

高频数据下的已实现高阶矩因子及改进

--- 因子选股系列之四

报告摘要：

本文作为对高频因子研究的初探。对于收益率分布类别中的高频因子进行研究，我们测试发现构造与改进的因子具有较好的选股表现。

具体来说，首先按文献中方法分别使用 5min 和 10min 的数据针对已实现方差、已实现偏度和已实现峰度构造了周度和月度因子，其分别表示收益率的波动程度，收益率的不对称性和收益率的峰态以及重尾程度。对其在中证 1000 股票池中进行回测，结果表明，5min 线构造的因子好于 10min 线，已实现波动率因子表现不稳定，已实现偏度与已实现峰度周度和月度因子表现较好。

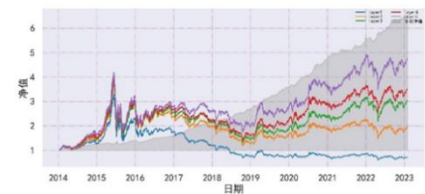
按照前述方式，对已实现高阶矩即已实现超偏度和已实现超尾度进行周度和月度因子构造，其均表示对尾部形态的描述，超偏度类似于偏度，度量重尾程度的同时也蕴含分布的不对称性；超尾度类似于峰度，更高阶矩表示对重尾程度的刻画更为严格。测试结果显示，在中证 1000 的测试中表现一般，可能由于高阶矩的估计对样本量要求较高，5min 线可能导致估计误差。

考虑到 1min 线噪声较大，我们通过下采样平均的方法，对前述构造的因子进行改进，使用中证 1000 股票池 2014 年 1 月至 2023 年 2 月数据进行回测，交易费用设置为双边万五，测试结果显示改进后的周度因子表现明显提升，并且对于更高阶矩构造的因子提升更加显著。其中，subRS_5min_week 因子 Rank IC 0.0588，ICIR 0.8148，多头年化超额 14.51%，多头换手率 1.58，多空对冲组合 Sharpe ratio 3.91；subRK_5min_week 因子 Rank IC 0.0545，ICIR 1.0382，多头年化超额 14.14%，多头换手率 1.31，多空对冲组合 Sharpe ratio 3.61；subRHS_5min_week 因子 Rank IC 0.0526，ICIR 0.8683，多头年化超额 13.83%，多头换手率 1.57，多空对冲组合 Sharpe ratio 3.88；subRHT_5min_week 因子 Rank IC 0.0517，ICIR 1.0187，多头年化超额 14.20%，多头换手率 1.37，多空对冲组合 Sharpe ratio 3.83。

从分年度测试结果来看，我们构造并改进的 subRK_5min_week 因子与 subRHT_5min_week 因子近年来表现持续优异，其中 2022 年多头超额收益分别为 13.03% 和 11.49%。另外改进后的因子与常见的风格因子相关性均较弱，相对而言略偏向流动性与波动率风格。

风险提示：以上分析基于模型结果和历史测算，存在模型失效风险。

subRHT_5min_week 因子分层回测



相关报告

《2022 年四季度公募主动权益基金持仓解析》

--20230206

《上月动量、技术、预期因子表现较优》

--20230203

《上月波动率、流动性、技术因子表现相对较优》

--20230104

《基于异常现金流模型的真实盈余管理因子（二）》

--20230103

《东北金工 2023 年年度策略报告：长风破浪会有时》

--20221224

证券分析师：王琦

执业证书编号：S0550521100001

021-61002390 wangqi_5636@nesc.cn

研究助理：贾英

执业证书编号：S0550122060006

13666061675 jiaying@nesc.cn

目 录

1.	引言	4
2.	已实现高阶矩因子构造及表现	4
2.1.	基本定义	4
2.2.	已实现矩因子	5
2.3.	已实现超偏度因子与已实现超尾度因子	7
3.	改进因子与测试	8
3.1.	下采样平均处理与因子表现对比	8
3.2.	改进后因子分年度测试及相关性	14
4.	总结	16
5.	参考文献	17
6.	风险提示	17

图表目录

图 1: subRS_5min_week 因子分层回测结果	9
图 2: subRS_5min_week 因子 IC 序列统计	9
图 3: subRK_5min_week 因子分层回测结果	9
图 4: subRK_5min_week 因子 IC 序列统计	9
图 5: subRHS_5min_week 因子分层回测结果	10
图 6: subRHS_5min_week 因子 IC 序列统计	10
图 7: subRHT_5min_week 因子分层回测结果	10
图 8: subRHT_5min_week 因子 IC 序列统计	10
图 9: 下采样处理后已实现矩周度因子表现对比	11
图 10: 下采样处理后已实现高阶矩周度因子表现对比	11
图 11: subRS_5min_month 因子分层回测结果	12
图 12: subRS_5min_month 因子 IC 序列统计	12
图 13: subRK_5min_month 因子分层回测结果	12
图 14: subRK_5min_month 因子 IC 序列统计	12
图 15: subRHS_5min_month 因子分层回测结果	12
图 16: subRHS_5min_month 因子 IC 序列统计	12
图 17: subRHT_5min_month 因子分层回测结果	13
图 18: subRHT_5min_month 因子 IC 序列统计	13
图 19: 下采样处理后已实现矩月度因子表现对比	13
图 20: 下采样处理后已实现高阶矩月度因子表现对比	14
图 21: 下采样处理后已实现各阶矩周度因子与 Barra CNE5 因子的相关性	16
图 22: 下采样处理后已实现各阶矩月度因子与 Barra CNE5 因子的相关性	16

表 1：已实现矩周度因子回测结果	6
表 2：已实现矩月度因子回测结果	7
表 3：已实现高阶矩周度因子回测结果	8
表 4：已实现高阶矩月度因子回测结果	8
表 5：1min 级别数据计算的已实现高阶矩因子回测结果.....	8
表 6：下采样平均改进后的已实现高阶矩周度因子回测结果	9
表 7：下采样平均改进后的已实现高阶矩月度因子回测结果	12
表 8：subRS_5min_week 因子分年度回测结果	14
表 9：subRK_5min_week 因子分年度回测结果	15
表 10：subRHS_5min_week 因子分年度回测结果	15
表 11：subRHT_5min_week 因子分年度回测结果.....	15

1. 引言

高频数据在量化研究领域的重要性日益增强。随着传统量价信息挖掘的深入，由低频量价信息构造的因子选股能力不断减弱，而高频信息对微观市场行为有更加精细的描述，其中可能蕴含传统数据不含有的 Alpha 信息。所以，从高频量价数据中有效地提取信息并构造因子来指导日内交易或将信息整合低频化指导低频交易，均是值得深入研究的领域。目前高频因子的分类主要有收益率分布相关因子、成交量分布相关因子和量价相关性因子等。本篇报告作为高频因子研究的初探，介绍并挖掘已实现高阶矩（Realised high-order moment）因子以及其相关衍生因子，描述高频收益率的分布特性。

矩 (Moment) 的概念起源于物理学。在概率论中矩母函数 (Moment-generating function) 与分布 (Distribution) 一一对应，矩是矩母函数生成，故矩是对分布的特征描述。一般来说，比较两个分布是否相似，除了比较分位数 (Quantile) 差异之外，还可以比较各阶矩的差异。其中，一阶矩期望 (Expectation) 表示数据的均值，二阶矩方差 (Variance) 表示数据偏离均值的程度，三阶矩偏度 (Skewness) 度量分布的不对称性，四阶矩峰度 (Kurtosis) 表示分布的峰态及尾部形态。

