

分析师:

郑兆磊

zhengzhaolei@xyzq.com.cn

S0190520080006

高频研究系列六—时序信息中的 Alpha

2023 年 6 月 29 日

报告关键点

本文中，我们将聚焦于时序信息，从序列自相关性与条件分布两个角度出发，构建了七类高频时序信息因子。在序列自相关性角度，我们基于一阶信息刻画基础高频因子，并基于高阶相关性刻画了个股的非同步交易程度因子。在条件分布中，我们以高波下日内收益率均值因子阐明该方法论的有效性，又进一步构建了股价自相关性下的 VaR 以及“闪电崩盘”概率因子。

相关报告

《高频漫谈》2022-01-04

《高频研究系列二—收益率分布因子构建》2022-01-23

《高频研究系列三—收益率分布中的 Alpha(2)》2022-05-04

《高频研究系列四—成交量分布中的 Alpha》2022-08-29

《高频研究系列五—市场微观结构剖析》2022-11-09

投资要点

- 2022 年以来，兴证金工团队先后推出了阐述高频研究方法论的《高频漫谈》，以及 3 篇高频因子深度研究。在高频漫谈中，我们阐述了高频因子的构建逻辑、因子的回测方法以及高频风险的识别。在后续三篇高频因子研究报告中，我们构建了约 27 个分布信息因子，其中不乏多个思路新颖、具有较强特异性的分布信息因子。
- 本文中，我们将聚焦于时序信息，从序列自相关性与条件分布两个角度出发，构建了七类高频时序信息因子。在序列自相关性角度，我们基于一阶信息刻画基础高频因子，并基于高阶相关性刻画了个股的非同步交易程度因子。在条件分布中，我们以高波下日内收益率均值因子阐明该方法论的有效性，又进一步构建了股价自相关性下的 VaR 以及“闪电崩盘”概率因子。
- 七类因子均展示出极强的选股能力：以 rtn_condVaR(股价自相关性下的 VaR) 为例进行阐述：该因子日度 Rank IC 均值为 4.01%，多空收益率约为 44%，多头收益率约为 21%，夏普比率约为 5.7。
- 进一步等权合成构建时序信息复合因子：我们以周度和月度换仓的形式，测试其在全市场等多个股票池内的表现。复合因子在全市场内周度 Rank IC 均值为 6.64%，ICIR 为 0.98；月度 Rank IC 均值为 8.07%，ICIR 为 1.32。周度测试下因子多头组合年化收益率为 26.7%，同期基准为 10.8%，多空夏普比率为 4.91。因子在不同股票池内的表现同样优秀。

核心表、时序信息复合因子周度测试结果

股票池	IC 均值	ICIR	Top 组年化超额	多空夏普比率
全市场	6.64%	0.98	14.3%	4.91
沪深 300	2.80%	0.23	3.6%	0.91
中证 500	5.21%	0.55	6.5%	1.92
中证 800	4.10%	0.41	5.3%	1.44
中证 1000	6.30%	0.88	11.4%	4.18
国证 2000	7.18%	1.06	14.0%	5.35

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

风险提示：模型结果基于历史数据的测算，在市场环境转变时模型存在失效的风险。

目 录

1、高频研究回顾与时序信息因子研究框架	- 4 -
1.1、高频研究回顾	- 4 -
1.2、时序信息简析与因子回测说明	- 4 -
2、时序自相关性基础因子	- 6 -
2.1、因子构建	- 6 -
2.2、因子表现	- 8 -
3、个股的非同步交易性与高阶自相关性	- 9 -
3.1、股票市场的非同步交易性	- 9 -
3.2、非同步交易程度与高阶时序稳定性	- 10 -
3.3、因子表现	- 11 -
4、条件分布基础因子	- 12 -
4.1、条件分布方法简述——以高波收益率均值为例	- 12 -
4.2、高波收益率均值因子表现与分析	- 12 -
5、考虑股价自相关性的风险度量因子	- 13 -
5.1、股价自相关性和 VaR 推导	- 13 -
5.2、因子表现	- 15 -
6、“闪电崩盘”、连续下跌和“崩盘概率”因子	- 16 -
6.1、“闪电崩盘”现象简析	- 16 -
6.2、连续下跌和泊松分布下的“崩盘”概率因子	- 17 -
6.3、因子表现	- 17 -
7、相关性检验与时序信息复合因子	- 18 -
7.1、时序信息因子正交化处理与相关性检验	- 18 -
7.2、时序信息复合因子构建与测试	- 20 -

图目录

图 1、时序信息高频因子研究方法论	- 5 -
图 2、常见时序自相关性因子多空组合净值	- 9 -
图 3、非同步交易程度较高的股票	- 11 -
图 4、非同步交易程度较低的股票	- 11 -
图 5、LBQ_Comb 因子 Rank IC 和累计 IC	- 12 -
图 6、LBQ_Comb 因子多空组合净值	- 12 -
图 7、股价相关性 VaR 因子 Rank IC 和累计 IC	- 15 -
图 8、股价相关性 VaR 因子多空组合净值	- 15 -
图 9、300659.SZ 日内价格走势	- 16 -
图 10、“崩盘”概率因子 Rank IC 和累计 IC	- 18 -
图 11、“崩盘”概率因子多空组合净值	- 18 -
图 12、时序信息因子正交化后相关性箱型图	- 19 -
图 13、正交化后的时序信息因子间相关性	- 20 -
图 14、时序信息复合因子周度分位数组合测试结果	- 21 -
图 15、时序信息复合因子月度分位数组合测试结果	- 22 -

表目录

表 1、兴证金工高频系列研究内容回顾	- 4 -
表 2、一阶自相关系数因子定义	- 7 -
表 3、D-W 统计量因子定义	- 7 -
表 4、残差自相关系数因子定义	- 7 -
表 5、常见时序自相关性因子 Rank IC 测试结果	- 8 -
表 6、常见时序自相关性因子组合测试结果	- 8 -
表 7、Q 统计量序列标准差与非同步交易逻辑说明	- 10 -
表 8、非同步交易性因子 Rank IC 测试结果	- 11 -
表 9、非同步交易性因子组合测试结果	- 11 -
表 10、高波收益率均值因子 Rank IC 测试结果	- 13 -
表 11、高波收益率均值因子组合测试结果	- 13 -
表 12、股价相关性 VaR 公式解析	- 15 -
表 13、股价相关性 VaR 因子 Rank IC 测试结果	- 15 -
表 14、股价相关性 VaR 因子组合测试结果	- 15 -
表 15、“崩盘”概率因子 Rank IC 测试结果	- 17 -
表 16、“崩盘”概率因子组合测试结果	- 18 -
表 17、时序信息因子定义	- 18 -
表 18、正交化后的时序信息因子 Rank IC 测试结果	- 19 -
表 19、正交化后的时序信息因子组合测试结果	- 19 -
表 20、时序信息复合因子周度 Rank IC 测试结果	- 20 -
表 21、时序信息复合因子月度 Rank IC 测试结果	- 21 -

报告正文

1、高频研究回顾与时序信息因子研究框架

1.1、高频研究回顾

2022 年以来，兴证金工团队先后推出了阐述高频研究方法论的《高频漫谈》，以及《收益率分布因子》、《收益率分布中的 Alpha(2)》与《成交量分布中的 Alpha》高频因子深度研究。在高频漫谈中，我们阐述了高频因子的构建逻辑、因子的回测方法以及高频风险的识别。该篇也是我们整个高频系列的基础篇和框架篇，后续所有的研究均建立在此基础上。据此我们介绍了四类高频指标的信息：分布信息、时序信息、关联信息与另类信息。在后续三篇高频因子研究报告中，我们基于量价数据，构建了约 27 个分布信息因子，其中不乏多个构建思路新颖、具有较强特异性的分布信息因子。**目前所有因子均实现日度“T+0”更新，并上传至云端数据库供下载。**本文中，我们将聚焦于第二类信息——时序信息。

表 1、兴证金工高频系列研究内容回顾

内容	简介
高频研究系列一：高频漫谈	简述高频因子构建逻辑：基于四类信息构造高频指标，两类时序操作生成高频因子，基于信号加权的多空构造方式；简述高频因子回测方法：通过回测结果判断因子有效性，基于时序相关性的特异性判断标准，通过基本面风险模型与统计风险模型结合衡量高频风险。
高频研究系列二：收益率分布因子构建	基于收益率分布信息构建六个常见收益率分布因子，以及两个收益率噪音偏离因子 nos 因子。
高频研究系列三：收益率分布中的 Alpha(2)	我们根据投资者对极端上涨和极端下跌的心理承受能力不同构建极端上涨和极端下跌因子，从日内跳价的信息中刻画大额投资者对于股票的操作能力并进一步构建高频因子，从日内震荡期和跳价期的信息差异中刻画股票的日内价格弹性并进一步构建选股 Alpha 因子。
高频研究系列四：成交量分布中的 Alpha	我们聚焦于日内成交量数据，并基于数据特征构建了七个常见成交量分布因子。进一步，我们根据分钟成交量的不稳定性与成交量极值来构造日内成交量分桶熵和极大值分布因子。此外，我们将日内成交量按收盘价异构，得到同价成交量分布，以此构建了三个同价成交量分布因子。
高频研究系列五：市场微观结构剖析	本文基于高频数据对市场微观结构进行观察与剖析。具体来说，我们从深度、紧密度和弹性三个维度共同衡量微观流动性。进一步，我们确定了指标在针对全市场、大小盘以及中信一级行业上的趋势分析、奇异点判断以及择时方面的应用。

