} - 0.5}{\delta}$$

$$新SNR = Weight_{Vol} \times SNR_{layer2} + (1 - Weight_{Vol}) \times SNR_{layer3}$$

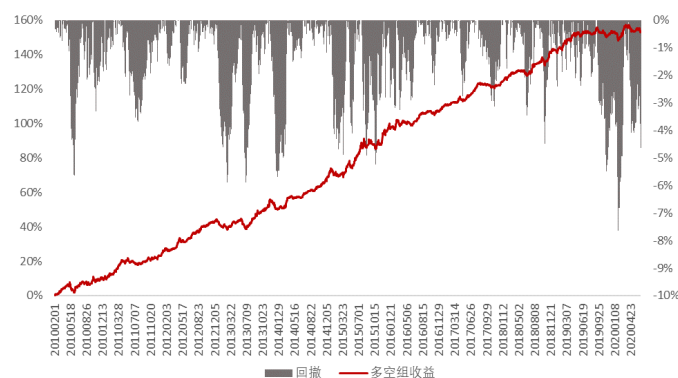
新 SNR 因子（ $\delta = 5$ 时）的月度 rankIC 均值为 0.045，年化 ICIR 为 2.265，IC 值为正的概率为 74%。多空组的年化收益为 15.40%，年化波动为 6.90%，年化夏普为 2.233，最大回撤在 7.64% 左右，对原始的 SNR_layer2 和 SNR_layer3 有一定程度的改进。