在统一的视角下，假设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是一个概率空间 (Probability space)， $X: \Omega \rightarrow \mathbb{R}$ 是一维随机变量 (Random variable)，如果考虑 k 阶原点矩的定义，

$$\mathbb{E}[X^k] = \int_{\Omega} X^k(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}} x^k dF(x),$$

第二个式子是 Lebesgue—Stieltjes 积分，其中 F 表示 X 的累积分布函数 (Cumulative distribution function; CDF)。从定义可以看出，高阶矩描述了尾部或者极端值的影响，可以作为对分布重尾现象 (Heavy tail) 的测度，阶数 k 可以理解作为一种权重，较大的 k 会放大尾部极端值的影响，这需要极小的极端值概率才能与之平衡。换句话说，当极端值附近的概率太小时，低阶矩难以有效地反映尾部极端值，而高阶矩对于尾部的积分权重更高，即使对于较小的尾部概率，尾部极端值也能更好地被反映出来。

高频收益率通常也存在重尾现象，高频数据的尾部信息中可能含有对未来收益率预测有价值的信息，值得被关注。不同于传统低频情景，在高频数据中高阶矩捕捉到的信息不是无效噪声，而可能是尾部形态中蕴含的风险信息。另外，对于高阶矩的估计需要更大的数据量作为支持，所以其更适用于高频情景。高频数据与高阶矩的应用较为契合。

本文的内容安排如下：第二章介绍已实现矩因子和已实现高阶矩因子的构造和测试表现；第三章对原始因子进行改进，并对因子表现进行了测试，另外选出绩优因子并进行相关性检验；第四章总结。

2. 已实现高阶矩因子构造及表现

2.1. 基本定义

在介绍因子构造之前，为了方便叙述，首先引入一些记号和定义，我们这里参考[1]中的设定。

假设股票的对数价格（Log-price）是一个跳扩散过程（Jump-diffusion process），i.e. 其满足以下随机微分方程（Stochastic differential equation; SDE）：

$$p_t = p_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \sum_{0 \leq \tau \leq t} \Delta J_\tau,$$

其中，漂移项（Drift） μ 是一个局部有界（Locally bounded）的可料（Predictable）过程，扩散项（Diffusion） σ 是一个正的 càdlàg 过程， J 是纯跳（Pure jump）过程，另外， $\Delta J_\tau = J_\tau - J_{\tau-}$ 。

考虑离散时间情形，将一天的时间分为 N 份，则 $p_{t,i}$ 表示第 t 天中第 i 个时点股票的对数价格。定义第 t 天中第 i 个时点股票的对数收益率（Log-return）为

$$r_{t,i} = p_{t,i} - p_{t,i-1}, \quad i = 1, \dots, N.$$

这些记号与定义是后文因子构造的基础。

2.2. 已实现矩因子

作为引入，本节首先介绍著名的已实现方差因子（Realised variance），可以参考[2]，然后介绍[1]中构造的已实现偏度因子（Realised skewness）以及已实现峰度因子（Realised kurtosis）。

已实现方差是对日内收益率波动的一个较好的度量，其定义为

$$RV_t = \sum_{i=1}^N r_{t,i}^2.$$

如果记第 t 天的时间是从0到 T ，那么[3]证明了 RV_t 依概率收敛（Convergence in probability）到积分波动率与日内对数价格跳跃之和，i.e.

$$\sum_{i=1}^N r_{t,i}^2 \xrightarrow{P} \int_0^T \sigma_s^2 ds + \sum_{0 \leq \tau \leq T} (\Delta p_\tau)^2.$$

[1]中构造了已实现偏度与已实现峰度，它们分别表示日内收益率的偏移程度和尾部形态。定义分别为

$$RS_t = \sqrt{N} \frac{\sum_{i=1}^N r_{t,i}^3}{RV_t^{3/2}}, \quad RK_t = N \frac{\sum_{i=1}^N r_{t,i}^4}{RV_t^2}.$$

[4]中证明了两者的极限情况

$$\sum_{i=1}^N r_{t,i}^3 \xrightarrow{P} \sum_{0 < \tau \leq T} (\Delta p_\tau)^3, \quad \sum_{i=1}^N r_{t,i}^4 \xrightarrow{P} \sum_{0 < \tau \leq T} (\Delta p_\tau)^4.$$

这表示日内收益率的高幂次和依概率收敛到日内对数价格跳跃高次幂的和。与已实现方差相比，极限情况时积分波动率那一项消失了。所以在理论上来说，已实现偏度和已实现峰度对日内对数价格跳跃的描述相比于已实现方差更为有效。另外可以发现，奇数幂次和考虑了日内跳跃的方向，即正负方向的跳跃存在抵消，描述了净的跳跃程度；偶数幂次和将正负跳跃方向合并考虑，描述了跳跃绝对量的程度。

依据以上三个已实现矩使用日内数据构建选股因子。已实现方差因子是由高频数据描述的收益率波动，在模型不含有跳过程（Jump process）时，它是日内积分波动率良好的估计量，由于存在低波异象，在逻辑上其因子方向为负。

已实现偏度因子是由高频数据描述的收益率不对称性，因子值为负时表示高频收益率存在左偏，此时较高的预期收益是对左偏的补偿，另外一般认为日内跳跃程度越大的资产表现越不稳定，已实现偏度的分子极限情形是对日内净跳跃程度的描述，所以因子值越低时，资产价格运行越稳定，倾向于有更好的预期表现，同时也可能获得收益分布左偏的补偿。所以在逻辑上其因子方向为负。

已实现峰度因子是由高频数据描述的收益率峰态与尾部形态。因子值较低时对应尾部重尾现象不明显，即尾部风险较小，倾向于有更好的预期表现，在逻辑上其因子方向亦为负。

使用2014年1月1日至2023年1月30日中证1000股票池5min级别与10min级别的高频量价数据计算因子值，基于以上日度因子值分别进行5日与21日移动平均处理构造周度与月度因子值。对周度和月度因子值进行去极值、标准化以及行业市值中性化处理，参见《因子模型理论与实践及因子检验的实证——因子选股系列之一》，之后统一乘以因子方向-1。对于周度因子进行周频调仓回测，对于月度因子进行月频调仓回测，分层回测时层数为5，交易费用按双边万五计算，以下是因子测试结果。

表 1：已实现矩周度因子回测结果

因子名称	Rank IC	ICIR	多头年化收益	多头年化超额	多头换手率	多空年化收益	多空年化波动率	多空Sharpe Ratio	多空最大回撤
RV_5min_week	0.0576	0.4035	21.21%	16.70%	0.7668	23.46%	20.16%	1.16	35.72%
RS_5min_week	0.0577	0.8823	18.02%	13.51%	1.5876	30.12%	7.33%	4.11	5.72%
RK_5min_week	0.0488	1.0068	17.98%	13.47%	1.4022	22.15%	5.88%	3.77	10.61%
RV_10min_week	0.0563	0.3964	21.27%	16.76%	0.8097	22.61%	19.91%	1.14	35.58%
RS_10min_week	0.0549	0.8370	18.42%	13.91%	1.6031	31.41%	7.36%	4.27	5.63%
RK_10min_week	0.0458	1.0502	15.60%	11.09%	1.5037	20.42%	5.67%	3.6	10.55%