资料来源：兴业证券经济与金融研究院整理

1.2、时序信息简析与因子回测说明

有效市场假说（Efficient Markets Hypothesis, EMH）是由美国著名经济学家尤金·法玛（Eugene Fama）于 1970 年提出并深化的理论。在弱式有效市场中，市场价格已充分反映出证券价格的所有过去历史信息，包括股票的成交价、成交量、卖空金额、融资金额等。如果弱式有效市场假说成立，则股票价格的技术分析失去作用，基本面分析也无法帮助投资者获得超额利润。然而，对于 A 股市场而言，目前主流观点认为市场并未达到弱式有效市场。对于流动性资产（如股票）而言，过去一段时间内的量价信息并不完全反映所有的过往信息。因此，该流动性资产前后信息的关联程度，以及在已知前序信息的基础上样本数据特征，则是衡量其

请务必阅读正文之后的信息披露和重要声明

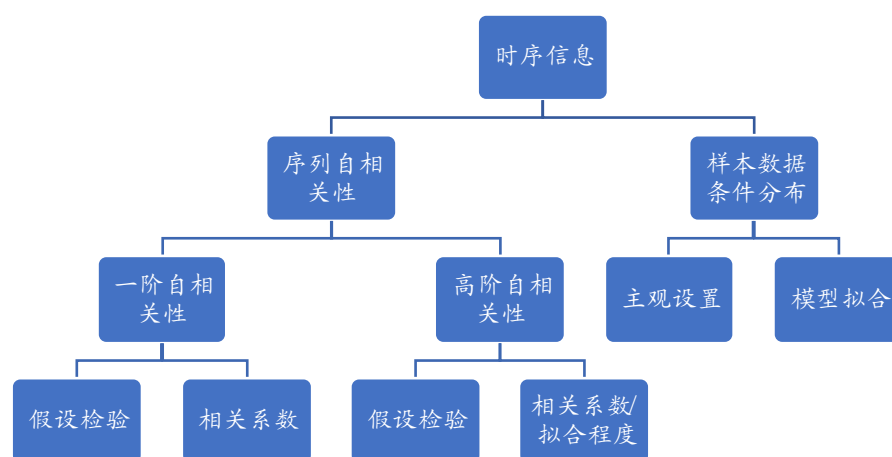
是否近似为弱式有效市场的重要维度。

在高频漫谈中，我们详细阐述了提取日内数据总体分布信息的四种方法。其中，我们定义时序信息需要满足下述定义：

$$g = f(\text{Reorder}(\text{data})) \neq f(\text{data}) \quad (1)$$

其中，**data**为日内的行情数据，如分钟级收盘价、成交量等。指标**g**对于时序重排序函数**Reorder**敏感，数据在时序上的变化会影响指标值，如收益率自相关性 $\rho(r_t, r_{t-1})$ 依赖于收益率序列的排序，重新排列收益率序列后，指标值会发生变化。在此基础之上，我们衍生出基于时序信息相关因子。在本文中，我们将时序信息因子分为两类构建方式：**序列自相关性**与**条件分布**。 r_t

图 1、时序信息高频因子研究方法论



资料来源：兴业证券经济与金融研究院整理

具体来说，时序自相关性是指时间序列中某一个时刻的值和另一个时刻的值之间的相关性，通常作为最为直观的时序信息统计量。在金融领域，学者们普遍认为资产收益率序列存在较强的自相关性与异方差性。为了同时考虑收益率的线性自相关性与方差之间的相关性，常见的 GARCH 模型被用作时间序列模型。本文中，我们针对时序自相关性，引入假设检验和相关系数/拟合程度两种统计量，以衡量日内时序数据自相关性的显著性与相关大小。此外，在非弱式有效市场的背景下，市场和个股内外部条件的变化使得收益率的统计分布特征发生变化，如之前提及的异方差性。对于日内收益率而言，其自身的非独立同分布特征更加明显（具体可参见报告《高频研究系列二——收益率分布因子构建》）。因此，我们在本文中引入条件分布，通过日内过去一段时间的量价信息特性对样本数据进行筛选和重塑，进一步计算筛选之下的样本数据统计特征。

为了更好体现高频因子对于短期行情信息把握的优势，我们采取日频调仓的方式测量因子收益率。具体来说， F_{Di} 是第*i*天日频调仓因子， F_j 是第*j*天的高频指标， n_s 取 15¹：

¹我们同时测试了基于 IC 的加权方式，在不控制换手率的情况下，IC 加权方式能够相对提升

$$F_Mean_{Di} = \frac{1}{n_s} \sum_{j=1}^{n_s} F_j \quad (2)$$

除此之外，我们也可以在时序上取标准差的方式构建因子，以衡量个股在该指标上的离散程度。

$$F_STD_{Di} = \frac{1}{n_s} \sum_{j=1}^{n_s} \left(F_j - \frac{1}{n_s} \sum_{j=1}^{n_s} F_j \right)^2 \quad (3)$$

在本文中如无额外说明，我们因子的回测规则设定如下：

1. 因子回测区间：2014 年 9 月 1 日—2023 年 5 月 4 日；
2. 回测规则：剔除当期不在市、涨跌停以及特殊处理的股票；
3. 回测结果说明：本文中，回测表格中提及的年化波动率、夏普比率、最大回撤、胜率均为因子回测时的多空净值对应统计量，多头换手率为多头组的单边换手率，且我们这里的多空是两分组并按照因子值进行加权（参见《高频漫谈》）。

本文的结构如下：

1. 我们首先聚焦于时序自相关性，从三个维度基于日内时序数据的一阶自相关性刻画因子，进一步从高阶时序稳定性的角度刻画个股的非同步交易性，进而构建因子；
2. 其次，我们聚焦于条件分布，首先以高波动率下的收益率均值角度，引入条件分布的常见刻画方式，并表明该方法可以刻画出具有特异性的因子；
3. 进一步，我们从股价相关性 VaR 和“闪电崩盘”两个角度构建两类时序信息因子；
4. 最后，我们进行中性化与相关性检验，并进一步合成高频时序信息复合因子，并在周度和月度层面测试其不同股票池内的表现。

2、时序自相关性基础因子

2.1、因子构建

前文提及，金融领域的时间序列通常存在着较强的自相关性。其中，一阶自相关性是最为基础和直观的统计量之一。在实际应用中，投资者或学者们通常通过计算相关系数和假设检验来判断时间序列的一阶自相关性是否显著。在本节中，我们同样以上述两个角度，衡量与构建日内数据的一阶相关性统计量。我们引入的数据分别是分钟级收益率与分钟级成交量占比。同时，由于个股在开盘和收盘阶段的量价信息波动较大，因此，我们均剔除了开盘和收盘 10 分钟数据。此外，

因子的多空与 IC 测试表现，但换手率较高。因此本文中根据换手率与因子表现进行调整，在时序上取 15、20 或 25 日均值/标准差。

对于每一类因子，我们均等权合成基于收益率和成交量占比构建的因子，作为该类的复合因子。

➤ 一阶自相关系数

首先，我们基于量价两个维度的时序数据，计算分钟级时序数据的一阶自相关系数，用于判断时序数据的前后相关性。以分钟级收益率数据为例，我们计算个股日内收益率序列的一阶差分相关系数作为高频指标，并进一步计算个股过去 15 日指标的均值作为最终的因子值。因子排序为升序。

表 2、一阶自相关系数因子定义

因子名	因子构造方法
rtn_foc	分钟收益率一阶自相关系数
vol_foc	分钟成交量占比一阶自相关系数
foc_Comb	等权合成 rtn_foc 和 vol_foc 因子

资料来源：兴业证券经济与金融研究院整理

➤ D-W 统计量

D-W 检验是统计分析中常用的一种检验序列一阶自相关的方法。我们直接计算收益率与成交量占比时序数据的 D-W 统计量，其统计量的绝对大小可以判定时序数据是否存在自相关性，以及相关性的正负性。以分钟级收益率数据为例，我们计算个股日内收益率序列的 D-W 统计量作为高频指标，并进一步计算个股过去 15 日指标的均值作为最终的因子值。因子排序为降序。

表 3、D-W 统计量因子定义

因子名	因子构造方法
rtn_DW	分钟收益率 D-W 统计量
vol_DW	分钟成交量占比 D-W 统计量
DW_Comb	等权合成 rtn_DW 和 vol_DW 因子

资料来源：兴业证券经济与金融研究院整理

➤ 残差自相关系数

上述分析中，我们仅检验日内时序数据的一阶自相关性。然而，对于分钟级数据而言，自身本身可能存在着高阶相关性。在高阶自相关性存在的情况下，以 AR(1) 作为模型的残差将具有自相关性，使得线性回归模型相对失效。在此角度之下，我们尝试测试时序数据在 AR(1) 模型下的残差一阶自相关回归模型的系数。我们进一步计算个股过去 15 日指标的标准差作为最终因子值。因子排序为升序。