图表 26 新SNR月度IC序列



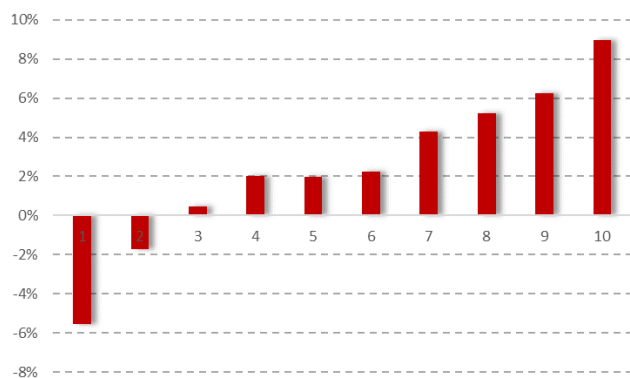
资料来源: wind, 华安证券研究所

图表 27 新SNR多空组收益与回撤



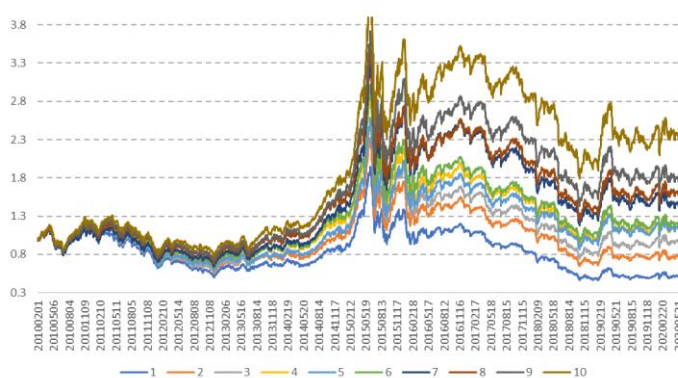
资料来源: wind, 华安证券研究所

图表 28 新SNR分 10 组年化收益



资料来源: wind, 华安证券研究所

图表 29 新SNR分 10 组净值

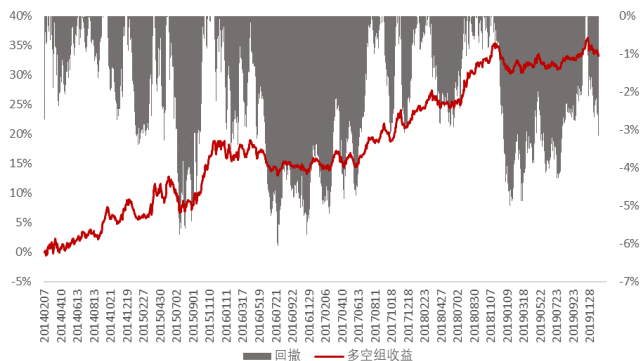


资料来源: wind, 华安证券研究所

2.6 信号与噪声的时间尺度

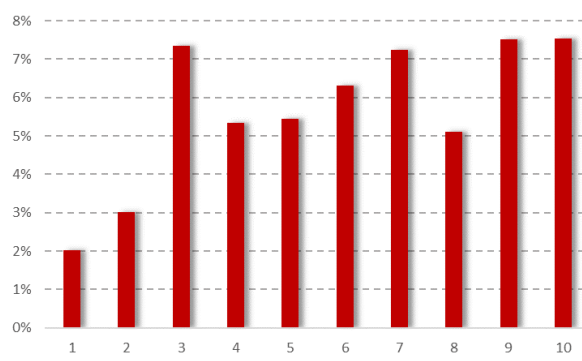
前文中, 我们对于信号与噪声的考虑是基于日内分钟频度上的, 根据道氏理论, 实际上价格的趋势和波动是基于不同时间尺度来识别的, 主要分为长期趋势 (持续一年以上)、中期趋势 (持续三个星期以上) 以及短期趋势 (持续不超过 6 天)。如果从长期趋势来看, 则日内的所有价格波动都被看作是噪音。

图表 30 60 日信噪比因子多空收益与回撤



资料来源: wind, 华安证券研究所

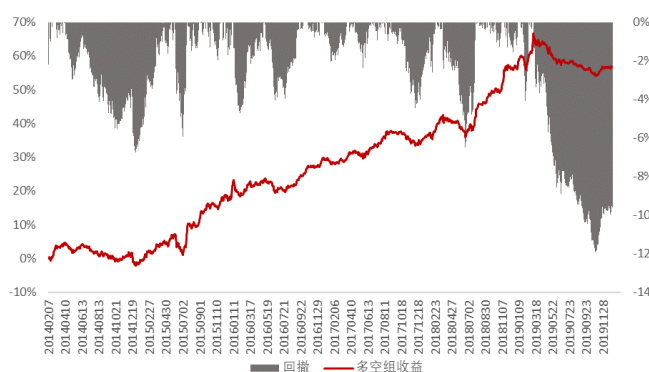
图表 31 60 日信噪比因子分组年化收益



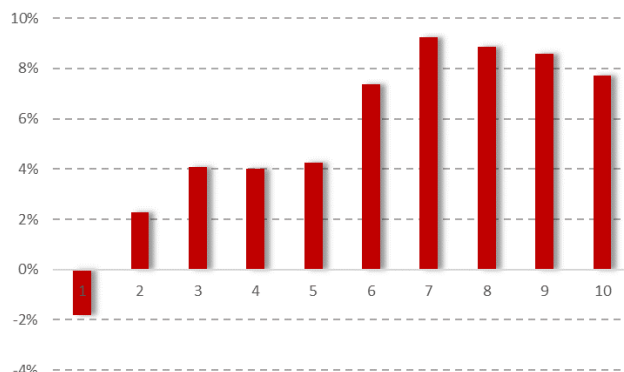
资料来源: wind, 华安证券研究所

我们以 20 日、40 日、60 日、120 日的日度价格序列分别构建日频价格下的信噪比因子。效果都十分有限，20 日、40 日的价格序列样本点较少，能分离出信噪比的股票覆盖率太低，图展示了 60 日度价格一层信噪比的效果，效果较差，主要是由于样本点较少，能够进行信噪分析的操作空间有限，这也侧面说明了高频分钟下的价格蕴藏了日度价格无法体现的独特交易行为，价格震荡在更加高频的计算方式下无所遁形。