数据来源：东北证券，wind，ricequant

表 2：已实现矩月度因子回测结果

因子名称	Rank IC	ICIR	多头年化收益	多头年化超额	多头换手率	多空年化收益	多空年化波动率	多空Sharpe Ratio	多空最大回撤
RV_5min_month	0.0734	0.5733	15.01%	10.50%	0.9205	12.62%	16.83%	0.75	36.34%
RS_5min_month	0.0690	1.0513	15.20%	10.69%	1.5527	18.03%	6.34%	2.84	7.53%
RK_5min_month	0.0683	1.2673	14.60%	10.09%	1.2929	15.37%	5.60%	2.74	6.08%
RV_10min_month	0.0727	0.5659	15.21%	10.70%	0.9388	12.30%	16.89%	0.73	36.76%
RS_10min_month	0.0599	1.0029	12.34%	7.83%	1.5832	14.05%	6.15%	2.28	7.92%
RK_10min_month	0.0638	1.2164	13.86%	9.35%	1.3936	14.90%	5.73%	2.6	8.53%

数据来源：东北证券，wind，ricequant

从测试结果可以看出：

1. 无论是周度因子和月度因子，使用 5min 级别高频数据效果要好于 10min 级别。这种数据精细度差异导致的因子效果差异可能是由于高精度数据中蕴含更多有价值的信息，并且使得对已实现矩的估计更为准确。
2. 对于已实现方差因子 RV，虽然其周度因子和月度因子贡献了较高的 Rank IC 以及多头年化收益，但是其 ICIR、多空对冲组合相关指标等明显较差，这表明因子有选股价值但表现不稳定，信息还需进一步提纯。
3. 关注 5min 级别的因子，在周度因子和月度因子中已实现偏度因子 RS 和已实现峰度因子 RK 表现较好，且 RS 因子表现相对更好，但 RK 月度因子从 ICIR 和多空组合最大回撤来看更为稳定。原因可能是 RS 三阶矩额外含有尾部的不对称性信息，虽然 RK 四阶矩对重尾现象的反映更为严格，但可能由于 5min 级别精细度不够即无法支持对 RK 进行较为精准的估计，对 RK 的表现有一定的影响。但是月度因子的移动平均处理可能在一定程度上减缓了这种估计误差的影响，这可能是 RK 月度因子表现更为稳定的原因。

2.3. 已实现超偏度因子与已实现超尾度因子

如引言所说，高阶矩捕捉到的尾部信息较低阶矩更为丰富。所以一个自然的想法就是仿照上一节已实现矩因子来构造更高阶的已实现矩因子，我们这里推广到了有一定物理意义的已实现超偏度因子与已实现超尾度因子。

高阶矩的物理意义可参考[5][6]，其中超偏度（Hyper-skewness）是五阶矩，超尾度（Hyper-tailedness）是六阶矩。超偏度可以理解为尾部相对于中心在贡献偏度上的相对重要性，更高的超偏度表示高偏度更多来自于尾部贡献。超尾度描述分布的峰态与尾部形态的更多信息。

在统一的视角来看，超偏度类似于偏度，描述了不对称性；而超尾度类似于峰度，描述了峰态以及尾部与中心的相对信息。但二者对于尾部的积分权重更大，更深刻的描述了尾部风险。二者均需要大量的数据来减少估计误差。

另外，已实现超偏度因子与已实现超尾度因子的构造如下：

$$RHS_t = N^{3/2} \frac{\sum_{i=1}^N r_{t,i}^5}{RV_t^{5/2}}, \quad RHT_t = N^2 \frac{\sum_{i=1}^N r_{t,i}^6}{RV_t^3}.$$

二者因子方向同样为负，这是因为二者均是对收益率重尾程度的描述且 RHT 对尾部的权重更大，当存在较低的极端尾部风险时，倾向于认为资产有更好的预期表现。

由于上一节的结果表明 5min 级别下构造的因子表现显著好于 10min 级别，所以这里仅使用 2014 年 1 月 1 日至 2023 年 1 月 30 日中证 1000 股票池 5min 级别量价数据构造日度因子。同样地，将日度因子分别做 5 日和 21 日移动平均处理得到周度和月度因子。按上一节的设定进行回测，回测结果如下：

表 3：已实现高阶矩周度因子回测结果

因子名称	Rank IC	ICIR	多头年化收益	多头年化超额	多头换手率	多空年化收益	多空年化波动率	多空Sharpe Ratio	多空最大回撤
RHS_5min_week	0.0506	0.9420	17.15%	12.64%	1.5742	24.22%	5.92%	4.09	4.79%
RHT_5min_week	0.0456	0.9754	17.26%	12.75%	1.4463	20.51%	5.50%	3.73	10.40%

数据来源：东北证券，wind，ricequant

表 4：已实现高阶矩月度因子回测结果

因子名称	Rank IC	ICIR	多头年化收益	多头年化超额	多头换手率	多空年化收益	多空年化波动率	多空Sharpe Ratio	多空最大回撤
RHS_5min_month	0.0650	1.0466	15.07%	10.56%	1.5414	16.94%	5.71%	2.97	6.75%
RHT_5min_month	0.0648	1.2991	14.47%	9.96%	1.3670	14.59%	5.22%	2.79	6.73%

数据来源：东北证券，wind，ricequant

结果表明，无论是周度还是月度因子，5min 级别数据构造的已实现高阶矩因子有着一定的选股能力，但效果不如低阶矩。高阶矩对尾部形态的描述更为深刻，但表现略差的原因可能是 5min 的精细度不足以准确估计高阶矩，使得捕捉到的收益率尾部信息中存在噪声。下一章对此问题做出改进。

3. 改进因子与测试

3.1. 下采样平均处理与因子表现对比

针对前一章提出的因数据量不足对已实现高阶矩产生的估计误差问题，一个初步的想法是直接使用 1min 级别数据，但 1min 收益率中噪声较大，并没有研究 5min 收益率分布性质稳定。以下回测结果也证明了这一点：

表 5：1min 级别数据计算的已实现高阶矩因子回测结果

因子名称	Rank IC	ICIR	多头年化收益	多头年化超额	多头换手率	多空年化收益	多空年化波动率	多空Sharpe Ratio	多空最大回撤
RHT_1min_week	0.0414	0.7820	18.16%	13.65%	1.3179	19.32%	5.66%	3.41	6.57%
RHS_1min_week	0.0412	0.8790	17.15%	12.64%	1.5372	20.15%	5.09%	3.96	6.57%
RHT_1min_month	0.0564	1.0475	14.37%	9.86%	1.2801	13.60%	5.10%	2.67	7.22%
RHS_1min_month	0.0625	1.3868	14.75%	10.24%	1.4929	14.71%	4.87%	3.02	4.51%

数据来源：东北证券，wind，ricequant

我们参考[7]中的做法，在更高频的数据网上做下采样（Subsampling）平均，具体来说，以 1min 级别的收盘价数据作为数据网，令采样步长为 5 对数据网进行下采样，得到一个 5min 级别的数据网，用来计算一次因子值，然后初始采样点前进一步，再次进行下采样计算因子值，如此往复并将得到的 5 个因子值平均作为新的因子值。