表 4、残差自相关系数因子定义

因子名	因子构造方法
rtn_rho	分钟收益率残差自相关系数
vol_rho	分钟成交量残差自相关系数
rho_Comb	等权合成 rtn_rho 和 vol_rho 因子

资料来源：兴业证券经济与金融研究院整理

2.2、因子表现

我们首先测试上述 9 个因子的表现。首先从日度 IC 测试结果上看，绝大多数的常见时序信息因子 IC 均值在 4% 以上，表现出较好的股价预测能力。此外，三类复合因子的表现更为优秀，以 foc_Comb 为例，该因子的日度 Rank IC 均值为 6.85%，ICIR 为 0.76，有效性较强。

表 5、常见时序自相关性因子 Rank IC 测试结果

因子名称	IC 均值	IC 标准差	ICIR	T 统计量
rtn_foc	6.19%	11.75%	0.53	24.19
vol_foc	5.77%	6.27%	0.92	42.22
rtn_DW	3.96%	10.51%	0.38	17.29
vol_DW	4.02%	11.45%	0.35	16.11
rtn_rho	4.62%	7.29%	0.63	29.11
vol_rho	1.87%	7.16%	0.26	11.99
foc_Comb	6.85%	9.00%	0.76	34.94
DW_Comb	4.35%	11.55%	0.38	17.28
rho_Comb	4.02%	6.80%	0.59	27.16

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从日度组合测试结果上看，我们构建的常见时序信息因子的表现均相对优秀，大多数因子的多空夏普大于 4，三类复合因子的表现更为优秀。具体来看，从多空组合测试上看，vol_foc 因子的多空收益率在 67% 左右，多头收益率高达 29%，夏普比率在 9.6 左右。

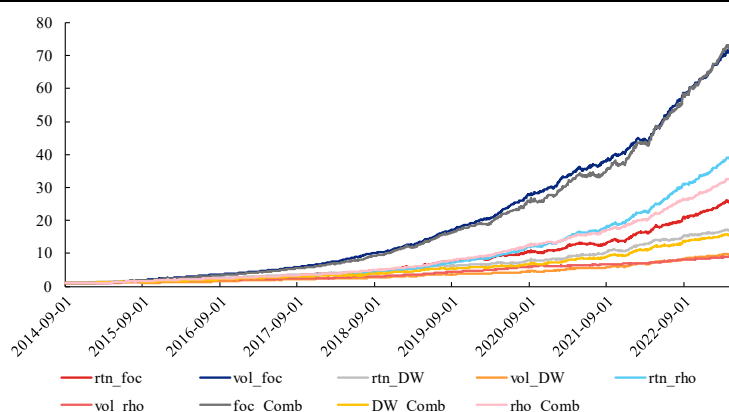
表 6、常见时序自相关性因子组合测试结果

因子名称	多空收益率	多头收益率	空头收益率	年化波动率	夏普比率	最大回撤	多头换手率
rtn_foc	47.8%	22.0%	-19.5%	11.9%	4.03	-8.8%	7.1%
vol_foc	67.0%	29.0%	-23.1%	7.0%	9.56	-5.4%	12.1%
rtn_DW	41.0%	20.6%	-16.0%	9.6%	4.29	-7.1%	7.0%
vol_DW	32.0%	17.0%	-13.3%	10.3%	3.11	-12.1%	7.8%
rtn_rho	55.3%	21.2%	-23.0%	9.1%	6.12	-7.2%	9.0%
vol_rho	30.1%	12.9%	-12.8%	6.5%	4.66	-6.5%	10.8%
foc_Comb	67.4%	28.3%	-24.7%	9.9%	6.80	-7.8%	9.7%
DW_Comb	39.7%	21.1%	-15.2%	10.3%	3.85	-9.0%	6.8%
rho_Comb	51.9%	21.7%	-20.2%	7.4%	7.02	-4.7%	10.5%

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从多空净值曲线上看，各个因子的多空净值长期呈现上升趋势，且最近几年无明显回撤，表现十分稳定。

图 2、常见时序自相关性因子多空组合净值



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

3、个股的非同步交易性与高阶自相关性

在前文中，我们从自相关性的相关系数与假设检验的角度出发，构建了三类常见的时序信息因子，其重点主要聚焦在对日内时序信息的一阶自相关性的分析上。结果表明，各类因子具有较好的选股能力。

上文更多地是从数据统计的角度出发。在本章中，我们基于股票对于信息到来以及股价反应速度不一致导致的非同步交易特性，从时序的高阶相关性角度出发，构建时序信息中，自相关性的特异性因子。

3.1、股票市场的非同步交易性

在有效的证券市场上，每一只证券的价格都应该迅速反映影响其自身价格的信息（来自于市场、交易情绪等）。但现实中，市场总是存在着一定的交易摩擦、投资者的预期也往往有一定程度的刚性：信息到达市场中不同投资者的时间会存在不同程度的差异、知情交易者的参与程度等等原因。这些原因往往会导致有些证券对信息的反映出现时滞，且在任一分钟内，不同投资者做出的交易决策并不基于相似的信息，导致当前时刻下的交易信息混杂，最终导致股价波动较大。这一现象在学术界可以类比于**非同步交易**。

对于某一证券而言，**信息的透明程度与个股的流动性是两个决定证券非同步交易程度的重要因素**。较多的知情交易者在获得了来自市场、公司和情绪面等多种信息后将做出异于非知情交易者的交易，高频次、不同信息来源的知情交易者导致股价波动较大，提高了非同步交易的程度；此外，流动性较低的股票往往代表着股票日内的价格发现过程相对低效，股价波动的随机性较小。从时序相关性的角度看，流动性差的股票吸收交易冲击和从冲击中恢复的能力较差，导致价格更加容易出现异于正常股价波动的趋势性走势，这进一步放大了非同步交易的程度。

3.2、非同步交易程度与高阶时序稳定性

因此，本质上来说，非同步交易是在衡量股票日内的信息丰富度以及对于信息的消化速度：信息丰富度高的股票，其当期股价可能与过去较长一段时间的信息均有关联，表现为在不同回望区间下，高阶相关性的显著性均较高。此外，个股信息消化速度的快慢反映为在不同回望区间下，股票是否都具有较高的相关性。若消化速度较差，说明更长的时段内，股价的高阶相关性显著性将逐步提升。综上，我们可以从时间序列的高阶相关性为出发点，刻画个股的非同步交易程度。

在具体计算层面，我们类比于 ACF 检验，引入多次 LB 检验来衡量个股在不同时间区间内高阶相关性的显著性与稳定性。我们计算 LB 检验中 Q 统计量序列的标准差，若标准差较大，说明不同回望区间下，Q 统计量大小差异较大，通常表现为较短的回望区间内 Q 统计量远小于较长回望区间内的 Q 统计量。因此，该股的信息丰富度较高，过去一段时间内的信息仍可以显著作用于当期的行情变动；此外，消化速度较差的股票其 Q 统计量数值的较大，进一步放大了长短回望区间下的 Q 统计量差异。因此，对于 Q 统计量序列标准差而言，其逻辑特征如下。

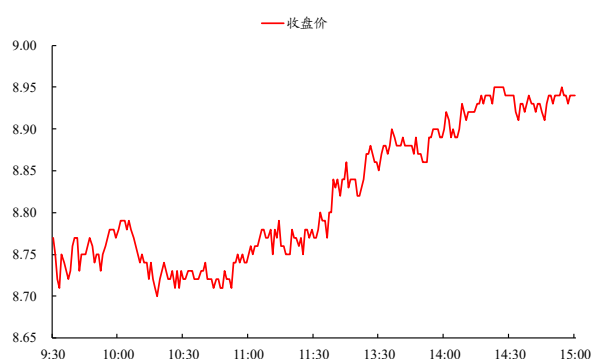
表 7、Q 统计量序列标准差与非同步交易逻辑说明

	流动性好	流动性差
证券透明度较低，知情交易者较多	标准差偏大 非同步交易程度较大	标准差大 非同步交易程度大
证券透明度较高，知情交易者较少	标准差小 非同步交易程度小	标准差偏小 非同步交易程度较小

资料来源：兴业证券经济与金融研究院整理

以上述逻辑为基础，我们找出在某日 Q 统计量序列标准差值较大与较小的股票，对比其股价走势。其中，左侧为当日标准差较高，即非同步交易程度较高的股票，右侧则为非同步交易程度较低的股票。两者在当日的价格走势相近，都存在着先小幅下跌，再上涨后收盘的走势。对比后我们可以发现，左侧本身的流动性逊于右侧，其部分时间段股价并未波动。此外，右侧的股票在上涨阶段的波动较为剧烈，在短时间大幅上涨，左侧则近似于缓步上涨。我们推测：左侧由于自身的关注度较低，且知情交易者占比偏高，当日该股股价被多次少量的冲高，其信息源可能来自于早期的某一时刻，多数投资者对于当日信息的解读与操作时间节点不一致，或者知情交易者通过多次少量交易隐藏自己的交易行为，导致非同步交易程度较大；右侧在当日出现大幅度的价格抬升，这或许是由于当日该股关注度较高，知情交易者占比较小，同时多数交易者在同一时间基于相近信息做出交易，导致价格波动较大。因此，我们认为 Q 统计量序列标准差可以从某一角度刻画股票的非同步交易性。