此外，我们还尝试了使用月度 20 个交易日的分钟价格数据连接起来，形成跨日的连续分钟序列，包含 4800 个分钟样本点，考察信噪比因子的表现。图 32、33 展示了多空组收益，和分 10 组年化收益。

图表 32 20 日分钟连续价格信噪比多空表现


资料来源：wind，华安证券研究所

图表 33 20 日分钟连续价格信噪比分组年化收益


资料来源：wind，华安证券研究所

我们发现价格连结序列信噪比因子的表现不尽人意，IC 均值 0.025，rankIC 均值 0.0318，年化 ICIR 为 1.48，多空组的年化收益为 9.70%，年化波动为 6.27%，年化夏普为 1.55，月度胜率约 62%，最大回撤在 11.96%左右。

图表 34 日间价格断层 (2019/12/31,000001.SZ)


资料来源：wind，华安证券研究所

原因由两方面带来：1、日度连结序列将月初和月末日期的价格序列等权地放入了信噪比计算中，而月初的股票表现对下月的影响极其有限；2、日间价格之间存在断层即跳价现象，这是隔夜信息的堆叠在开盘价的反应体现，技术方法容易将这类高震荡识别为噪音，导致信噪比的计算出现偏差。

3 因子的深入剖析与应用

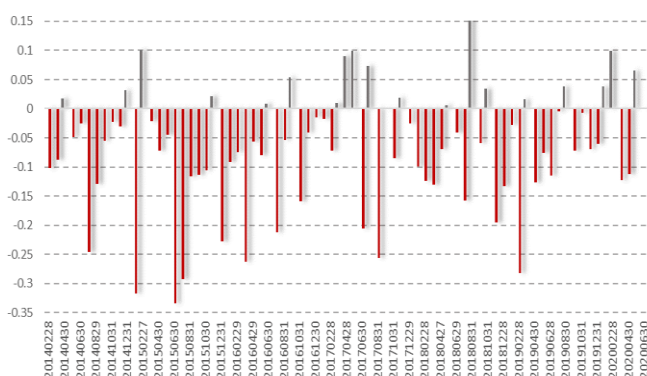
3.1 动量/反转因子的困境

传统的反转因子近年来遭遇了多头端失效的问题，以 20 日反转因子为例。

$$Reversal_{20} = \frac{P_t}{P_{t-20}} - 1$$

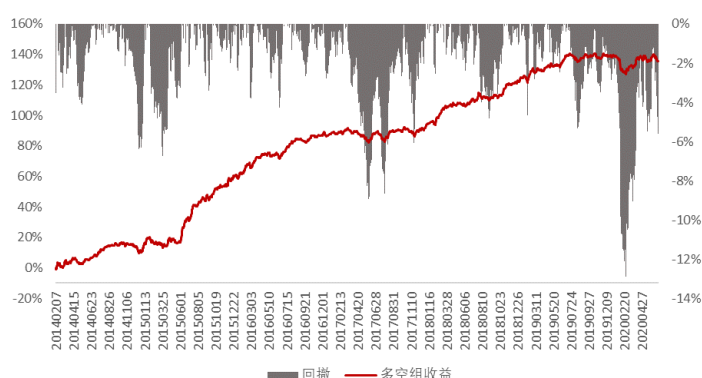
在 2014 年至 2020 年的区间内，因子的 IC 为-0.055，rankIC 为-0.068，年化 ICIR 为-1.91，多空组的年化收益为 22.77%，年化波动为 9.79%，年化夏普为 2.32，月度胜率约 702%，最大回撤在 12.91%左右，仅从 IC 来看是一个选股效果显著的因子。