这种方式对高频数据充分利用，在不改变 5min 收益率作为研究对象的同时降低了对高阶矩的估计误差，另外平均过程也使得因子更加稳定，逻辑上对因子效果有一

定的提升，并期望在一定程度上缓解上一章提出的问题。

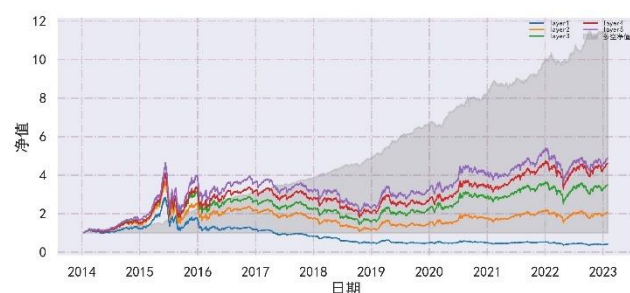
在下采样平均处理后得到的因子是日度因子，按前一章的做法取 5 日和 21 日移动平均分别作为周度因子和月度因子。对因子进行去极值、标准化以及市值行业中性化后乘以因子方向-1。回测的设定与上一章一致，对周度因子进行周频回测，对月度因子进行月频回测，交易费用设定为双边万五。以下是经过下采样平均改进后的因子回测结果。

表 6：下采样平均改进后的已实现高阶矩周度因子回测结果

因子名称	Rank IC	ICIR	多头年化收益	多头年化超额	多头换手率	多空年化收益	多空年化波动率	多空Sharpe Ratio	多空最大回撤
subRS_5min_week	0.0588	0.8148	19.02%	14.51%	1.5845	30.98%	7.93%	3.91	6.40%
subRK_5min_week	0.0545	1.0382	18.65%	14.14%	1.3096	23.40%	6.48%	3.61	12.42%
subRHS_5min_week	0.0526	0.8683	18.34%	13.83%	1.5654	25.82%	6.65%	3.88	5.58%
subRHT_5min_week	0.0517	1.0187	18.71%	14.20%	1.3661	22.92%	5.98%	3.83	10.24%