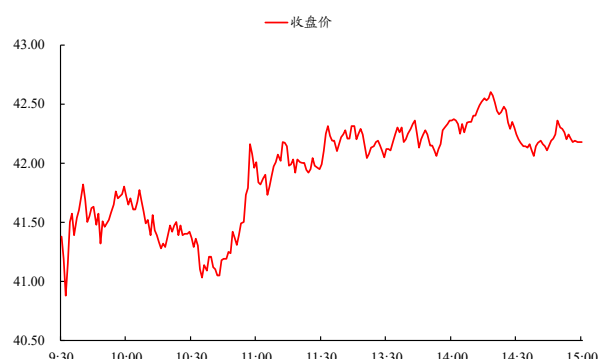
图 3、非同步交易程度较高的股票（000156.SZ）



资料来源：上交所、深交所行情数据，兴业证券经济与金融研究院整理

数据日期：2023 年 3 月 31 日

图 4、非同步交易程度较低的股票（002368.SZ）



资料来源：上交所、深交所行情数据，兴业证券经济与金融研究院整理

数据日期：2023 年 3 月 31 日

在具体计算维度，我们同时以分钟级收益率和分钟级成交量占比作为时间序列数据，统计其在量和价上的高阶稳定性，并最终取时序上 15 日均值作为最终因子，分为叫做 rtn_LBQ 与 vol_LBQ 。均值越大，说明该股的非同步交易程度越高，个股在低流动性下股价被操纵的可能性较高，下行风险较高。此外，我们延续上一章节的想法，等权合成两个因子，记为 LBQ_Comb 。

3.3、因子表现

从日度 IC 测试结果上看，非同步交易性因子表现较好，其中以成交量占比构建的因子更为出色，Rank IC 均值为 5.28%，ICIR 大于 0.5，表现出较好的股价预测能力。

表 8、非同步交易性因子 Rank IC 测试结果

因子名称	IC 均值	IC 标准差	ICIR	T 统计量
rtn_LBQ	2.55%	5.05%	0.51	23.19
vol_LBQ	5.28%	10.35%	0.51	23.41
LBQ_Comb	4.86%	8.16%	0.60	27.33

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从日度组合测试结果上看，非同步交易性因子的表现十分优秀，多空夏普均较高，且无明显回撤，多头收益率较高，基于成交量序列构建的 LBQ 因子优于基于收益率构建的因子。具体来看，从多空组合测试上看， LBQ_Comb 的多空收益率在 49% 左右，多头收益率约为 23%，夏普比率在 6 左右。

表 9、非同步交易性因子组合测试结果

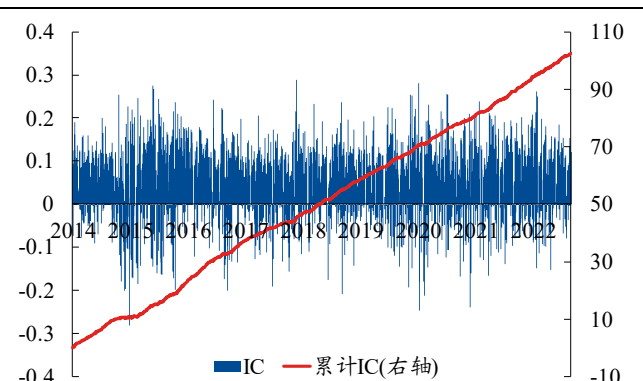
因子名称	多空收益率	多头收益率	空头收益率	年化波动率	夏普比率	最大回撤	多头换手率
rtn_LBQ	33.7%	16.8%	-13.3%	5.8%	5.78	-7.0%	17.3%
vol_LBQ	43.3%	23.8%	-15.4%	9.7%	4.45	-8.5%	9.8%
LBQ_Comb	49.2%	23.1%	-19.0%	8.6%	5.75	-6.7%	14.2%

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从多空净值曲线以及分位数组合测试结果上看， LBQ_Comb 因子的多空净值

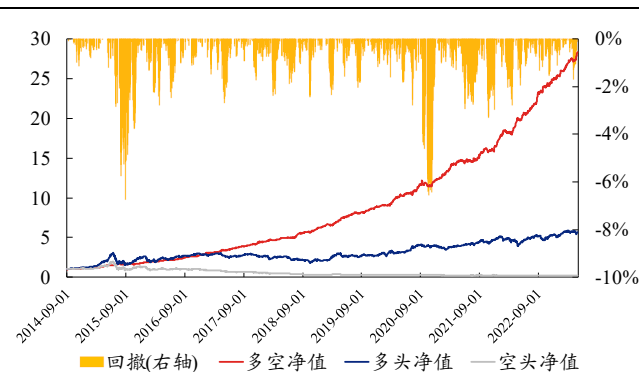
长期呈现上升趋势，且最近几年无明显回撤，表现十分稳定。

图 5、LBQ_Comb 因子 Rank IC 和累计 IC



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

图 6、LBQ_Comb 因子多空组合净值



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

4、条件分布基础因子

4.1、条件分布方法简述——以高波收益率均值为例

在前序章节中，我们从时序自相关性的角度出发构建了多类因子。在之后的篇幅中，我们将从条件分布的角度出发，从基础因子逐步延伸至两个特异因子。

长期以来，在大多数研究中，流动性资产的收益率分布均假设其服从正态分布。然而，其假设大多建立在市场为有效市场的背景之下。在非弱式有效市场的背景之下，市场和个股内外部条件的变化使得收益率的统计分布特征发生变化。对于日内收益率而言，其自身的非独立同分布特征更加明显（具体可参见报告《高频研究系列二——收益率分布因子构建》）。因此，我们在本章中引入条件分布，以此假设日内收益率在一定已知条件下服从某一分布。举例来说，我们可以假设收益率在一定条件下服从正态分布。

$$r_{tn} | Condition \sim N(\mu, \sigma^2) \quad (4)$$

因此，在条件分布的基础上，对于已知条件的刻画以及对于分布的假设便是提取时序信息构建相关因子的重要步骤。条件分布类因子本质上基于过去一段时间的数据特征，对高频样本数据进行分类、筛选或重构，并进一步刻画新构建的数据特征。在本章中，我们以已实现波动率作为条件，计算各个分钟节点过去 30 分钟的 5 分钟滚动收益率标准差，并筛选出标准差处于日内该股 80% 分位数以上的时间节点，并统计该时间节点中 5 分钟滚动收益率的均值，最终构建因子 `highStdRtn_mean`。

4.2、高波收益率均值因子表现与分析

细心的读者可能会发现，该因子与此前构建的常见收益率分布因子——5 分钟滚动收益率 `rtn5_mean` 因子相近。两者在构建方式上几乎一致，仅是高波收益率均值因子在样本选择上有所筛选，这也是条件分布的重点。我们对比发现，两者

的相关性仅 0.61，且与底层其他所有高频因子的相关性均低于 0.6。在测试中，我们也将该因子与 5 分钟滚动收益率因子进行正交化处理，记为 highStdRtn_meanN，并进行回测测试。经过正交化后的因子与底层所有高频因子的相关性最高仅 0.35。

从日度 IC 测试结果上看，highStdRtn_mean 因子表现较好，Rank IC 均值为 3.34%，ICIR 为 0.72，表现出较好的股价预测能力。经过 rtn5_mean 正交化后的因子保持着较好的选股能力，ICIR 为 0.41。

表 10、高波收益率均值因子 Rank IC 测试结果

因子名称	IC 均值	IC 标准差	ICIR	T 统计量
highStdRtn_mean	3.34%	4.65%	0.72	33.00
highStdRtn_meanN	1.76%	4.27%	0.41	18.84

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从日度组合测试结果上看，highStdRtn_mean 因子的表现同样优秀。具体来看，从多空组合测试上看，该因子的多空收益率在 41% 左右，多头收益率约为 14%，夏普比率在 7 左右。此外，经过 rtn5_mean 正交化后的因子仍保持着较好的选股能力，其多空的夏普比率约为 4.8。因此，我们认为：只要选择有效且适合的条件进行筛选，同样的统计量也能提取日内高频数据不同的数据特征，进而构造出具有一定特异性的因子。

表 11、高波收益率均值因子组合测试结果

因子名称	多空收益率	多头收益率	空头收益率	年化波动率	夏普比率	最大回撤	多头换手率
highStdRtn_mean	41.2%	14.3%	-19.4%	5.5%	7.48	-3.9%	20.3%
highStdRtn_meanN	20.7%	10.3%	-9.1%	4.4%	4.76	-4.2%	19.7%

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

5、考虑股价自相关性的风险度量因子

在上一章节中，我们简单介绍了条件分布下的因子构建逻辑，并以高波收益率均值因子作为样例，表明即使是相同的统计量，使用不同的筛选条件也能构造出具有一定特异性的因子。在本章中，我们引入股价的自相关性，通过考虑自相关来修正目前的风险度量指标 VaR，并由此构造在股价自相关性显著的条件下的 VaR 因子。