图表 35 20 日反转因子月度 IC



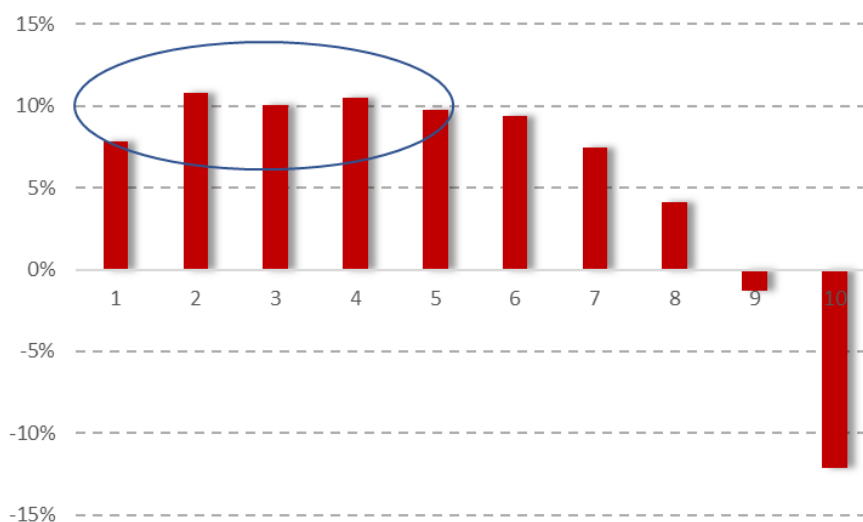
资料来源：wind，华安证券研究所

图表 36 20 日反转因子多空收益与回撤



资料来源：wind，华安证券研究所

图表 37 20 日反转因子多头收益趋 0 (2014 年后)



资料来源：wind，华安证券研究所

但从分组效果来看，多头端并没有表现出超额收益（图中蓝色圆圈），空头端的分组表现出了良好的线性关系。如图表 38 所示，反转因子 2013 年前尚且有多头收益，2014 年后表现极其不稳定，近年来多头收益已经渐渐消失。因子的选股效果基本上是由

空头贡献，即超涨的股票在下个月有负超额收益，对于超跌和涨跌幅不大的股票，反转因子难以预测。

图表 38 20 日反转因子逐年超额收益

反转因子	多头收益	10组收益均值	多头超额收益
2007年	209.11%	192.61%	16.50%
2008年	-56.87%	-57.60%	0.73%
2009年	155.18%	137.84%	17.34%
2010年	13.08%	9.70%	3.38%
2011年	-29.91%	-34.06%	4.14%
2012年	9.70%	-0.23%	9.93%
2013年	30.38%	23.92%	6.46%
2014年	32.25%	43.25%	-11.00%
2015年	116.34%	95.39%	20.96%
2016年	-9.36%	-10.92%	1.56%
2017年	-26.35%	-19.87%	-6.48%
2018年	-30.51%	-32.99%	2.47%
2019年	23.91%	21.96%	1.96%

资料来源：wind，华安证券研究所

3.2 信噪比与反转双分组解析

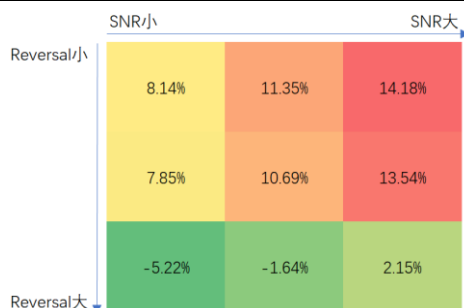
对传统反转因子的改进有许多方向，比如去除噪声较大的早盘，回溯过去 N 日的极大/极小值来构建反转因子等等，目的是去除市场中影响判断的噪声，得到更加有效的涨跌幅信息。

回看信噪比的计算方法，使用的是信号部分的波动率与噪声部分波动率的比值，对于日内涨幅/跌幅大的股票，分离出的信号部分的变化也会大一些，导致信噪比与涨跌幅绝对值的相关性有弱正相关，计算发现，分离三层的信噪比日度因子与当日涨跌幅绝对值的截面相关系数均值在 0.13 左右，即大涨大跌的股票在分离信噪比时，信号部分的波动天然地比其他股票更大，因此 SNR 因子值也会相对更高。