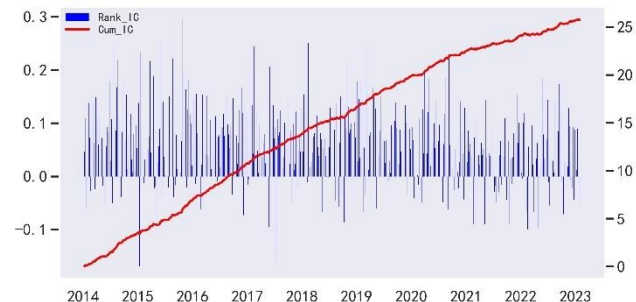
数据来源：东北证券，wind，ricequant

图 1：subRS_5min_week 因子分层回测结果



数据来源：东北证券，wind，ricequant

图 2：subRS_5min_week 因子 IC 序列统计



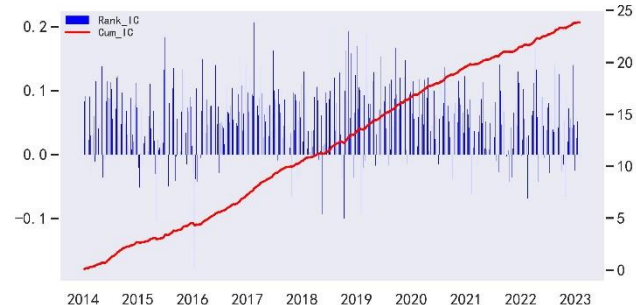
数据来源：东北证券，wind，ricequant

图 3：subRK_5min_week 因子分层回测结果



数据来源：东北证券，wind，ricequant

图 4：subRK_5min_week 因子 IC 序列统计



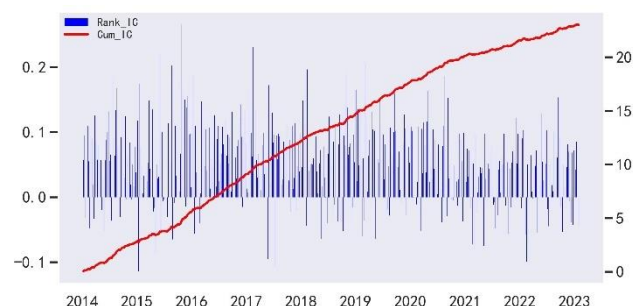
数据来源：东北证券，wind，ricequant

图 5: subRHS_5min_week 因子分层回测结果



数据来源: 东北证券, wind, ricequant

图 6: subRHS_5min_week 因子 IC 序列统计



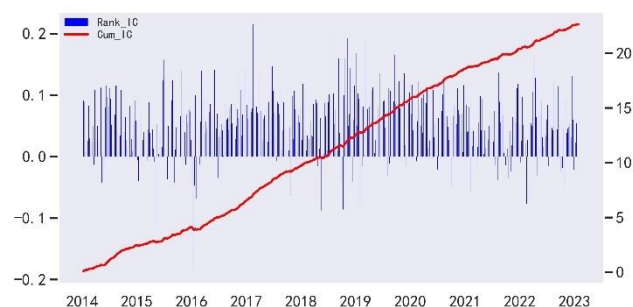
数据来源: 东北证券, wind, ricequant

图 7: subRHT_5min_week 因子分层回测结果



数据来源: 东北证券, wind, ricequant

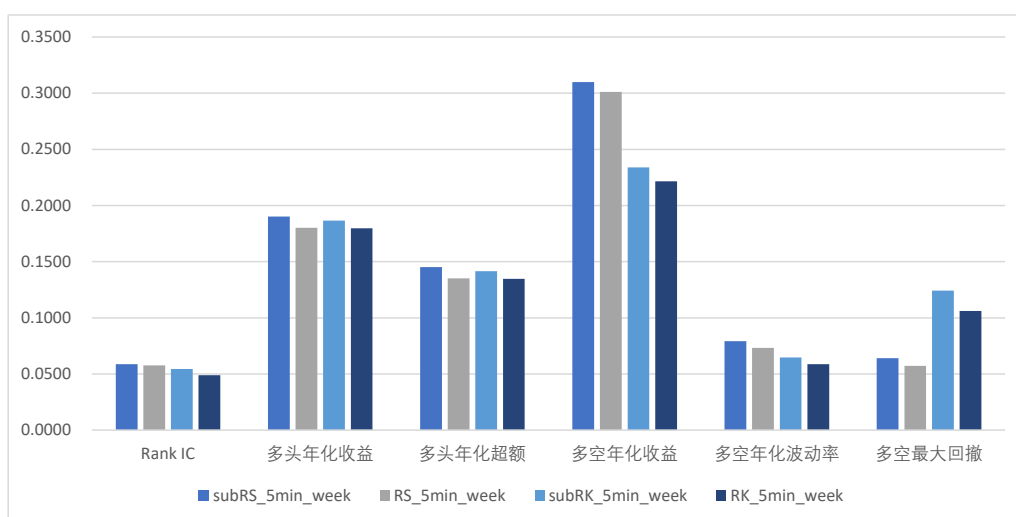
图 8: subRHT_5min_week 因子 IC 序列统计



数据来源: 东北证券, wind, ricequant

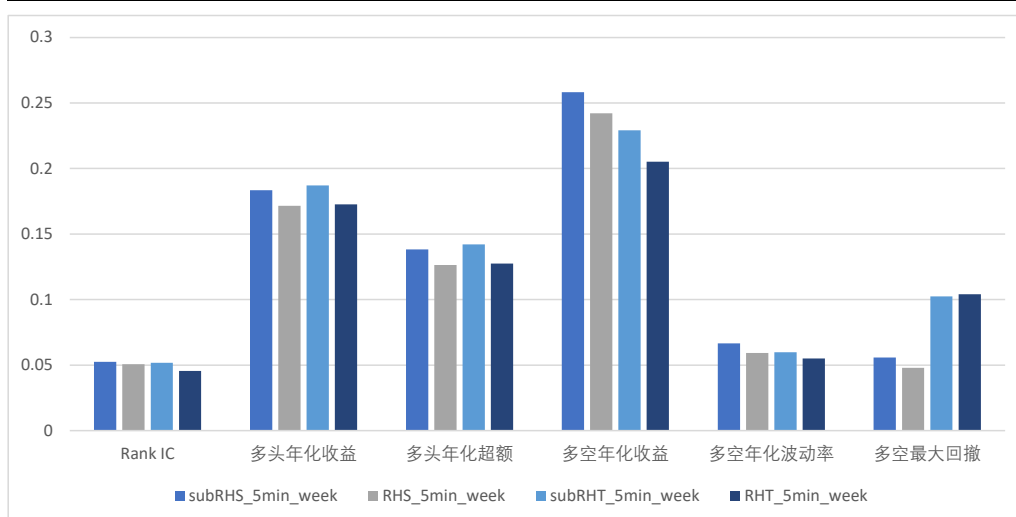
从周度因子表现来看, 在 Rank IC 和 ICIR 方面, 四个因子均有较好的选股能力 subRS_5min_week 和 subRK_5min_week 有着更高的 Rank IC 平均值, 而 subRK_5min_week 和 subRHT_5min_week 的 Rank IC 序列更为稳定。在分层回测方面, subRS_5min_week 和 subRHT_5min_week 多头组合表现优异, 分别取得 14.51% 和 14.20% 的年化超额收益; 改进后的因子单调性均较优, 且已实现高阶矩因子的分层效果更为明显; 多空对冲组合净值曲线平稳上升, 表明因子选股效果持续稳定且优异; subRS_5min_week 和 subRHS_5min_week 的多空对冲组合表现更好, 多空组合年化收益分别为 30.98% 和 25.82%, 多空组合 Sharpe Ratio 分别为 3.91 和 3.88, 多空组合最大回撤分别为 6.40% 和 5.58%。

图 9：下采样处理后已实现矩周度因子表现对比



数据来源：东北证券，wind，ricequant

图 10：下采样处理后已实现高阶矩周度因子表现对比



数据来源：东北证券，wind，ricequant

将下采样平均改进后的周度因子表现与原始周度因子对比，可以看出改进后的因子表现明显提升，在不过多增大风险的情况下，显著的提升了因子收益，使得我们改进后的因子取得优异的表现。另外可以发现，下采样对于高阶矩因子的提升效果相对于低阶矩因子更为明显，这与改进的逻辑一致，即下采样平均降低了对高阶矩的估计误差，提升了稳定性。另外高阶矩捕捉极端尾部信息、度量重尾程度，其估计相对于低阶矩对数据量要求较高，因此下采样平均处理对由高阶矩构造的因子提升更为显著。

表 7：下采样平均改进后的已实现高阶矩月度因子回测结果

因子名称	Rank IC	ICIR	多头年化收益	多头年化超额	多头换手率	多空年化收益	多空年化波动率	多空Sharpe Ratio	多空最大回撤
subRS_5min_month	0.0742	1.0852	13.01%	8.50%	1.5260	17.38%	6.85%	2.54	10.57%
subRK_5min_month	0.0697	1.1994	13.38%	8.87%	1.1978	15.72%	6.00%	2.62	6.32%
subRHS_5min_month	0.0693	1.0693	12.39%	7.88%	1.5139	15.71%	6.13%	2.56	8.96%
subRHT_5min_month	0.0673	1.2319	13.73%	9.22%	1.2927	15.36%	5.73%	2.68	6.22%