5.1、股价自相关性和 VaR 推导

无论是在学术界还是投资界，如何度量风险都是难以绕开的问题。VaR (Value at Risk) 便是其中一种最为常见的度量方式。其通常指在一定的假设以及给定的置信区间内，投资组合在既定的时期内可能遭受的最大价值损失。其核心功能是针对流动性资产的风险度量。在往期的报告中，我们就曾参考 VaR 的定义，构建了极值收益率分布因子，详见报告《高频研究系列三——收益率分布中的 Alpha(2)》。

然而，对于基础的 VaR 定义而言，其假设流动性资产的收益率在某一时间段内服从正态分布，并进一步假设流动性资产的价格波动是随机的，不存在自相关性。然而，在此前的报告中我们也提及，收益率在不同的时间区间内，尤其是在

日内极容易偏离正态分布。此外，在实际计算中我们也发现，对于日内股票价格序列而言，其价格的自相关性相对显著。因此，若以 VaR 作为股票日内价格波动的风险度量可能并不准确。在本章中，我们将假设价格序列存在自相关性，即股价波动并不随机的情况下，通过二元正态分布刻画日内的个股 VaR 值，并最终构建因子。

具体来说，我们首先假设股票价格服从以下分布

$$\ln P_{t+n} - \ln P_t \sim N \left[\left(\mu - \frac{\sigma^2}{2} \right) \cdot n, \sigma \cdot \sqrt{n} \right] \quad (5)$$

其中， P_{t+n} 为未来时刻的股票价格， P_t 为当前时刻的股票价格。我们设置 $n=1$ ，即价格时差为 1 分钟。由于需要考虑到分钟级股价存在自相关性，我们进一步假设其 $\ln P_{t+1}$ 与 $\ln P_t$ 之间的相关系数为 ρ 。

在此基础之上，我们假设 $(\ln P_{t+1}, \ln P_t)$ 服从二元正态分布，且两者具有一样的均值和标准差。由此，我们可得

$$(\ln P_{t+1}, \ln P_t) \sim N(\mu, \mu, \sigma^2, \sigma^2, \rho) \quad (6)$$

我们构建了前后对数价格的二元正态分布 $\mathbf{X} = \begin{pmatrix} \ln P_{t+1} \\ \ln P_t \end{pmatrix}$ ，其均值 $\boldsymbol{\mu} = \begin{pmatrix} \mu \\ \mu \end{pmatrix}$ ， $\boldsymbol{\Sigma} = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ 。

进一步，我们计算对数收益率 $R_t = \ln P_{t+1} - \ln P_t$ ，由此得到

$$\mathbf{Y} = \begin{pmatrix} R_t \\ \ln P_t \end{pmatrix} = \begin{pmatrix} \ln P_{t+1} - \ln P_t \\ \ln P_t \end{pmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \mathbf{X} = \mathbf{C}\mathbf{X} \quad (7)$$

其中 $\mathbf{C} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ 。由此，我们得到 \mathbf{Y} 的分布为 $N(\mathbf{C}\boldsymbol{\mu}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}')$ 。由此，我们在考虑了对数价格的前后自相关性后，得到了收益率与对数股票价格的二元正态分布。最后，在股票价格 $\ln P_t$ 已知的情况下，我们最终得到在考虑了股票价格自相关性下的 VaR。其中， ρ 为对数价格的前后分钟相关性， μ 为日内分钟级对数价格的均值， σ 为日内分钟级对数价格的标准差，同时我们假设 P_t 为当日收盘价， $Z_\alpha = 1.96$ 。

基于上述推导，我们可以知道：假设自相关的绝对值相同的情况下，自相关为正（趋势性更强）且收盘价小于当日均价的股票可能处于下行通道，风险较大；自相关为正（趋势性更强）但收盘价大于当日均价的股票可能处于上行通道，风险较小；反之，自相关为负（反转性更强）且收盘价小于当日均价的股票可能处于向上反弹通道，风险小；自相关为负（反转性更强）但收盘价大于当日均价的股票可能处于向下反弹通道，风险大。

表 12、股价相关性 VaR 公式解析

$\mu - \ln P_t > 0$		$\mu - \ln P_t < 0$
$\rho > 0, \quad \rho = a$	风险偏大	风险偏小
$\rho < 0, \quad \rho = a$	风险小	风险大

资料来源：兴业证券经济与金融研究院整理

计算得到个股每日的股价相关性 VaR 之后，我们取时序上 15 日标准差作为最终因子，叫做 rtn_condVaR。标准差越大，说明该股过去一段时间内的风险不稳定性越高，个股下行风险较高。

5.2、因子表现

从日度 IC 测试结果上看，股价相关性 VaR 因子表现较好，Rank IC 均值为 4.01%，ICIR 大于 0.5，表现出较好的股价预测能力。

表 13、股价相关性 VaR 因子 Rank IC 测试结果

因子名称	IC 均值	IC 标准差	ICIR	T 统计量
rtn_condVaR	4.01%	7.71%	0.52	23.87

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从日度组合测试结果上看，股价相关性 VaR 因子的表现十分优秀，因子的多空收益率在 44% 左右，多头收益率约为 21%，夏普比率在 5.7 左右。

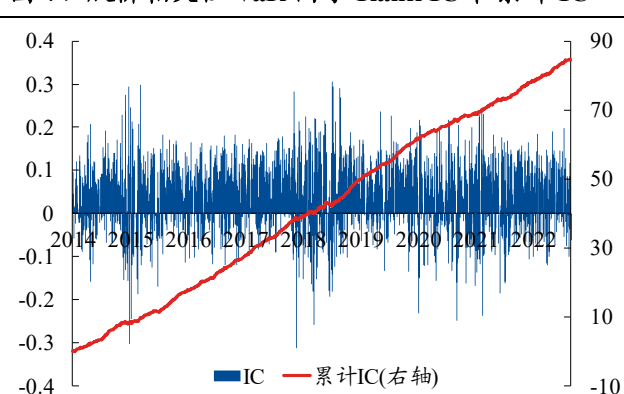
表 14、股价相关性 VaR 因子组合测试结果

因子名称	多空收益率	多头收益率	空头收益率	年化波动率	夏普比率	最大回撤	多头换手率
rtn_condVaR	44.1%	21.1%	-16.7%	7.8%	5.66	-7.6%	10.4%

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

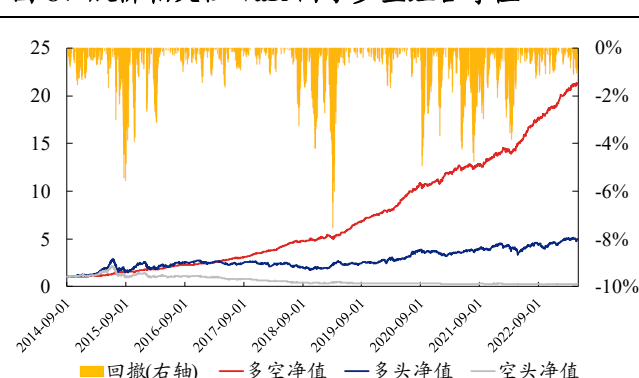
从多空净值曲线以及分位数组合测试结果上看，股价相关性 VaR 因子的多空净值长期呈现上升趋势，且最近几年无明显回撤，表现十分稳定。

图 7、股价相关性 VaR 因子 Rank IC 和累计 IC



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

图 8、股价相关性 VaR 因子多空组合净值



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

6、“闪电崩盘”、连续下跌和“崩盘概率”因子

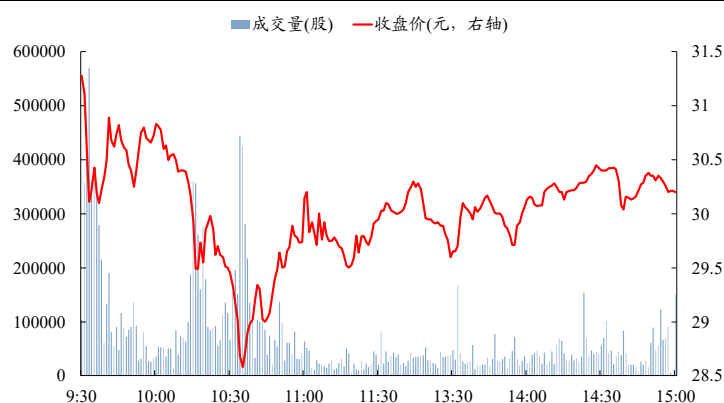
前序文章中，我们简单介绍了条件分布下的因子构建逻辑，并以高波收益率均值因子作为样例，表明即使是相同的统计量，使用不同的筛选条件也能构造出具有一定特异性的因子。此外，我们引入股价的自相关性并构造在股价自相关性显著的条件下的 VaR 因子。在本章中，我们从“闪电崩盘”的角度出发，通过观察“闪电崩盘”的特征，设置前一时点收益率的正负性异构得到个股日内连续上涨/下跌次数的样本数据，并进一步构建出“崩盘概率”因子。