$$SNR = \log \left(\frac{std(r_n(t))}{std(P(t) - r_n(t))} \right)$$

股价的信号部分可分为向上涨的趋势与向下跌的趋势，如果价格的噪声振幅非常大，意味着多空方向的不明确，这种情况下提供的反转是质量较低的。我们将反转因子与信噪比因子结合分析，进一步理解股价拉升和骤降时，噪声的含量高低在其中扮演的角色。将所有股票先按照反转因子由低到高排序，分为 3 组，再在各个组中按照信噪比因子由低到高也分为 3 组，形成两个因子的 3×3 双排序组，计算各个分组的年化收益。很明显地看出反转因子多头端收益向信噪比高的股票倾斜，即反转因子值小，信噪比高的股票下个月的预期收益较高（右上角分组），主要原因在于噪声含量高的股票提供的反转其实是无意义的，信号强的股票反转效果更强。通过在反转因子中加入信噪比，无形之中是加入了对反转的路径的考察。

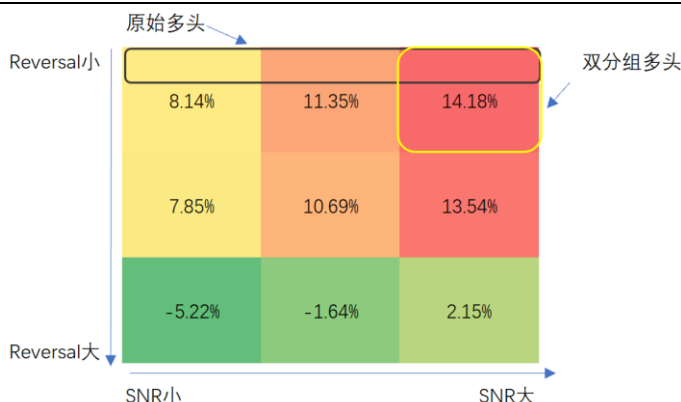
图表 39 Reversal 和 SNR 分组收益示意图



资料来源：wind，华安证券研究所

这个发现可以增强反转因子的多头端收益，如图，选择反转&信噪比双分组右上角3×3(图中黄色框出)作为新的多头组股票，和原始反转因子多头(图中黑色框出)作对比。新多头组，即反转因子小，信噪比大的股票分组年化收益约14.18%，净值走势如图41红色曲线所示，我们将动量因子分为同样9组，多头端年化收益约8.37%，净值走势如图41灰色曲线所示，经过信噪比增强的反转因子显著地跑赢了反转的原始多头。

图表 40 双分组多头收益对比 (2014 年以来)



资料来源：wind，华安证券研究所

图表 41 多头收益对比 (2014 年以来)



资料来源：wind，华安证券研究所

3.3 改进反转因子：信噪比增强反转

虽然日内高频的信噪比因子拥有良好的 Alpha 特性，但同时也面临着换手率较高导致的费率问题，因此作为基础因子来选股时费率较高。信噪比因子本质上是一个反应交易质量与噪声扰动度的因子，将其叠加在一些传统因子上，高频信息带来的增量 Alpha 可以修正传统因子失效的部分，使其重焕生机。

改进的基本思路是强化高信噪比股票的反转因子值，即对于高信号趋势、低噪声扰动的股票，对其 20 日反转因子值给予更高的信噪比乘数，因为日内买卖力量博弈程度高的股票，下一周期方向不确定的可能性依然很高，而日内信号含量高的股票，反转的可能更强。具体操作方法：在横截面上，对信噪比因子归一标准化作为信噪比乘数，然后对调整方向以及标准化后的反转因子加权。