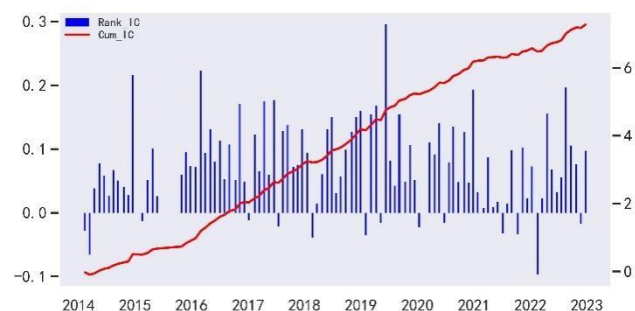
数据来源：东北证券，wind，ricequant

图 11：subRS_5min_month 因子分层回测结果



数据来源：东北证券，wind，ricequant

图 12：subRS_5min_month 因子 IC 序列统计



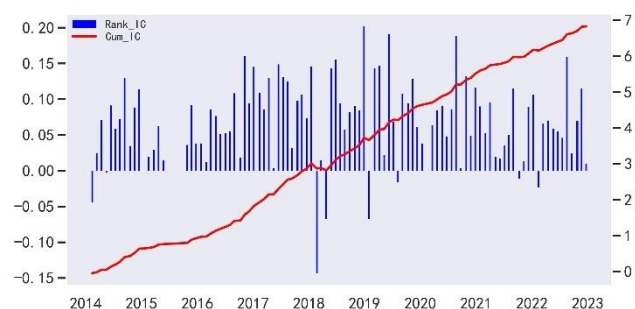
数据来源：东北证券，wind，ricequant

图 13：subRK_5min_month 因子分层回测结果



数据来源：东北证券，wind，ricequant

图 14：subRK_5min_month 因子 IC 序列统计



数据来源：东北证券，wind，ricequant

图 15：subRHS_5min_month 因子分层回测结果



数据来源：东北证券，wind，ricequant

图 16：subRHS_5min_month 因子 IC 序列统计



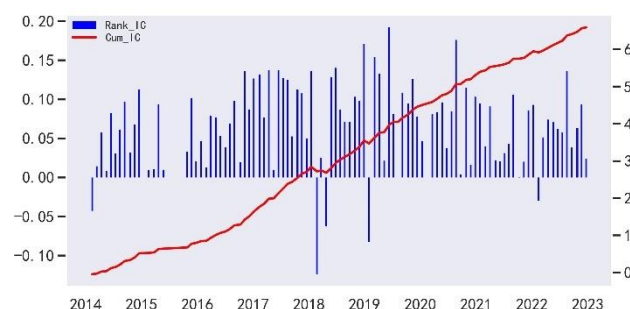
数据来源：东北证券，wind，ricequant

图 17: subRHT_5min_month 因子分层回测结果



数据来源：东北证券，wind，ricequant

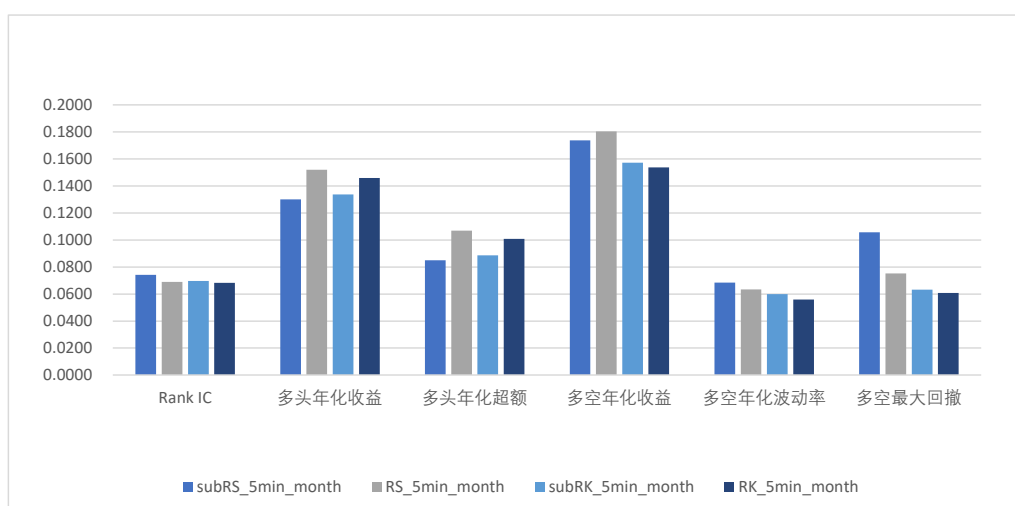
图 18: subRHT_5min_month 因子 IC 序列统计



数据来源：东北证券，wind，ricequant

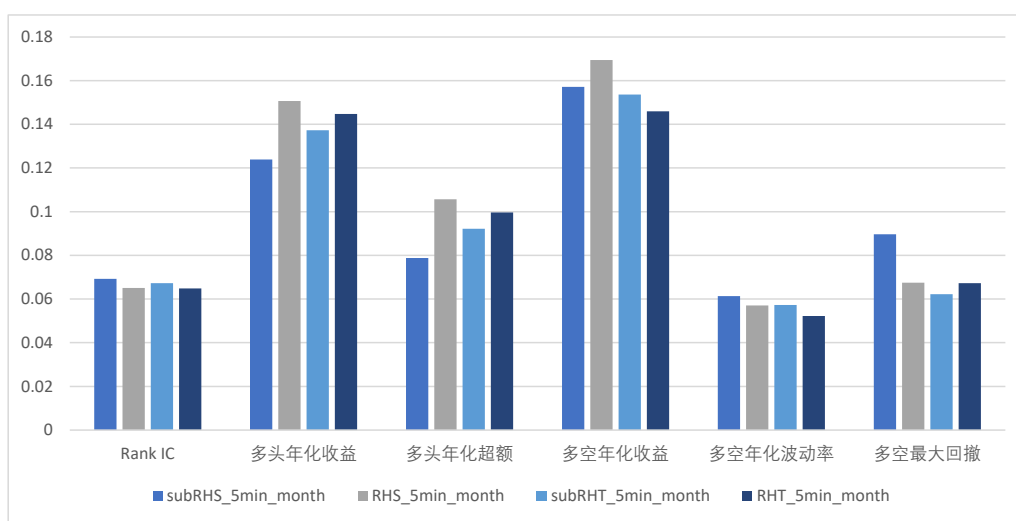
从月度因子表现来看，Rank IC 方面和 ICIR 方面四个因子表现均较为优秀，subRS_5min_month 因子 Rank IC 平均值达到了 0.0742，subRHT_5min_month 因子 ICIR 达到了 1.2319，因子值与未来收益率相关性显著且稳定。因子分层回测方面，subRHT_5min_month 与 subRK_5min_month 分层单调性较好，分别取得了 9.22% 与 8.87% 的多头年化超额，多空对冲组合净值相对于其他两个因子更为平稳，回撤更小。

图 19: 下采样处理后已实现矩月度因子表现对比



数据来源：东北证券，wind，ricequant

图 20：下采样处理后已实现高阶矩月度因子表现对比



数据来源：东北证券，wind，ricequant

将下采样平均改进后的月度因子表现与原始月度因子对比，Rank IC 方面有一定的提升，但因子收益方面并没有提升。这可能是因为月度因子构造窗口期比较长，移动平均可能已经降低了估计误差，导致下采样平均不再有提升效果。

结果表明，下采样平均改进对周度因子的表现有大幅的提升，其对高阶矩因子的提升效果强于低阶矩，而对月度因子没有提升效果。

3.2. 改进后因子分年度测试及相关性

鉴于改进后周度因子有更好的表现，本节对前述周度因子进行了分年度回测，以下是详细的回测结果。

表 8：subRS_5min_week 因子分年度回测结果

年度	Rank IC	ICIR	多头年化收益	多头年化超额	多空年化收益	多空年化波动率	多空Sharpe Ratio
2014	0.0717	1.0258	66.34%	34.02%	37.31%	6.77%	5.5126
2015	0.0736	0.7792	135.56%	60.92%	65.40%	11.73%	5.5742
2016	0.0791	1.3687	0.90%	13.23%	35.46%	8.56%	4.1422
2017	0.0625	0.8637	-12.69%	5.25%	24.39%	5.35%	4.5555
2018	0.0551	0.7642	-28.73%	8.68%	25.78%	6.46%	3.9920
2019	0.0686	1.0209	40.83%	14.79%	36.59%	6.89%	5.3124
2020	0.0553	0.7820	23.53%	6.52%	23.75%	8.44%	2.8134
2021	0.0336	0.5681	28.85%	11.11%	19.66%	7.58%	2.5925
2022	0.0323	0.4581	-16.94%	4.37%	14.33%	7.90%	1.8134
汇总	0.0588	0.8148	19.02%	14.51%	30.98%	7.93%	3.9100

数据来源：东北证券，wind，ricequant

表 9: subRK_5min_week 因子分年度回测结果

年度	Rank IC	ICIR	多头年化收益	多头年化超额	多空年化收益	多空年化波动率	多空Sharpe Ratio
2014	0.0561	1.2543	63.99%	31.67%	31.63%	4.34%	7.2956
2015	0.0393	0.6913	89.72%	15.08%	8.49%	11.64%	0.7298
2016	0.0547	0.9823	-2.64%	9.69%	15.27%	5.47%	2.7922
2017	0.0670	1.4655	-8.49%	9.45%	34.64%	4.66%	7.4378
2018	0.0562	0.8746	-25.34%	12.08%	22.23%	5.31%	4.1883
2019	0.0703	1.2271	40.15%	14.11%	35.57%	5.57%	6.3858
2020	0.0611	1.5122	31.49%	14.48%	33.52%	6.10%	5.4929
2021	0.0403	0.8957	30.15%	12.41%	18.15%	5.93%	3.0608
2022	0.0464	0.8678	-8.29%	13.03%	16.75%	6.19%	2.7075
汇总	0.0545	1.0382	18.65%	14.14%	23.40%	6.48%	3.6100

数据来源：东北证券，wind，ricequant

表 10: subRHS_5min_week 因子分年度回测结果

年度	Rank IC	ICIR	多头年化收益	多头年化超额	多空年化收益	多空年化波动率	多空Sharpe Ratio
2014	0.0573	1.0256	62.01%	29.69%	25.69%	5.86%	4.3811
2015	0.0599	0.7345	128.68%	54.03%	42.37%	10.00%	4.2386
2016	0.0721	1.4457	2.30%	14.62%	30.65%	7.07%	4.3319
2017	0.0626	1.0567	-10.85%	7.09%	28.20%	4.44%	6.3542
2018	0.0504	0.8081	-28.03%	9.39%	23.03%	5.42%	4.2476
2019	0.0627	1.0089	36.46%	10.42%	27.99%	6.06%	4.6199
2020	0.0515	0.9374	23.83%	6.82%	27.06%	6.54%	4.1397
2021	0.0304	0.6233	27.06%	9.32%	17.64%	6.44%	2.7403
2022	0.0276	0.5062	-16.64%	4.67%	10.46%	6.68%	1.5656
汇总	0.0526	0.8683	18.34%	13.83%	25.82%	6.65%	3.8800