6.1、“闪电崩盘”现象简析

闪电崩盘的概念在 2010 年起开始逐步受到重点关注。最为著名的例子便是 2010 年 5 月 6 日美股的“千点大跌”事件。在当日下午 2 点 42 分至 47 分，道琼斯指数下跌约 1000 点。现如今，对于当时出现崩盘的原因众说纷纭，但不可否认的是，闪电崩盘对投资者心理和交易行为带来的负面影响是极大且深远的。至此，闪电崩盘开始迅速进入市场投资者的眼中。

如今，闪电崩盘通常指流动性资产的价格在日内短时间出现大幅超跌，同时后续允许价格出现恢复，即在日内出现深“V”的价格走势。在 A 股市场中同样存在着这样的价格走势，尤其是在 2022 年以来市场出现大幅回调的时期。以 300659.SZ 2023 年 3 月 31 日的价格走势为例，当日该股开盘约 32 元，在 10 点半左右，该股收盘价跌至 28.58 元，相对于 9 点 30 分收盘价跌幅达 8.63%。此后股价攀升并持续震荡，最终价格收于 30.2 元，不难看出，当日便是相对明显的价格深“V”走势，这便是接近于“闪电崩盘”的一个交易日样例。

图 9、300659.SZ 日内价格走势



资料来源：上交所、深交所行情数据，兴业证券经济与金融研究院整理

注：数据日期为 2023 年 3 月 31 日

虽说“闪电崩盘”通常伴随着价格后期的恢复，然而，其出现时在日内造成的恐慌情绪则不可小觑。因此，本文将通过对日内连续下行收益率进行特征刻画，来衡量个股“崩盘”的概率，进而构造因子。

6.2、连续下跌和泊松分布下的“崩盘”概率因子

在上述对于闪电崩盘的刻画中我们发现：在日内出现闪电崩盘时，日内的分钟级别收益率大多会出现多次的连续负数。我们将该现象刻画为：分钟级别的连续下跌次数。以 2022 年 3 月 9 日的沪深 300 指数的价格走势为例，当日沪深 300 指数在 11:22 至 11:29 的 8 个分钟内，分钟级别收益率连续为负，跌幅为-0.5%。同时，日内多次出现类似的情况，如 9:53 至 9:57；10:35 至 10:39 等。我们统计了当日连续下跌次数的样本值，具体计算方式为：

1. 按照时间顺序逐个寻找：若一个分钟级收益率为负数，则开始统计连续分钟收益率为负数的个数，直至下一个收益率非负，则停止计数，该个数则为一个连续下跌次数的样本点；
2. 跳过之前被计数的分钟收益率样本点，继续寻找负数分钟收益率并按照步骤 1 计数，直至当日无未统计的分钟收益率。

我们假设日内分钟级收益率连续上涨和下跌次数（连续上涨的统计方式与下跌类似，只是将负数收益率改为正数收益率）出现的概率近似服从泊松分布，此外，一段时间内连续下跌趋势越明显，未来市场的波动率也越大。因此，若当日连续下跌次数相对较多，则未来价格在日内大概率会出现连续下跌的次数较多、跌幅较大、且波动率较大的特性。因此，我们可以基于日内的收益率样本数据，来估计未来连续下跌次数相对于连续上涨次数的大小，若连续下跌与连续上涨的差异较大，未来出现闪电崩盘的概率则较大。由此，我们的计算方式如下：

1. 首先基于前一个交易日的分钟级收益率序列，计算得到的连续上涨/下跌次数的样本数据，并进一步计算得到对于连续下跌泊松分布中参数 λ 的估计；
2. 计算个股连续下跌和连续上涨的差异： x 为全市场当日连续上涨 λ_t^{pos} 中位数， $x + k$ 为当日连续下跌 λ_t^{neg} 前 25%分位数；
3. 最终引入泊松分布的累计分布函数，计算得到“崩盘”概率。

最终，我们计算个股过去 15 日“崩盘”概率的标准差，记为 flashCrashProb 因子。标准差越大，说明该股下行风险较高，预期收益较低。

6.3、因子表现

从日度 IC 测试结果上看，“崩盘”概率因子表现良好，Rank IC 均值为 1.65%，ICIR 约为 0.3，表现出一定的股价预测能力。

表 15、“崩盘”概率因子 Rank IC 测试结果

因子名称	IC 均值	IC 标准差	ICIR	T 统计量
flashCrashProb	1.65%	5.45%	0.30	13.89

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从日度组合测试结果上看，“崩盘”概率因子的表现十分优秀，因子的多空收益率在 25%左右，多头收益率约为 13%，夏普比率在 5.2 左右。

请务必阅读正文之后的信息披露和重要声明

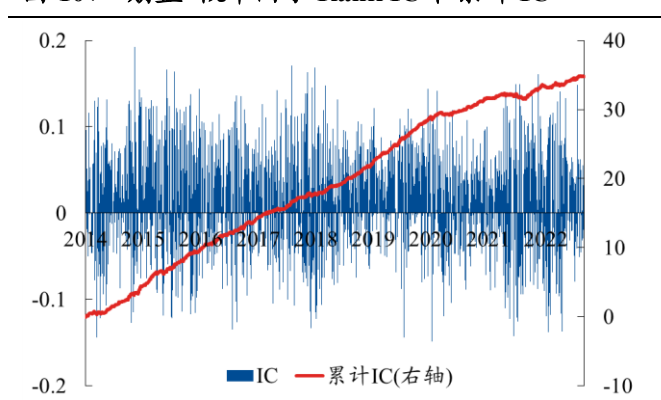
表 16、“崩盘”概率因子组合测试结果

因子名称	多空收益率	多头收益率	空头收益率	年化波动率	夏普比率	最大回撤	多头换手率
flashCrashProb	25.1%	13.3%	-9.1%	4.8%	5.21	-3.7%	11.1%

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

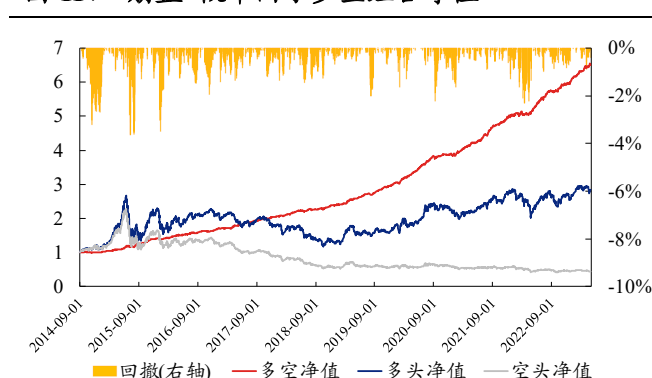
从多空净值曲线以及分位数组合测试结果上看，“崩盘”概率因子的多空净值长期呈现上升趋势，且最近几年无明显回撤，表现十分稳定。

图 10、“崩盘”概率因子 Rank IC 和累计 IC



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

图 11、“崩盘”概率因子多空组合净值



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

7、相关性检验与时序信息复合因子

7.1、时序信息因子正文化处理与相关性检验

由此，我们从时序相关性和条件分布两个角度，最终构建了 7 个时序信息因子，具体定义如下。在本节中，我们首先测试下述 7 个因子与兴证金工团队底层多个高频因子的时序相关性，并与底层相关性最高的因子进行正文化处理，保证正文化后的时序因子的最高相关性约在 0.6 以下。在测试中发现，“闪电崩盘”概率因子的相关性本身较低，最高相关性仅 0.4，不需要进行正文化处理。

表 17、时序信息因子定义

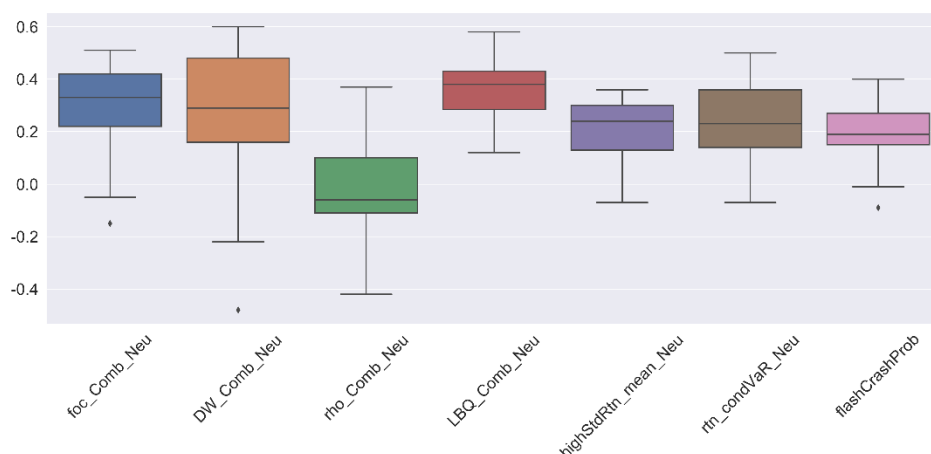
因子名称	因子定义
foc_Comb	收益率与成交量占比的一阶自相关系数
DW_Comb	收益率与成交量占比的 D-W 统计量
rho_Comb	收益率与成交量占比的残差自相关系数
LBQ_Comb	基于收益率与成交量占比刻画非同步交易程度
highStdRtn_mean	高波动率下的收益率均值
rtn_condVaR	存在股价相关性下的 VaR
flashCrashProb	“闪电崩盘”概率因子