$$Weight_i = \frac{SNR_i - \min(SNR)}{\max(SNR) - \min(SNR)}$$

$$\text{加强Reversal}_i = Weight_i \times \text{Reversal}_i$$

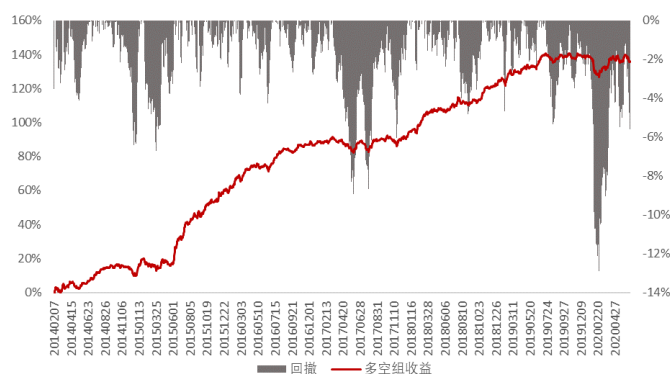
增强后的反转因子 rankIC 从-0.068 提升至-0.074，年化 ICIR 从-2.270 提升至-2.727，多空组的年化收益提升至为 25.01%，年化波动为 9.04%，年化夏普为 2.77，月度胜率约 75.3%，改进后分组的单调线性效果大大加强，多空组夏普和几次大幅回撤都有显著改善，多头年化收益从 7.83% 提升到 11.90%，净值走势对比如图 30 所示。

图表 42 反转因子强化前后对比

因子改进	rankIC	年化ICIR	多空组				多头年化收益	
			年化收益	年化波动	年化IR	胜率		最大回撤
原始反转	-0.068	-2.270	22.77%	9.80%	2.325	72.73%	12.92%	7.83%
增强反转	-0.074	-2.727	25.01%	9.04%	2.766	75.32%	13.00%	11.90%

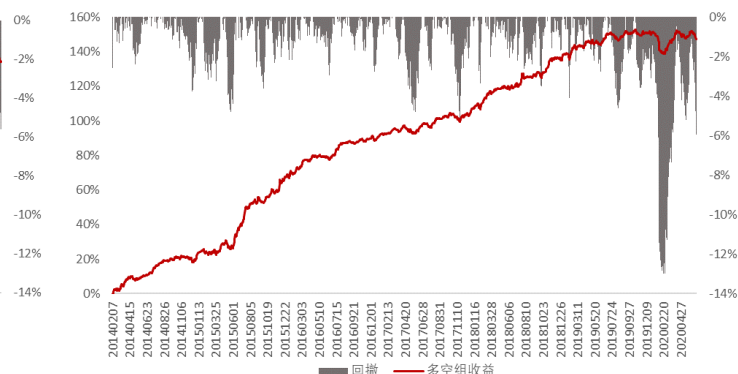
资料来源：wind，华安证券研究所

图表 43 原始 20 日反转因子多空收益与回撤



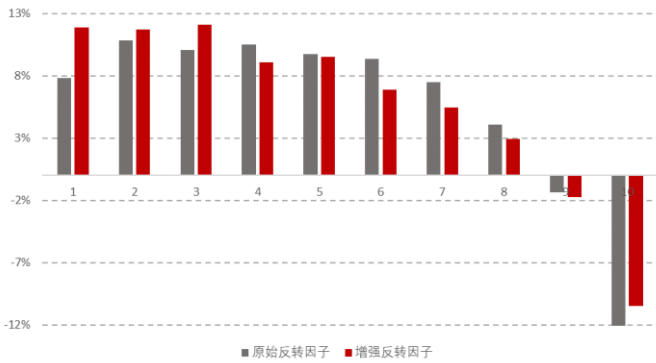
资料来源：wind，华安证券研究所

图表 44 增强 20 日反转因子分组年化收益



资料来源：wind，华安证券研究所

图表 45 反转因子强化前后分组年化收益



资料来源: wind, 华安证券研究所

图表 46 反转因子强化前后多头净值

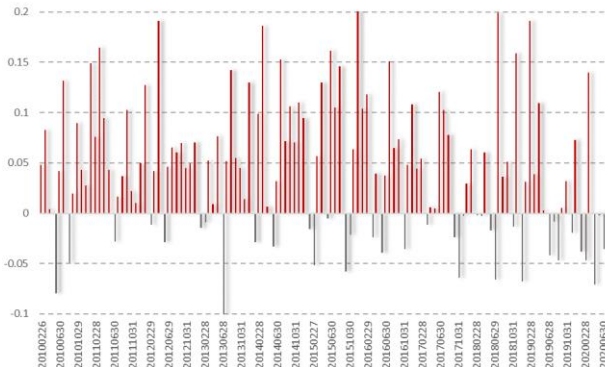


资料来源: wind, 华安证券研究所

4 总结

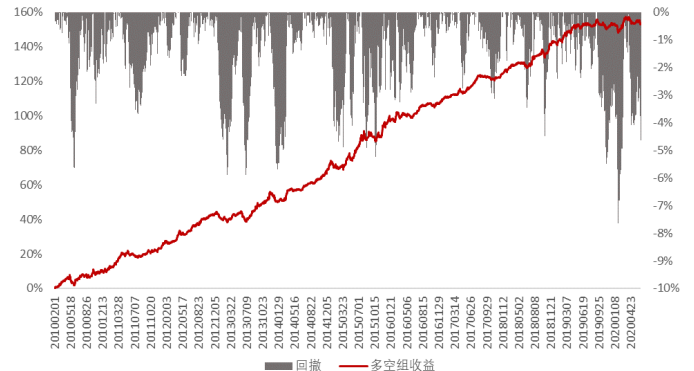
本文将日内的股价序列通过经验模态分解 (EMD), 分离成信号 (趋势) 和噪声 (震荡) 部分, 计算两者波动率的比值作为股票的信噪比因子, 测试其选股效果。合成信噪比因子选择 15 日左右的加权移动平均法合成, 月度 rankIC 均值为 0.045, 年化 ICIR 为 2.265, IC 值为正的概率为 74%。多空组的年化收益为 15.40%, 年化波动为 6.90%, 年化夏普为 2.233, 最大回撤在 7.64%左右, 表现出良好的 Alpha 特性。

图表 47 新SNR月度IC序列



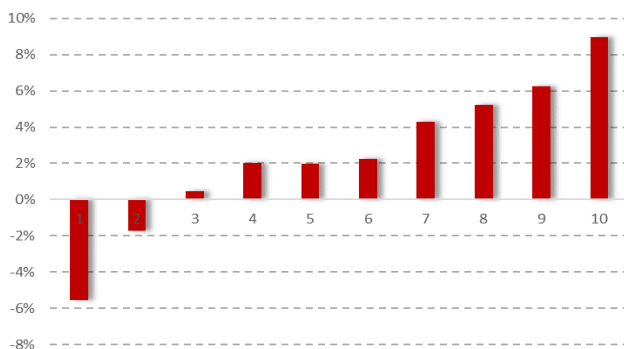
资料来源: wind, 华安证券研究所

图表 48 新SNR多空组收益与回撤



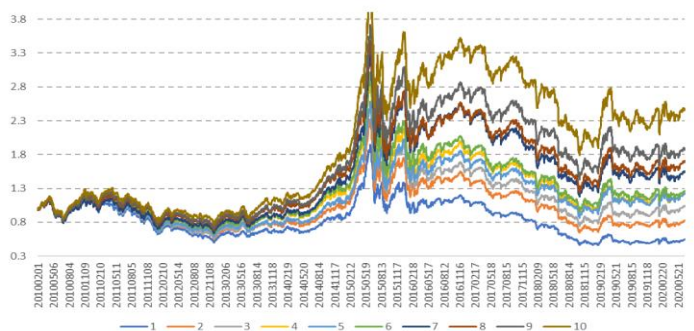
资料来源: wind, 华安证券研究所

图表 49 新SNR分 10 组年化收益



资料来源: wind, 华安证券研究所

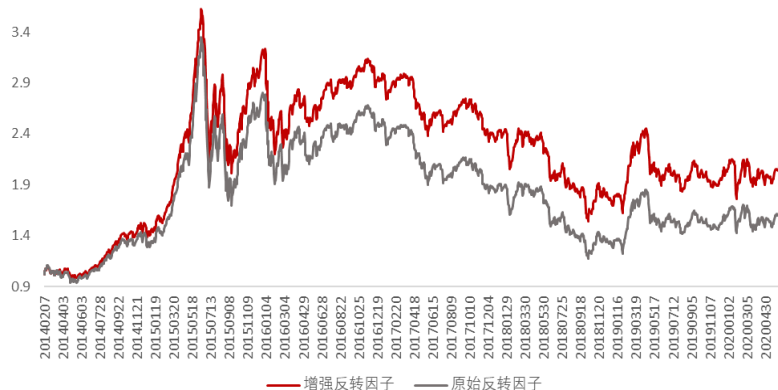
图表 50 新SNR分 10 组净值



资料来源: wind, 华安证券研究所

此外，进一步探索了信噪比因子在强化传统反转因子方向的应用，高频信息带来的增量 Alpha 可以修正传统因子多头失效的部分，改造后的反转因子稳定性上升，多头年化收益从 7.83% 提升到 11.90%，多空组的夏普显著改善。

图表 51 反转因子改进前后多头对比



资料来源：wind，华安证券研究所

回测说明：

- (1) 如果没有特别说明，本文回测一律采用市值、行业中性后的纯净因子。
- (2) 本文的回测时间段均为 2010/01/01~2020/06/30，改进反转因子时采用的回测时间段为 2014/01/01~2020/06/30，月度调仓时第一个月为因子值计算月，真实持仓日期为下个月月初开始。
- (3) 无特别说明时，手续费设为双边 3‰。
- (4) 在全部 A 股中，剔除涨停、跌停、停牌、日成交额低于 100 万、上市不满 180 天的新股等特殊股票。
- (5) 单因子等权组合测试中没有做组合优化和换手率控制，收益仍有提升空间。

5 风险提示

本报告基于历史个股高频数据进行测试，历史回测结果不代表未来收益。未来市场风格可能切换，微观交易结构可能发生变化，Alpha 因子可能失效，本文内容仅供参考。

重要声明

分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的执业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，本报告所采用的数据和信息均来自市场公开信息，本人对这些信息的准确性或完整性不做任何保证，也不保证所包含的信息和建议不会发生任何变更。报告中的信息和意见仅供参考。本人过去不曾与、现在不与、未来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收任何形式的补偿，分析结论不受任何第三方的授意或影响，特此声明。

免责声明

华安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。本报告中的信息均来源于合规渠道，华安证券研究所力求准确、可靠，但对这些信息的准确性及完整性均不做任何保证，据此投资，责任自负。本报告不构成个人投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。客户应考虑本报告中的任何意见或建议是否符合其特定状况。华安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。

本报告仅向特定客户传送，未经华安证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。如欲引用或转载本文内容，务必联络华安证券研究所并获得许可，并需注明出处为华安证券研究所，且不得对本文进行有悖原意的引用和删改。如未经本公司授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司并保留追究其法律责任的权利。

投资评级说明

以本报告发布之日起 6 个月内，证券（或行业指数）相对于同期沪深 300 指数的涨跌幅为标准，定义如下：

行业评级体系

- 增持—未来 6 个月的投资收益率领先沪深 300 指数 5%以上；
- 中性—未来 6 个月的投资收益率与沪深 300 指数的变动幅度相差-5%至 5%；
- 减持—未来 6 个月的投资收益率落后沪深 300 指数 5%以上；

公司评级体系

- 买入—未来 6-12 个月的投资收益率领先市场基准指数 15%以上；
- 增持—未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%；
- 中性—未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；
- 减持—未来 6-12 个月的投资收益率落后市场基准指数 5%至 15%；
- 卖出—未来 6-12 个月的投资收益率落后市场基准指数 15%以上；
- 无评级—因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。市场基准指数为沪深 300 指数。