数据来源：东北证券，wind，ricequant

表 11: subRHT_5min_week 因子分年度回测结果

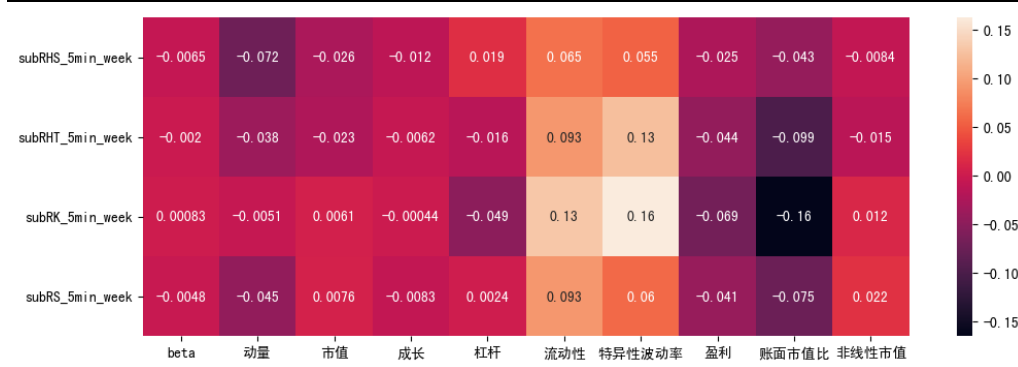
年度	Rank IC	ICIR	多头年化收益	多头年化超额	多空年化收益	多空年化波动率	多空Sharpe Ratio
2014	0.0513	1.1968	64.03%	31.71%	29.44%	4.14%	7.1088
2015	0.0354	0.6920	94.08%	19.43%	9.71%	10.21%	0.9506
2016	0.0486	0.8544	-4.07%	8.25%	13.67%	5.37%	2.5440
2017	0.0656	1.4843	-7.77%	10.16%	35.04%	4.39%	7.9884
2018	0.0562	0.9094	-24.85%	12.57%	22.19%	4.99%	4.4418
2019	0.0687	1.2666	40.19%	14.15%	34.68%	5.22%	6.6434
2020	0.0576	1.3303	28.93%	11.92%	32.42%	5.82%	5.5752
2021	0.0387	0.9025	32.58%	14.84%	17.60%	5.48%	3.2140
2022	0.0438	0.8795	-9.82%	11.49%	16.57%	5.83%	2.8394
汇总	0.0517	1.0187	18.71%	14.20%	22.92%	5.98%	3.8300

数据来源：东北证券，wind，ricequant

以上结果表明，subRS_5min_week 因子与 subRHS_5min_week 因子表现相似，均是由奇数阶矩所构造，蕴含分布的不对称性信息即对尾部的描述均含有方向性。它们在回测期间各年度中均有正的多头超额收益，虽然在 2022 年因子超额收益有所降低，但仍然具有不错的选股能力。subRK_5min_week 因子与 subRHT_5min_week 因子表现相似，其均是由偶数阶矩所构造，蕴含尾部信息即是对双侧重尾程度的刻画。无论从 Rank IC、ICIR 或是因子收益角度来看，它们在回测区间的每个年度均表现稳定且优异，每年的超额收益可观，与前述两个因子不同，在近年来 subRK_5min_week 因子与 subRHT_5min_week 因子表现依然稳健，在 2022 年因子多头超额分别为 13.03%与 11.49%，延续优异的表现。

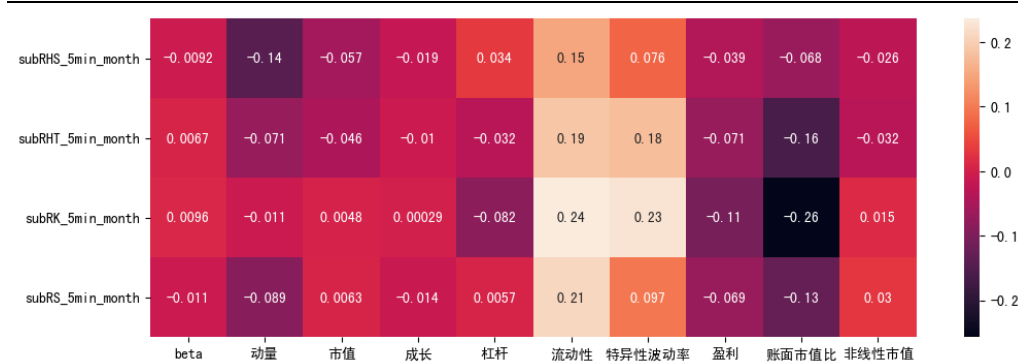
下面计算改进后的周度因子与月度因子与 Barra CNE5 因子的相关性。

图 21：下采样处理后已实现各阶矩周度因子与 Barra CNE5 因子的相关性



数据来源：东北证券，wind，ricequant

图 22：下采样处理后已实现各阶矩月度因子与 Barra CNE5 因子的相关性



数据来源：东北证券，wind，ricequant

结果表明，我们从高频数据中构造并改进的已实现各阶矩因子与常见的大类风格因子相似度较低。subRS 因子与 subRHS 因子与流动性因子有相对较强的相似度，subRK 因子与 subRHT 因子与流动性因子与特异性波动率因子有相对较强的相关性，但总体来说，我们构造并改进的因子与常见风格并没有显著相关性。

4. 总结

首先按文献中方法分别使用 5min 和 10min 的数据针对已实现方差、已实现偏度和已实现峰度构造了周度和月度因子，其分别表示收益率的波动程度，收益率的不对称性和收益率的重尾程度。对其在中证 1000 股票池中进行回测，结果表明，5min 线构造的因子好于 10min 线，已实现波动率因子表现不稳定，已实现偏度与已实现峰度周度和月度因子表现较好。

按照前述方式，对已实现高阶矩即已实现超偏度和已实现超尾度进行周度和月度因子构造，其均表示对尾部形态的描述，超偏度类似于偏度，度量重尾程度的同时也蕴含分布的不对称性；超尾度类似于峰度，更高阶矩表示对重尾程度的刻画更为严格。在中证 1000 的测试中表现一般，可能由于高阶矩的估计对样本量要求较高，5min 线可能导致估计误差。

考虑到 1min 线噪声较大，我们应用下采样平均的方法，对前述构造的因子进行改进，改进后的周度因子在中证 1000 中的表现明显提升，并且对于更高阶矩构造的因

子提升更加显著。从分年度测试结果来看，我们构造并改进的 subRK_5min_week 因子与 subRHT_5min_week 因子近年来表现持续优异。另外改进后的因子与常见的风格因子相关性均较弱。

之后的工作可以对各阶矩构造的因子进行合成，并剔除可能存在干扰的信息。另外也可以考虑已实现高阶矩因子与其他高频因子的联动效应，形成更有效的选股策略。

5. 参考文献

- [1] Amaya, D., P. Christoffersen, K. Jacobs and A. Vasquez, 2013, Does Realized Skewness Predict the Cross-Section of Equity Returns? Journal of Financial Economics 118.
- [2] Andersen, T.G., and T. Bollerslev, 1998, Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, International Economic Review 39, 885-905.
- [3] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, 2003, Modeling and Forecasting Realized Volatility, Econometrica, 71, 529-626.
- [4] Jacod, J., 2012, Statistics and High Frequency Data, In: M. Kessler, A. Lindner, and M. Sorensen (Eds.), Statistical Methods for Stochastic Differential Equations, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Volume 124.
- [5] Khadem, S., P. Narayan and S. Sharma, 2019, Higher Moments and Exchange Rate Behavior, Financial Review 54, 201-229.
- [6] Kostakis, A., K. Muhammad and A. Siganos, 2012, Higher co-moments and asset pricing on London Stock Exchange. English. Journal of Banking and Finance 36, 913-922. ISSN: 0378-4266.
- [7] Zhang, L., P. Mykland, and Y. Ait-Sahalia, 2005, A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data, Journal of the American Statistical Association 100, 1394-1411.

6. 风险提示

分析基于历史数据与模型，存在模型失效风险，历史数据回测结果不代表未来表现。

研究团队简介:

王琦：帝国理工学院数学与金融荣誉硕士，南开大学统计学学士。2021 年加入东北证券上海证券研究咨询分公司任金融工程首席分析师，研究方向为金融工程。曾任职于兴业财富资产管理有限公司，任 FOF 投资经理。

李严：中山大学统计学本科、复旦大学应用统计硕士，主要研究基金产品、行业轮动方向，2020 年加入东北证券，现任东北证券上海证券研究咨询分公司金融工程组分析师。

贾英：伦敦大学学院金融数学荣誉硕士，厦门大学数学与应用数学本科。2022 年加入东北证券，研究方向为金融工程、因子选股。现任东北证券上海证券研究咨询分公司金融工程组研究助理。

张栋梁：复旦大学金融硕士，南京大学金融学本科，2022 年加入东北证券，研究方向为因子选股，现任东北证券上海证券研究咨询分公司金融工程组研究助理。

重要声明

本报告由东北证券股份有限公司（以下称“本公司”）制作并仅向本公司客户发布，本公司不会因任何机构或个人接收到本报告而视其为本公司的当然客户。

本公司具有中国证监会核准的证券投资咨询业务资格。

本报告中的信息均来源于公开资料，本公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅反映本公司于发布本报告当日的判断，不保证所包含的内容和意见不发生变化。

本报告仅供参考，并不构成对所述证券买卖的出价或征价。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的证券买卖建议。本公司及其雇员不承诺投资者一定获利，不与投资者分享投资收益，在任何情况下，我公司及其雇员对任何人使用本报告及其内容所引发的任何直接或间接损失概不负责。

本公司或其关联机构可能会持有本报告中涉及到的公司所发行的证券头寸并进行交易，并在法律许可的情况下不进行披露；可能为这些公司提供或争取提供投资银行业务、财务顾问等相关服务。

本报告版权归本公司所有。未经本公司书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，须在本公司允许的范围内使用，并注明本报告的发布人和发布日期，提示使用本报告的风险。

若本公司客户（以下称“该客户”）向第三方发送本报告，则由该客户独自为此发送行为负责。提醒通过此途径获得本报告的投资者注意，本公司不对通过此种途径获得本报告所引起的任何损失承担任何责任。

分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师。本报告遵循合规、客观、专业、审慎的制作原则，所采用数据、资料的来源合法合规，文字阐述反映了作者的真实观点，报告结论未受任何第三方的授意或影响，特此声明。

投资评级说明

股票 投资 评级 说明	买入	未来 6 个月内，股价涨幅超越市场基准 15%以上。	投资评级中所涉及的市场基准： A 股市场以沪深 300 指数为市场基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为市场基准；香港市场以摩根士丹利中国指数为市场基准；美国市场以纳斯达克综合指数或标普 500 指数为市场基准。
	增持	未来 6 个月内，股价涨幅超越市场基准 5%至 15%之间。	
	中性	未来 6 个月内，股价涨幅介于市场基准-5%至 5%之间。	
	减持	未来 6 个月内，股价涨幅落后市场基准 5%至 15%之间。	
	卖出	未来 6 个月内，股价涨幅落后市场基准 15%以上。	
行业 投资 评级 说明	优于大势	未来 6 个月内，行业指数的收益超越市场基准。	
	同步大势	未来 6 个月内，行业指数的收益与市场基准持平。	
	落后大势	未来 6 个月内，行业指数的收益落后于市场基准。	

东北证券股份有限公司

 网址: <http://www.nesc.cn> 电话: 400-600-0686

地址	邮编
中国吉林省长春市生态大街 6666 号	130119
中国北京市西城区锦什坊街 28 号恒奥中心 D 座	100033
中国上海市浦东新区杨高南路 799 号	200127
中国深圳市福田区福中三路 1006 号诺德中心 34D	518038
中国广东省广州市天河区冼村街道黄埔大道西 122 号之二星辉中心 15 楼	510630

机构销售联系方式

姓名	办公电话	手机	邮箱
公募销售			
华东地区机构销售			
王一 (副总监)	021-61001802	13761867866	wangyi1@nesc.cn
吴肖寅	021-61001803	17717370432	wuxiaoyin@nesc.cn
李瑞暄	021-61001802	18801903156	lirx@nesc.cn
周嘉茜	021-61001827	18516728369	zhoujq@nesc.cn
陈梓佳	021-61001887	19512360962	chen_zj@nesc.cn
屠诚	021-61001986	13120615210	tucheng@nesc.cn
康杭	021-61001986	18815275517	kangh@nesc.cn
丁园	021-61001986	19514638854	dingyuan@nesc.cn
吴一凡	021-20361258	19821564226	wuyifan@nesc.cn
王若舟	021-61002073	17720152425	wangrz@nesc.cn
华北地区机构销售			
李航 (总监)	010-58034553	18515018255	lihang@nesc.cn
殷璐璐	010-58034557	18501954588	yinlulu@nesc.cn
曾彦戈	010-58034563	18501944669	zengyg@nesc.cn
吕奕伟	010-58034553	15533699982	lyyw@nesc.com
孙伟豪	010-58034553	18811582591	sunwh@nesc.cn
陈思	010-58034553	18388039903	chen_si@nesc.cn
徐鹏程	010-58034553	18210496816	xupc@nesc.cn
曲浩蕴	010-58034555	18810920858	quhy@nesc.cn
华南地区机构销售			
刘璇 (总监)	0755-33975865	13760273833	liu_xuan@nesc.cn
刘曼	0755-33975865	15989508876	liuman@nesc.cn
王泉	0755-33975865	18516772531	wangquan@nesc.cn
王谷雨	0755-33975865	13641400353	wanggy@nesc.cn
张瀚波	0755-33975865	15906062728	zhang_hb@nesc.cn
王熙然	0755-33975865	13266512936	wangxr_7561@nesc.cn
阳晶晶	0755-33975865	18565707197	yang_jj@nesc.cn
张楠淇	0755-33975865	13823218716	zhangnq@nesc.cn
钟云柯	0755-33975865	13923804000	zhongyk@nesc.cn
杨婧	010-63210892	18817867663	yangjing2@nesc.cn
梁家滢	0755-33975865	13242061327	liangjy@nesc.cn
非公募销售			
华东地区机构销售			
李茵茵 (总监)	021-61002151	18616369028	liyinyin@nesc.cn
杜嘉琛	021-61002136	15618139803	dujiachen@nesc.cn
王天鸽	021-61002152	19512216027	wangtg@nesc.cn
王家豪	021-61002135	18258963370	wangjiahao@nesc.cn
白梅柯	021-20361229	18717982570	baimk@nesc.cn
刘刚	021-61002151	18817570273	liugang@nesc.cn
曹李阳	021-61002151	13506279099	caoly@nesc.cn
曲林峰	021-61002151	18717828970	qulf@nesc.cn
华北地区机构销售			
温中朝 (副总监)	010-58034555	13701194494	wenzc@nesc.cn
王动	010-58034555	18514201710	wang_dong@nesc.cn
闫琳	010-58034555	17862705380	yanlin@nesc.cn
张煜苑	010-58034553	13701150680	zhangyy2@nesc.cn