资料来源：兴业证券经济与金融研究院整理

首先我们展示各个因子正文化后，与底层高频因子的相关性。从结果上看，经过正文化后各个因子的时序相关性均降低至 0.6 以下，其中残差自相关系数、高波收益率均值等因子的相关性较低，最高相关性在 0.5 以下。

请务必阅读正文之后的信息披露和重要声明

图 12、时序信息因子正交化后相关性箱型图



资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从日度 IC 测试结果上看，经过正交化后各个因子仍保持着较好的选股能力。其中，一阶自相关系数、非同步交易程度与 D-W 统计量因子的 Rank IC 均值较高，几乎都在 3% 以上。

表 18、正交化后的时序信息因子 Rank IC 测试结果

因子名称	IC 均值	IC 标准差	ICIR	T 统计量
foc_CombN	4.27%	6.93%	0.62	28.28
DW_CombN	2.98%	10.07%	0.30	13.59
rho_CombN	1.24%	6.51%	0.19	8.75
LBQ_CombN	3.04%	4.53%	0.67	30.79
highStdRtn_meanN	1.76%	4.27%	0.41	18.91
rtn_condVaRN	2.33%	5.63%	0.41	19.01
flashCrashProb	1.65%	5.45%	0.30	13.89

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从日度组合测试结果上看，“经过正交化后各个因子的组合测试结果优秀，几乎所有因子的多空夏普比率大于 3。其中，部分因子在正交化后的夏普比率和多空年化收益率保持优秀。以非同步交易程度 LBQ 因子为例，该因子的多空年化收益率为 34%，多头年化收益率约为 19%，多空夏普比率为 6.4。此外，一阶序列相关系数因子的多头年化收益率约为 27%。

表 19、正交化后的时序信息因子组合测试结果

因子名称	多空收益率	多头收益率	空头收益率	年化波动率	夏普比率	最大回撤	多头换手率
foc_CombN	43.7%	26.8%	-12.5%	7.0%	6.27	-7.2%	12.0%
DW_CombN	21.8%	17.1%	-5.8%	9.0%	2.42	-9.1%	8.9%
rho_CombN	22.8%	15.4%	-5.6%	5.9%	3.84	-4.8%	14.2%
LBQ_CombN	34.2%	18.6%	-12.0%	5.3%	6.40	-6.2%	17.4%
highStdRtn_meanN	20.7%	10.3%	-9.1%	4.4%	4.76	-4.2%	19.7%
rtn_condVaRN	19.6%	15.3%	-3.7%	5.6%	3.48	-7.9%	14.1%
flashCrashProb	25.1%	13.3%	-9.1%	4.8%	5.21	-3.7%	11.1%

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

7.2、时序信息复合因子构建与测试

最后，我们尝试基于本文最终推荐的七个高频时序因子，构建时序信息复合因子，并在多种股票池内进行周度和月度换仓的测试，以测试高频因子在中低频换仓和不同股票池内的表现。在合成之前，我们先测试这七个因子的相关性。从测试结果上看，各个因子的时序相关性较低，最高为 0.5。这从侧面反映出我们构建的因子在不同的角度提取了高频数据中的时序信息。

图 13、正交化后的时序信息因子间相关性

	foc_CombN	DW_CombN	rho_CombN	LBQ_CombN	highStdRtn_meanN	rtn_condVaRN	flashCrashProb
foc_CombN		0.50	0.16	0.37	0.36	0.14	0.10
DW_CombN	0.50		-0.23	0.42	0.26	-0.06	-0.07
rho_CombN	0.16	-0.23		0.04	0.15	0.37	0.33
LBQ_CombN	0.37	0.42	0.04		0.45	0.33	0.37
highStdRtn_meanN	0.36	0.26	0.15	0.45		0.25	0.34
rtn_condVaRN	0.14	-0.06	0.37	0.33	0.25		0.23
flashCrashProb	0.10	-0.07	0.33	0.37	0.34	0.23	

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

我们等权合成上述七个因子，记为时序信息复合因子。我们首先以周度换仓的形式，测试其在全市场、沪深 300、中证 500、中证 800、中证 1000 与国证 2000 里的表现。具体地，我们在各个股票池进行 Rank IC 和分位数组合测试。其中，在沪深 300、中证 500 与中证 800 中我们进行五分位组合测试，其余进行十分位组合测试。

首先展示时序信息复合因子周度调仓的 Rank IC 测试结果。从测试结果上看，该因子表现十分优秀，尤其是在中小盘股票池内。具体来说，该因子在全市场内周度 Rank IC 均值为 6.64%，ICIR 为 0.98。此外，因子在中证 1000 与国证 2000 内的表现十分出色，ICIR 接近或超过 1。

表 20、时序信息复合因子周度 Rank IC 测试结果

股票池	IC 均值	IC 标准差	ICIR	T 统计量
全市场	6.64%	6.75%	0.98	20.72
沪深 300	2.80%	12.40%	0.23	4.76
中证 500	5.21%	9.54%	0.55	11.49
中证 800	4.10%	10.04%	0.41	8.60
中证 1000	6.30%	7.16%	0.88	18.39
国证 2000	7.18%	6.80%	1.06	22.21

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从复合因子的周度分位数组合测试结果上看，该因子在不同的股票池内均呈现出明显的单调特征，且多头组合均明细战胜基准，除沪深 300 对应测试，其多空组合的夏普比率均大于 1。具体来看，该因子在全市场十分位组合测试中，多头组合年化收益率为 26.7%，同期基准为 10.8%，多空夏普比率为 4.91。此外，该因子在国证 2000 内的多空夏普比率为 5.35，在中证 1000 内的多空夏普比率为 4.18，因子在中小盘股票上的表现优秀。

图 14、时序信息复合因子周度分位数组合测试结果

股票池	组合	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率	年化超额收益率	股票池	组合	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率	年化超额收益率
沪深300	Top	12.0%	20.9%	0.57	63.6%	34.6%	3.6%	中证800	Top	14.3%	23.0%	0.62	65.2%	37.7%	5.3%
	1	10.4%	21.5%	0.49	114.7%	37.3%	2.4%		1	14.3%	23.6%	0.61	116.6%	36.6%	5.5%
	2	7.7%	22.3%	0.35	121.2%	44.8%	0.1%		2	10.0%	23.7%	0.42	122.9%	45.0%	1.6%
	3	8.9%	23.6%	0.38	112.5%	48.2%	1.5%		3	6.4%	24.7%	0.26	113.0%	51.9%	-1.6%
	Bottom	-0.5%	24.5%	-0.02	63.2%	63.0%	-7.2%		Bottom	-2.3%	26.1%	-0.09	65.3%	69.9%	-9.4%
	市场	7.7%	21.4%	0.36		45.7%			市场	8.5%	22.9%	0.37		49.1%	0.0%
中证500	L-S	10.8%	11.9%	0.91		36.9%		中证1000	L-S	15.4%	10.7%	1.44		25.0%	0.3%
	Top	15.9%	25.0%	0.64	68.5%	37.4%	6.5%		Top	20.5%	28.2%	0.73	72.0%	41.6%	11.4%
	1	15.3%	25.7%	0.59	119.4%	39.4%	6.3%		1	17.4%	29.4%	0.59	120.3%	45.0%	8.9%
	2	10.7%	26.1%	0.41	126.1%	46.3%	2.1%		2	12.7%	29.6%	0.43	126.7%	50.4%	4.6%
	3	7.7%	27.0%	0.28	117.0%	53.8%	-0.6%		3	3.8%	30.2%	0.12	118.1%	67.3%	-3.6%
	Bottom	-5.7%	28.1%	-0.20	70.9%	75.9%	-12.7%		Bottom	-2.0%	30.6%	-0.39	73.7%	84.6%	-18.3%
全市场	市场	8.8%	24.6%	0.36		51.6%	0.0%	国证2000	市场	8.4%	27.2%	0.31		56.2%	
	L-S	21.1%	11.0%	1.92		15.6%	4.2%		L-S	35.5%	8.5%	4.18		5.4%	
	Top	26.7%	26.1%	1.02	85.7%	35.9%	14.3%		Top	29.0%	28.0%	1.03	88.1%	38.2%	14.0%
	1	22.8%	26.9%	0.85	140.1%	36.6%	11.1%		1	27.1%	28.9%	0.94	142.0%	37.8%	12.6%
	2	22.1%	27.2%	0.81	153.4%	37.5%	10.5%		2	24.2%	29.4%	0.82	154.4%	43.2%	10.3%
	3	18.8%	27.6%	0.68	158.5%	39.3%	7.6%		3	20.9%	29.6%	0.71	159.8%	43.6%	7.4%
	5	17.5%	27.3%	0.64	159.6%	41.9%	6.3%		5	21.1%	29.5%	0.72	161.0%	43.2%	7.5%
	6	13.2%	27.8%	0.48	159.5%	49.9%	2.6%		6	14.8%	30.3%	0.49	160.8%	48.8%	2.1%
	7	9.1%	27.9%	0.32	156.8%	56.5%	-1.1%		7	9.9%	29.8%	0.33	157.9%	59.4%	-2.3%
	8	2.4%	28.7%	0.08	150.3%	67.4%	-7.0%		8	3.9%	30.6%	0.13	151.8%	66.5%	-7.6%
	9	-1.9%	28.7%	-0.07	135.8%	69.1%	-10.9%		9	-2.2%	30.8%	-0.07	137.5%	72.0%	-13.0%
	Bottom	-17.5%	29.3%	-0.60	86.2%	90.2%	-25.1%		Bottom	-17.4%	30.4%	-0.57	88.9%	90.2%	-26.6%
	市场	10.8%	25.8%	0.42		50.1%	0.0%		市场	13.3%	27.0%	0.49		50.4%	
	L-S	51.0%	10.4%	4.91		11.0%	26.7%		L-S	53.9%	10.1%	5.35		7.2%	

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

我们进一步展示时序信息复合因子月度调仓的 Rank IC 测试结果。从测试结果上看，该因子表现同样优秀，尤其是在中小盘股票池内。具体来说，该因子在全市场内月度 Rank IC 均值为 8.07%，ICIR 为 1.32。此外，因子在中证 1000 与国证 2000 内的表现十分出色，ICIR 均超过 1。

表 21、时序信息复合因子月度 Rank IC 测试结果

股票池	IC 均值	IC 标准差	ICIR	T 统计量
全市场	8.07%	6.12%	1.32	13.38
沪深 300	4.26%	11.38%	0.37	3.80
中证 500	6.13%	9.32%	0.66	6.67
中证 800	5.27%	9.34%	0.56	5.72
中证 1000	7.93%	6.82%	1.16	11.74
国证 2000	8.33%	5.99%	1.39	14.12

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

从复合因子的月度分位数组合测试结果上看，该因子在不同的股票池内均呈现出明显的单调特征，且多头组合均明细战胜基准，除沪深 300 对应测试，其多空组合的夏普比率均大于 1。具体来看，该因子在全市场十分位组合测试中，多头组合年化收益率为 17.6%，同期基准为 7.9%，多空夏普比率为 3.28。此外，该因子在国证 2000 内的多空夏普比率为 3.42，因子在中小盘股票上的表现优秀。

图 15、时序信息复合因子月度分位数组合测试结果

股票池	组合	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率	年化超额收益率	股票池	组合	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率	年化超额收益率
沪深300	Top	11.1%	22.9%	0.48	114.3%	34.3%	4.5%	中证800	Top	11.1%	23.8%	0.46	117.6%	36.1%	4.5%
	1	7.3%	21.7%	0.34	148.4%	41.7%	0.7%		1	9.1%	23.7%	0.38	149.9%	42.3%	2.6%
	2	6.3%	23.0%	0.27	151.5%	44.4%	0.2%		2	8.2%	24.4%	0.34	153.7%	44.2%	2.0%
	3	6.2%	24.5%	0.26	147.9%	48.0%	0.4%		3	4.9%	24.3%	0.20	149.5%	51.3%	-1.1%
	Bottom	0.7%	25.8%	0.03	114.7%	59.6%	-4.7%		Bottom	-1.6%	25.7%	-0.06	119.3%	66.7%	-7.1%
	市场	6.2%	22.7%	0.27		45.6%			市场	6.2%	23.7%	0.26		49.9%	
	L-S	8.9%	11.4%	0.78		24.5%			L-S	11.8%	9.7%	1.22		15.7%	
中证500	Top	10.9%	25.7%	0.43	122.3%	41.2%	4.6%	中证1000	Top	12.2%	29.8%	0.41	130.0%	45.4%	6.6%
	1	9.8%	25.8%	0.38	153.0%	39.0%	3.5%		1	12.8%	31.1%	0.41	154.1%	45.3%	7.7%
	2	7.6%	26.4%	0.29	156.5%	47.2%	1.6%		2	5.5%	30.2%	0.18	156.7%	54.3%	0.4%
	3	4.9%	25.4%	0.19	152.1%	54.1%	-1.1%		3	2.2%	31.2%	0.07	153.1%	65.6%	-2.4%
	Bottom	-2.7%	26.6%	-0.10	128.3%	72.1%	-8.1%		Bottom	-7.2%	31.1%	-0.23	135.0%	77.0%	-11.5%
	市场	6.1%	25.1%	0.24		52.7%			市场	5.2%	29.6%	0.18		58.4%	
	L-S	13.1%	9.7%	1.35		14.3%			L-S	19.8%	7.9%	2.50		4.2%	
全市场	Top	17.6%	28.1%	0.63	146.3%	34.4%	9.0%	国证2000	Top	19.0%	29.7%	0.64	149.4%	36.1%	8.2%
	1	14.8%	27.8%	0.53	170.9%	35.5%	6.4%		1	5.9%	30.7%	0.52	172.7%	38.7%	5.8%
	2	15.2%	28.5%	0.53	174.6%	37.6%	6.9%		2	17.4%	30.7%	0.57	176.1%	39.8%	7.2%
	3	12.7%	28.9%	0.44	177.2%	41.7%	4.8%		3	14.4%	31.2%	0.46	177.9%	43.2%	4.6%
	5	10.3%	28.6%	0.36	177.1%	46.8%	2.4%		5	1.1%	31.1%	0.36	178.0%	45.8%	1.5%
	6	9.1%	28.6%	0.32	177.5%	47.9%	1.3%		6	1.2%	30.9%	0.36	179.0%	46.2%	1.6%
	7	6.7%	29.0%	0.23	176.8%	56.7%	-0.8%		7	7.5%	31.5%	0.24	178.3%	55.0%	-1.6%
	8	4.1%	28.6%	0.14	174.7%	61.0%	-3.3%		8	5.5%	30.8%	0.18	176.5%	60.3%	-3.7%
	9	1.7%	29.1%	-0.06	171.1%	71.0%	-8.6%		9	0.2%	31.1%	-0.01	173.5%	68.1%	-8.9%
	Bottom	-9.7%	29.4%	-0.33	154.0%	79.3%	-16.1%		Bottom	-9.1%	30.9%	-0.29	158.4%	78.1%	-17.1%
	市场	7.9%	27.8%	0.28		51.9%	0.0%		市场	9.9%	29.6%	0.33		48.4%	
	L-S	28.9%	8.8%	3.28		3.3%	10.5%		L-S	29.5%	8.6%	3.42		3.7%	

资料来源：上交所、深交所行情数据，Wind，聚源，兴业证券经济与金融研究院整理

风险提示：模型结果基于历史数据的测算，在市场环境转变时模型存在失效的风险。

分析师声明

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登记为证券分析师，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了本人的研究观点。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收到任何形式的补偿。

投资评级说明

投资建议的评级标准	类别	评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后的12个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅。其中：A股市场以沪深300指数为基准；新三板市场以三板成指为基准；香港市场以恒生指数为基准；美国市场以标普500或纳斯达克综合指数为基准。	股票评级	买入	相对同期相关证券市场代表性指数涨幅大于15%
		增持	相对同期相关证券市场代表性指数涨幅在5%~15%之间
		中性	相对同期相关证券市场代表性指数涨幅在-5%~5%之间
		减持	相对同期相关证券市场代表性指数涨幅小于-5%
		无评级	由于我们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，致使我们无法给出明确的投资评级
	行业评级	推荐	相对表现优于同期相关证券市场代表性指数
		中性	相对表现与同期相关证券市场代表性指数持平
		回避	相对表现弱于同期相关证券市场代表性指数

信息披露

本公司在知晓的范围内履行信息披露义务。客户可登录 www.xyzq.com.cn 内幕交易防控栏内查询静默期安排和关联公司持股情况。

使用本研究报告的风险提示及法律声明

兴业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。

/，本公司不会因接收人收到本报告而视其为客户。

本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约，投资者自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效，任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以本公司向客户发布的本报告完整版本为准。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本报告所载资料的来源被认为是可靠的，但本公司不保证其准确性或完整性，也不保证所包含的信息和建议不会发生任何变更。本公司并不对使用本报告所包含的材料产生的任何直接或间接损失或与此相关的其他任何损失承担任何责任。

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，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌，过往表现不应作为日后的表现依据；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；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现。过往的业绩表现亦不应作为日后回报的预示。我们不承诺也不保证，任何所预示的回报会得以实现。分析中所做的回报预测可能是基于相应的假设。任何假设的变化可能会显著地影响所预测的回报。

本公司的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告并非针对或意图发送予或为任何就发送、发布、可得到或使用此报告而使兴业证券股份有限公司及其关联子公司等违反当地的法律或法规或可致使兴业证券股份有限公司受制于相关法律或法规的任何地区、国家或其他管辖区域的公民或居民，包括但不限于美国及美国公民（1934年美国《证券交易所》第15a-6条例定义为本「主要美国机构投资者」除外）。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本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。未经授权的转载，本公司不承担任何转载责任。

特别声明

在法律许可的情况下，兴业证券股份有限公司可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到兴业证券股份有限公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

兴业证券研究

上海	北京	深圳
地址：上海浦东新区长柳路36号兴业证券大厦15层	地址：北京市朝阳区建国门大街甲6号SK大厦32层01-08单元	地址：深圳市福田区皇岗路5001号深业上城T2座52楼
邮编：200135	邮编：100020	邮编：518035
邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn