

7月看好估值和量价，关注 GAT 模型

华泰研究

2021 年 7 月 09 日 | 中国内地

量化投资月报

研究员 林晓明
 SAC No. S0570516010001 linxiaoming@htsc.com
 SFC No. BPY421 +86-755-82080134

研究员 何康, PhD
 SAC No. S0570520080004 hegang@htsc.com
 SFC No. BRB318 +86-21-28972039

联系人 王晨宇
 SAC No. S0570119110038 wangchenyu@htsc.com
 +8602138476179

7月因子观点：看好估值和量价因子，关注图注意力网络选股模型

综合内生变量、外生变量、因子周期三重逻辑，7月看好估值因子和反转、波动率、换手率为代表的量价因子。华泰图注意力网络（GAT）选股模型2021年上半年样本外表现出色，历史上量价因子占优的市场环境下，GAT选股模型表现更好。内生变量视角短期看好估值、波动率、换手率因子，不看好成长、盈利、财务质量因子。外生变量视角短期看好反转、波动率、换手率因子，不看好财务质量、小市值因子。因子周期视角中长期看好大市值、估值、盈利、财务质量、波动率、换手率因子。

图注意力网络（GAT）模型 2021 年样本外表现出色

传统因子选股模型通常将股票视作相互独立的样本，但股票间显然存在复杂关联，GAT 的优势在于将股票间关系作为增量信息纳入预测模型。华泰 GAT 选股模型发布于研报《图神经网络选股与 Qlib 实践》（2021-02-21），采用 158 个量价因子对沪深 300 成分股进行日频选股和调仓，截至 6 月底模型样本外跟踪表现出色。2021 年上半年 GAT 模型相对基准沪深 300 年化超额收益为 26.4%，信息比率为 2.67；对照组 LightGBM 模型年化超额收益为 6.1%，信息比率为 0.52。从历史上看，量价因子占优的市场环境下，GAT 选股模型表现更好。

内生变量视角：短期看好估值/波动率/换手率，不看好成长/盈利/质量

内生变量视角基于因子动量、因子离散度、因子拥挤度三类指标，预测风格因子未来表现，因子动量反映因子过去一段时间内表现；因子离散度相当于因子估值；因子拥挤度反映因子是否处于拥挤状态。内生变量对大类风格因子 7 月表现预测结果如下：看好估值、波动率、换手率因子，不看好成长、盈利、财务质量因子。其中，估值、波动率、换手率因子在动量指标上占优；估值、小市值、波动率、换手率因子在离散度指标上占优，处于过度忽视状态；估值、波动率、换手率因子在拥挤度指标上占优。

外生变量视角：短期看好反转/波动率/换手率，不看好质量/小市值

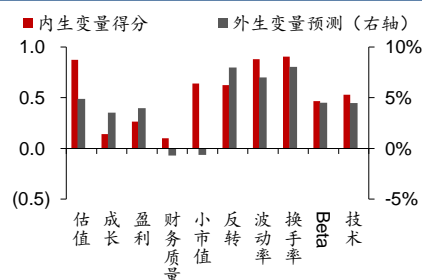
外生变量视角假设风格因子表现受市场变量、宏观变量等外生变量驱动，对每个大类风格因子筛选出具备预测能力的外生变量，使用线性回归预测未来一个月的 Rank IC 值。外生变量对大类风格因子 7 月 Rank IC 预测结果如下：反转、波动率、换手率因子 Rank IC 预测值超过 5%，可能较为有效，建议短期看多上述因子表现；不看好财务质量、小市值因子未来一个月表现。

因子周期视角：中长期看好大市值/估值/盈利/质量/波动率/换手率

因子周期视角将华泰金工经济周期研究成果应用于风格配置，自上而下构建经济系统周期运行状态与因子表现对应关系的“因子投资时钟”，根据当前周期状态预测因子未来表现的整体趋势。预计未来一年处于库兹涅茨周期向上，经济长期趋势向上的环境，同时朱格拉周期向下，流动性趋于宽松。建议配置大市值、估值、盈利、财务质量因子。另外，波动率、换手率因子在不同的宏观环境下表现都较好，这两个因子建议长期配置。

风险提示：风格因子的效果与宏观环境和大盘走势密切相关，历史结果不能预测未来。金融周期规律被打破。市场出现超预期波动，导致拥挤交易。

风格因子 7 月表现预测



资料来源：Wind，华泰研究预测

正文目录

7 月因子观点：看好估值和量价因子，关注 GAT 选股模型	4
图注意力网络（GAT）选股模型	5
GAT 选股模型 2021 年上半年样本外表现出色	5
GAT 选股模型测试流程	6
GAT 选股模型原理：量价因子结合图注意力网络	7
量价因子占优环境下，GAT 选股模型表现更好	8
因子表现回顾	9
6 月大类风格因子表现	9
近 1 年大类风格因子表现	9
近期细分因子表现	10
因子表现计算方法	10
内生变量视角	11
内生变量综合预测结果	11
因子动量	12
因子离散度	13
因子拥挤度	14
外生变量视角	15
外生变量选取	15
外生变量预测方法	15
条件期望模型	16
AIC 和 AICc 准则	16
外生变量筛选流程	16
回归预测模型构建	17
外生变量预测结果	17
因子预测结果解读	17
因子周期视角	20
自上而下预测整体趋势：因子投资时钟	20
自下而上预测边际变化：因子周期回归拟合	21
风险提示	22
附录	23
附录一：所有细分因子计算方法	23
附录二：风格因子合成、Rank IC 值计算、因子收益率计算方法	24

图表目录

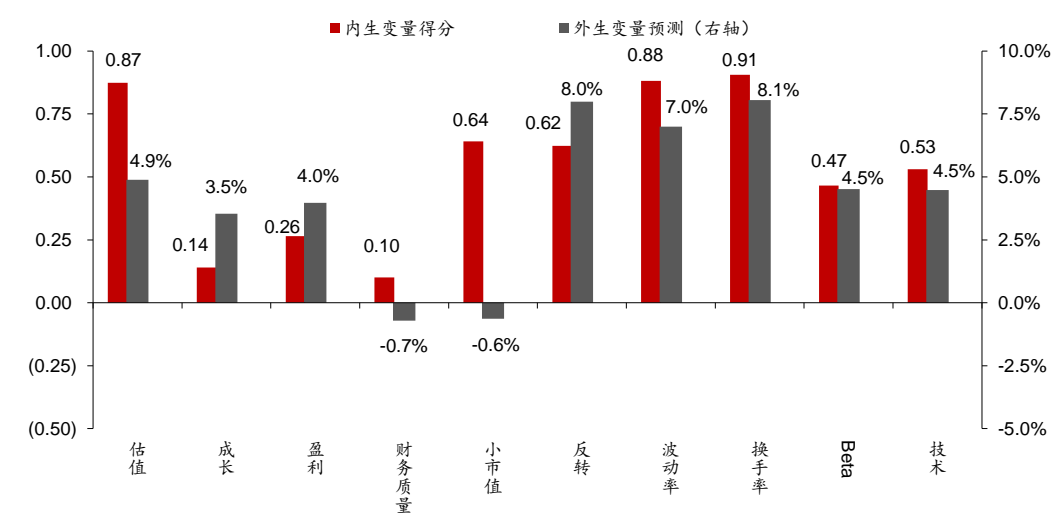
图表 1：内生变量和外生变量对风格因子 7 月表现的预测	4
图表 2：不同模型 TopkDropout 策略各年度回测指标（含交易费用，基准为沪深 300 指数，回测期 2010-01-04 至 2021-06-30）	5
图表 3：GAT 模型回测净值和超额净值表现（全回测期）	5
图表 4：GAT 模型回测净值和超额净值表现（2021 年上半年）	5

图表 5: GAT 模型参数设置	6
图表 6: GAT 原理示意图	7
图表 7: GAT 选股模型示意图	8
图表 8: GAT 模型和波动率因子 Rank IC 值相关性 (2019~2021 年)	8
图表 9: GAT 模型和换手率因子 Rank IC 值相关性 (2019~2021 年)	8
图表 10: 6 月大类风格因子表现	9
图表 11: 近 1 年大类风格因子累计 Rank IC 值	9
图表 12: 近 1 年细分因子表现	10
图表 13: 内生变量预测大类风格因子 7 月表现综合预测得分	11
图表 14: 大类风格因子动量历史序列	12
图表 15: 大类风格因子 6 月底截面动量和 2007 年以来历史分位数	12
图表 16: 大类风格因子离散度历史序列	13
图表 17: 大类风格因子 6 月底截面离散度和 2007 年以来历史分位数	13
图表 18: 大类风格因子拥挤度历史序列	14
图表 19: 大类风格因子 6 月底截面拥挤度和 2007 年以来历史分位数	14
图表 20: 外生变量列表	15
图表 21: 外生变量筛选流程图	16
图表 22: 外生变量对大类风格因子 7 月 Rank IC 预测值	17
图表 23: 估值因子 Rank IC 预测值及外生变量贡献情况	17
图表 24: 成长因子 Rank IC 预测值及外生变量贡献情况	17
图表 25: 盈利因子 Rank IC 预测值及外生变量贡献情况	18
图表 26: 财务质量因子 Rank IC 预测值及外生变量贡献情况	18
图表 27: 小市值因子 Rank IC 预测值及外生变量贡献情况	18
图表 28: 反转因子 Rank IC 预测值及外生变量贡献情况	18
图表 29: 波动率因子 Rank IC 预测值及外生变量贡献情况	18
图表 30: 换手率因子 Rank IC 预测值及外生变量贡献情况	18
图表 31: Beta 因子 Rank IC 预测值及外生变量贡献情况	19
图表 32: 技术因子 Rank IC 预测值及外生变量贡献情况	19
图表 33: 不同经济周期下的因子表现分类	20
图表 34: 实体经济三周期滤波结果	21
图表 35: 估值因子去趋势后累计收益率与三周期回归拟合	21
图表 36: 成长因子去趋势后累计收益率与三周期回归拟合	21
图表 37: 盈利因子去趋势后累计收益率与三周期回归拟合	21
图表 38: 财务质量因子去趋势后累计收益率与三周期回归拟合	21
图表 39: 小市值因子去趋势后累计收益率与三周期回归拟合	22
图表 40: 反转因子去趋势后累计收益率与三周期回归拟合	22
图表 41: 波动率因子去趋势后累计收益率与三周期回归拟合	22
图表 42: 换手率因子去趋势后累计收益率与三周期回归拟合	22
图表 43: Beta 因子去趋势后累计收益率与三周期回归拟合	22
图表 44: 技术因子去趋势后累计收益率与三周期回归拟合	22
图表 45: 报告中涉及的所有细分因子及其计算方式	23

7 月因子观点：看好估值和量价因子，关注 GAT 选股模型

综合内生变量、外生变量、因子周期三重逻辑，7 月看好估值因子和反转、波动率、换手率为代表的量价因子。华泰图注意力网络（GAT）选股模型 2021 年上半年样本外表现出色，历史上量价因子占优的市场环境下，GAT 选股模型表现更好。内生变量视角短期看好估值、波动率、换手率因子，不看好成长、盈利、财务质量因子。外生变量视角短期看好反转、波动率、换手率因子，不看好财务质量、小市值因子。因子周期视角中长期看好大市值、估值、盈利、财务质量、波动率、换手率因子。

图表1：内生变量和外生变量对风格因子 7 月表现的预测



资料来源：Wind，华泰研究预测

内生变量视角基于因子动量、因子离散度、因子拥挤度三类指标，预测风格因子未来表现，因子动量反映因子过去一段时间内表现；因子离散度相当于因子估值；因子拥挤度反映因子是否处于拥挤状态。内生变量对大类风格因子 7 月表现预测结果如下：看好估值、波动率、换手率因子，不看好成长、盈利、财务质量因子。其中，估值、波动率、换手率因子在动量指标上占优；估值、小市值、波动率、换手率因子在离散度指标上占优，处于过度忽视状态；估值、波动率、换手率因子在拥挤度指标上占优。

外生变量视角假设风格因子表现受市场变量、宏观变量等外生变量驱动，对每个大类风格因子筛选出具备预测能力的外生变量，使用线性回归预测未来一个月的 Rank IC 值。外生变量对大类风格因子 7 月 Rank IC 预测结果如下：反转、波动率、换手率因子 Rank IC 预测值超过 5%，可能较为有效，建议短期看多上述因子表现；不看好财务质量、小市值因子未来一个月表现。

因子周期视角将华泰金工经济周期研究成果应用于风格配置，自上而下构建经济系统周期运行状态与因子表现对应关系的“因子投资时钟”，根据当前周期状态预测因子未来表现的整体趋势。预计未来一年处于库兹涅茨周期向上，经济长期趋势向上的环境，同时朱格拉周期向下，流动性趋于宽松。建议配置大市值、估值、盈利、财务质量因子。另外，波动率、换手率因子在不同的宏观环境下表现都较好，这两个因子建议长期配置。

图注意力网络（GAT）选股模型

GAT 选股模型 2021 年上半年样本外表现出色

华泰金工《人工智能 42：图神经网络选股与 Qlib 实践》（2021-02-21）研报中，我们介绍图神经网络及图注意力网络（GAT）概念，并通过微软 Qlib 平台测试选股效果。传统因子选股模型中，通常将股票视作相互独立的样本，但股票间显然存在复杂关联，如产业链上下游关系、相关行业主题等。GAT 的优势在于能将股票间关系作为增量信息纳入预测模型。

微软 AI 量化投资开源平台 Qlib 已实现动态图注意力网络（GATs_ts），其本质是时序神经网络 LSTM 和图注意力网络 GAT 的融合。前述研究中，我们测试 GATs_ts 方法（以下简称 GAT）在沪深 300 成分股量价因子日频选股上的表现。组合构建采用 TopKDropout 策略，每日等权持有 50 只股票，每日根据预测打分更换 5 只股票。回测期内（2010 年至 2021 年 2 月初）相比对照组 LightGBM 模型具备显著优势。

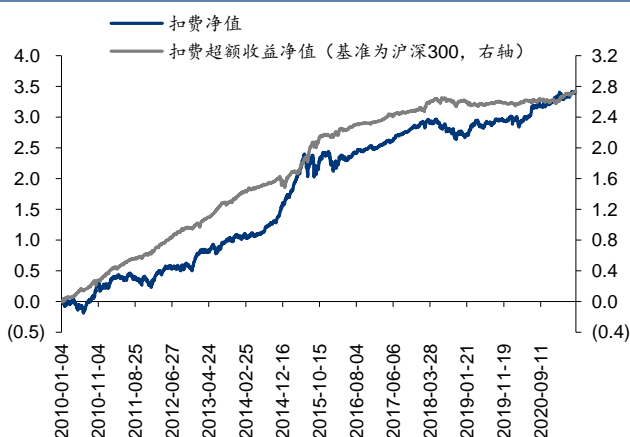
本研究中，我们将模型回测更新至 2021 年 6 月 30 日（其中 2021 年 2 月~6 月为样本外），GAT 选股模型样本外表现出色。相关回测指标如下表所示。2021 年上半年 GAT 模型相对基准沪深 300 指数年化超额收益为 26.4%，信息比率为 2.67，超额收益最大回撤为 -3.2%；对照组 LightGBM 模型年化超额收益为 6.1%，信息比率为 0.52，超额收益最大回撤为 -6.2%。

图表2：不同模型 TopkDropout 策略各年度回测指标（含交易费用，基准为沪深 300 指数，回测期 2010-01-04 至 2021-06-30）

模型	指标	2010-2021	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
GAT	年化收益率 (%)	30.8%	29.9%	-1.7%	51.3%	37.2%	61.6%	62.0%	22.5%	45.8%	-17.5%	40.4%	27.1%	26.8%
	年化波动率 (%)	24.7%	27.6%	22.2%	23.9%	22.4%	17.4%	43.3%	27.9%	11.4%	25.3%	19.8%	23.5%	23.0%
	夏普比率	1.25	1.08	(0.07)	2.14	1.66	3.55	1.43	0.81	4.03	(0.69)	2.04	1.15	1.17
	最大回撤 (%)	-37.4%	-22.0%	-21.0%	-12.3%	-15.4%	-8.1%	-40.1%	-21.8%	-3.7%	-28.3%	-14.6%	-19.3%	-16.0%
	年化超额收益率 (%)	24.6%	37.4%	30.0%	39.5%	43.6%	13.4%	58.9%	24.1%	24.5%	13.4%	3.8%	-0.2%	26.4%
	超额收益年化波动率 (%)	9.4%	8.5%	6.2%	7.5%	8.7%	11.2%	16.2%	10.6%	6.9%	9.7%	7.2%	8.2%	9.8%
	信息比率	2.62	4.42	4.86	5.25	5.01	1.20	3.63	2.27	3.53	1.37	0.52	(0.03)	2.67
	超额收益最大回撤 (%)	-14.3%	-5.9%	-3.0%	-2.3%	-6.4%	-14.6%	-9.0%	-4.4%	-3.3%	-9.1%	-6.3%	-6.2%	-3.2%
LightGBM	年化收益率 (%)	28.5%	21.6%	-14.1%	36.3%	29.5%	72.1%	55.5%	19.0%	46.9%	-14.1%	45.6%	31.3%	6.5%
	年化波动率 (%)	26.0%	27.9%	22.0%	23.2%	23.1%	19.3%	47.3%	28.2%	10.6%	24.4%	20.1%	24.8%	24.3%
	夏普比率	1.09	0.78	(0.64)	1.57	1.28	3.75	1.17	0.67	4.42	(0.58)	2.27	1.26	0.27
	最大回撤 (%)	-50.0%	-28.3%	-25.5%	-13.8%	-13.1%	-8.9%	-50.1%	-23.8%	-4.2%	-25.1%	-12.9%	-18.5%	-18.2%
	年化超额收益率 (%)	22.3%	29.2%	17.6%	24.6%	35.9%	23.9%	52.4%	20.6%	25.6%	16.8%	9.0%	4.0%	6.1%
	超额收益年化波动率 (%)	9.3%	8.5%	6.0%	6.6%	9.3%	8.5%	16.8%	10.2%	5.4%	8.9%	6.9%	8.6%	11.6%
	信息比率	2.40	3.43	2.93	3.69	3.84	2.81	3.11	2.02	4.74	1.87	1.31	0.47	0.52
	超额收益最大回撤 (%)	-10.5%	-7.0%	-4.2%	-2.1%	-8.3%	-6.8%	-10.5%	-4.9%	-2.1%	-8.5%	-3.2%	-4.6%	-6.2%

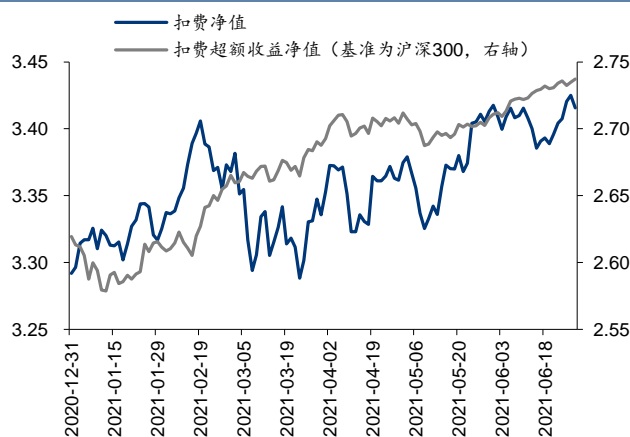
资料来源：Qlib, Wind, 华泰研究

图表3：GAT 模型回测净值和超额净值表现（全回测期）



资料来源：Qlib, Wind, 华泰研究

图表4：GAT 模型回测净值和超额净值表现（2021 年上半年）



资料来源：Qlib, Wind, 华泰研究

全回测期和 2021 年上半年策略净值和超额收益净值如上图所示。GAT 模型在 2019 年和 2020 年超额收益相对较弱，但是在样本外的 2021 年上半年 2~6 月表现出色。模型在 2019 年和 2020 年表现不佳，原因可能为：1) 所使用的 Alpha158 量价因子在沪深 300 成分股票池出现衰减；2) 策略为 50 只股票等权，相对于基准沪深 300 存在小市值暴露，若市场表现为少数权重股领涨的结构性行情，对等权策略不利。2021 年以来抱团逐渐瓦解，模型取得不错的超额收益。

GAT 选股模型测试流程

GAT 选股模型测试流程如下：

1. 数据获取
 - 1) 获取 Wind 中的 A 股数据，按照华泰金工《人工智能 40：微软 AI 量化投资平台 Qlib 体验》(2020-12-22) 中 dump_all 转换用户数据格式的方式，转换为 Qlib 的 bin 数据存储格式。
 - 2) 股票池：沪深 300 成分股。
 - 3) 回测区间：2010-01-04 至 2021-06-30。
2. 特征和标签提取：特征采用 Qlib 内置的因子库 Alpha158vwap 中的 158 个因子特征，将标签定义为 t+2 日 vwap 复权均价相对于 t+1 日 vwap 复权均价的涨跌幅，相当于 t 日收盘后发信号，t+1 日以日内均价开仓，t+2 日以日内均价平仓。
3. 特征预处理：
 - 1) 训练集和验证集：首先剔除标签为缺失值的样本 (Qlib 的 DropnaLabel 类)，再对标签进行截面标准化 (Qlib 的 CSRrankNorm 类)，即对每个截面的标签先转换为 rank 序数，最后 Z 分数标准化至标准正态分布。
 - 2) 测试集：首先对特征进行标准化 (Qlib 的 RobustZScoreNorm 类)，即对因子做稳健 Z 分数标准化，对原始数据减去中位数除以 1.48 倍 MAD 统计量，再将因子特征取值限制在 -3 到 3 之间 (clip_outlier 设置为 True)，最后将因子缺失值填充为 0 (Qlib 的 Fillna 类)。
4. 数据集设置：训练集采用样本内数据的前五年，验证集采用样本内数据的最后一年，测试集采用接下来的样本外一年。例如以 2005~2009 年为训练集，2010 年为验证集，2011 年为测试集，每年重新训练一次模型。

图表5： GAT 模型参数设置

参数	含义	参数取值
step_len	时间序列长度	20
d_feat	原始特征维度	158
hidden_size	隐藏层神经元数	64
num_layers	隐藏层层数	2
dropout	神经元被临时删除的概率	0.7
n_epochs	训练轮数	200
lr	学习率	0.0001
optimizer	优化器	adam
early_stop	提前结束轮数	10
metric	提前结束训练的度量方法	loss
loss	损失计算方式	mse
base_model	循环神经网络类型	LSTM
with_pretrain	是否导入预训练模型	False
GPU	用来训练的 GPU 编号	0
num_layers_gat	GATs 层数	1

资料来源：Qlib，华泰研究

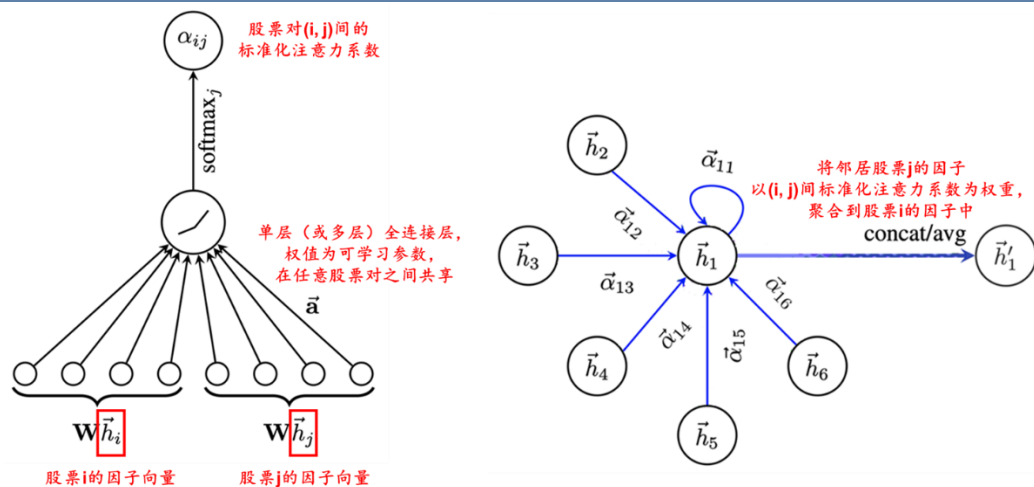
5. 样本内训练：使用自定义 `pytorch_gats_ts_layers.py` 的 GATs 类训练。将回测区间按年份划分为 12 个子区间，因此需要配置相应的训练集区间进行滚动训练。
6. 验证集调参：训练过程中，当验证集上的评价分数在连续 10 轮迭代后都没有提升时，停止模型训练，选取验证集评价分数最高的一组参数作为模型的最优参数，用来对测试集进行预测。这里的评价分数使用 Qlib 中默认的方式，即均方误差的相反数。
7. 样本外测试和回测：
 - 1) 得到最优参数后，进行模型预测得到测试集中每一天股票的预测收益率。
 - 2) 使用 TopkDropout 策略，每日持有 `topk=50` 只股票，同时每日卖出持仓股票中最新预测收益最低的 `n_drop=5` 只股票，买入未持仓股票中最新预测收益最高的 `n_drop=5` 只股票。买入每只股票金额为 95% 的剩余现金除以需购买的股票数量，95% 是可调参数，用来控制仓位。开仓交易费率为 0.05%，平仓交易费率为 0.15%。

GAT 选股模型原理：量价因子结合图注意力网络

图神经网络的核心思想是对样本间关系进行建图，每个节点对应一条样本，再将邻居节点的特征聚合到中心节点，从而更新节点特征。图注意力网络 GAT 是图神经网络中的代表性方法，采用全局注意力机制，无需对股票市场进行显式建图，而是隐式学习所有节点对中心节点的影响。

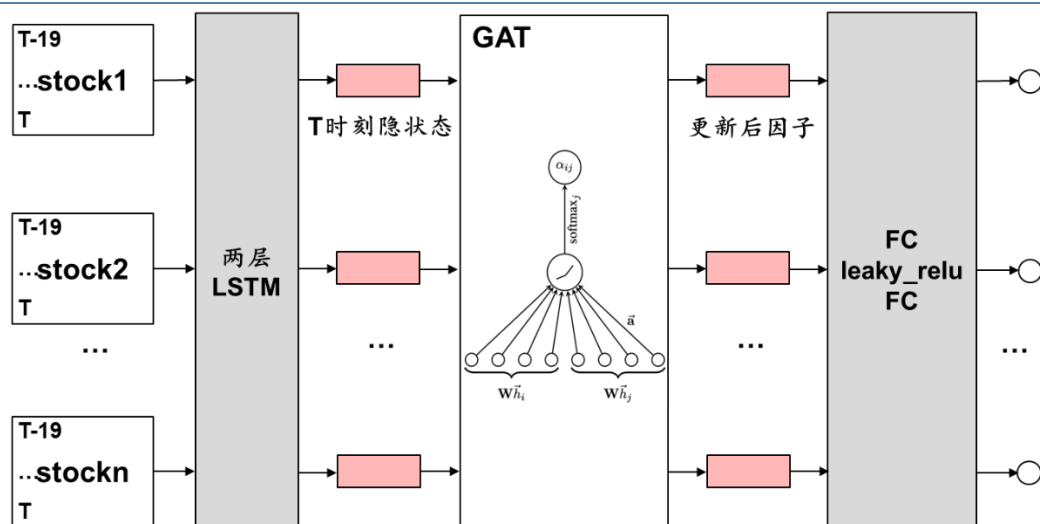
具体而言，首先对任意两只股票 i 和 j 的因子值向量与可学习的注意力参数向量计算内积，得到两只股票间的注意力系数；随后以注意力系数为权重，将邻居股票 j 的因子聚合到股票 i 的因子中，从而更新股票 i 的因子；最后将更新的因子值送入后续网络层，预测股票收益，将预测损失进行反向传播，实现网络参数的优化。

图表6： GAT 原理示意图



资料来源：华泰研究

图表7: GAT 选股模型示意图



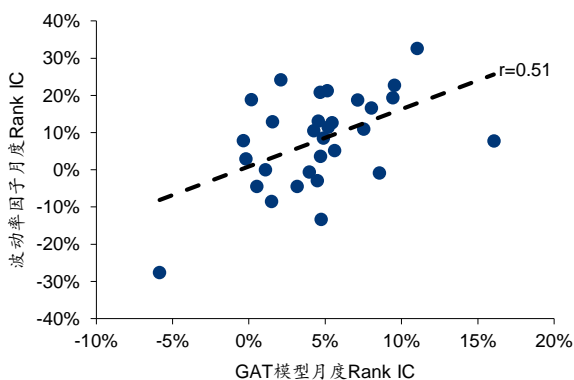
资料来源: Qlib, 华泰研究

Qlib 中的 GATs_ts (简称 GAT) 模型有两层 LSTM 层、GAT 层、两层全连接层构成。模型输入为每只股票 T-19~T 日共 20 个交易日的 158 个量价因子;通过 LSTM 层提取时序信息,得到 T 日 64 个隐状态因子;再通过 GAT 层提取股票间关系信息,得到更新后的 64 个因子;最终通过全连接层,模型输出为对股票收益的预测。

量价因子占优环境下, GAT 选股模型表现更好

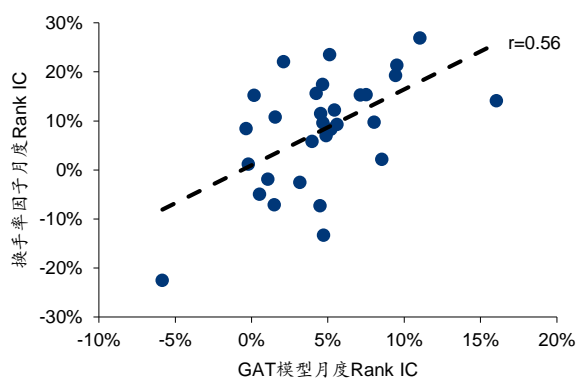
由于 GAT 模型使用的选股因子为 Qlib 预定义的 158 个量价因子, GAT 模型表现应与量价因子整体表现具有相关性。我们统计 2019 年至 2021 年上半年 GAT 模型月度 Rank IC 值与波动率、换手率因子月度 Rank IC 值的相关性,相关系数为 0.51 和 0.56。预计在量价因子占优的市场环境下, GAT 选股模型表现更好。

图表8: GAT 模型和波动率因子 Rank IC 值相关性 (2019~2021 年)



资料来源: Qlib, Wind, 华泰研究

图表9: GAT 模型和换手率因子 Rank IC 值相关性 (2019~2021 年)



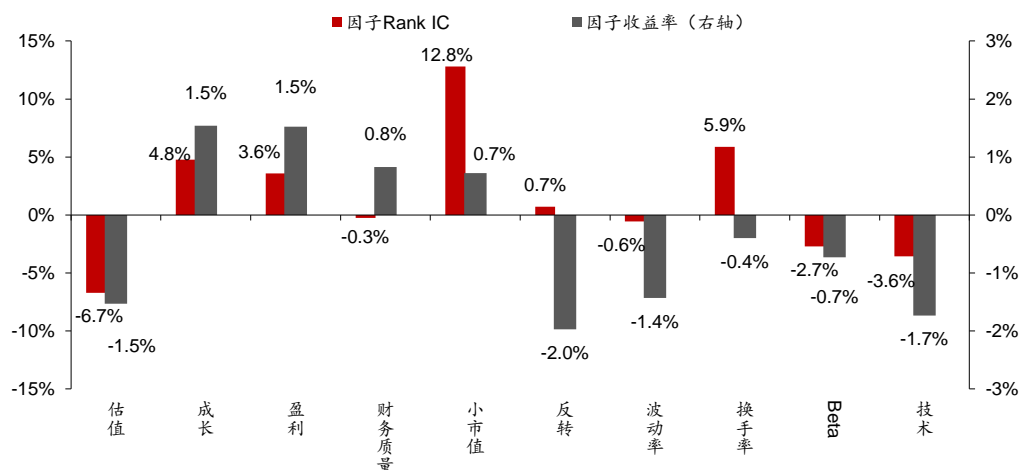
资料来源: Qlib, Wind, 华泰研究

因子表现回顾

6 月大类风格因子表现

大类风格因子 6 月在全 A 股票池的 Rank IC 值和因子收益率如下图所示。成长、盈利、小市值因子表现较好，其中小市值因子 Rank IC 值超过 10%。估值、波动率、Beta、技术因子出现回撤。反转、换手率因子 Rank IC 值为正，但是因子收益率为负。财务质量因子 Rank IC 值为负，但是因子收益率为正。

图表10： 6 月大类风格因子表现

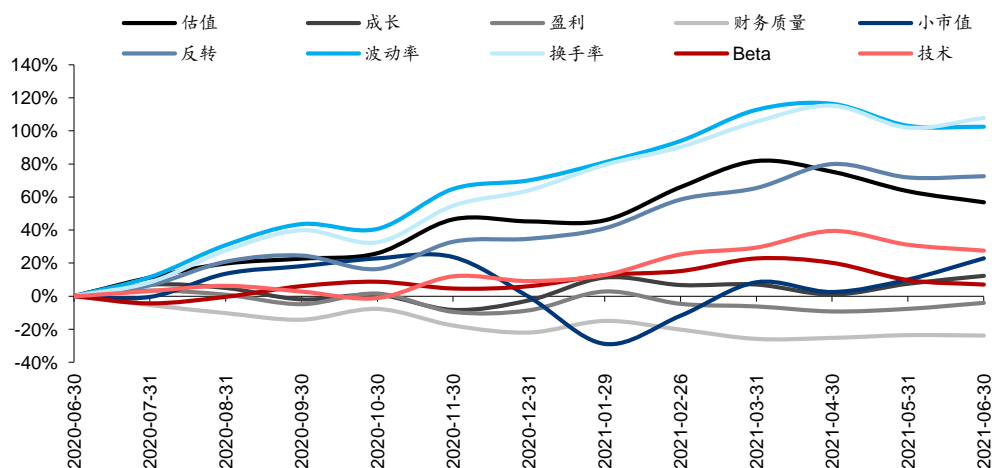


资料来源：Wind，华泰研究

近 1 年大类风格因子表现

对大类风格因子近 12 个月的 Rank IC 值做累加，得到各月累计 Rank IC 值，如下图所示。总的来看，波动率、换手率、反转、估值因子近 1 年表现良好，但近两个月表现相对平淡；技术因子表现尚可；成长、盈利、财务质量、Beta 因子表现相对平淡。小市值因子波动较大，2020 年 12 月及 2021 年 1 月出现较大回撤，随后止跌反弹。

图表11： 近 1 年大类风格因子累计 Rank IC 值

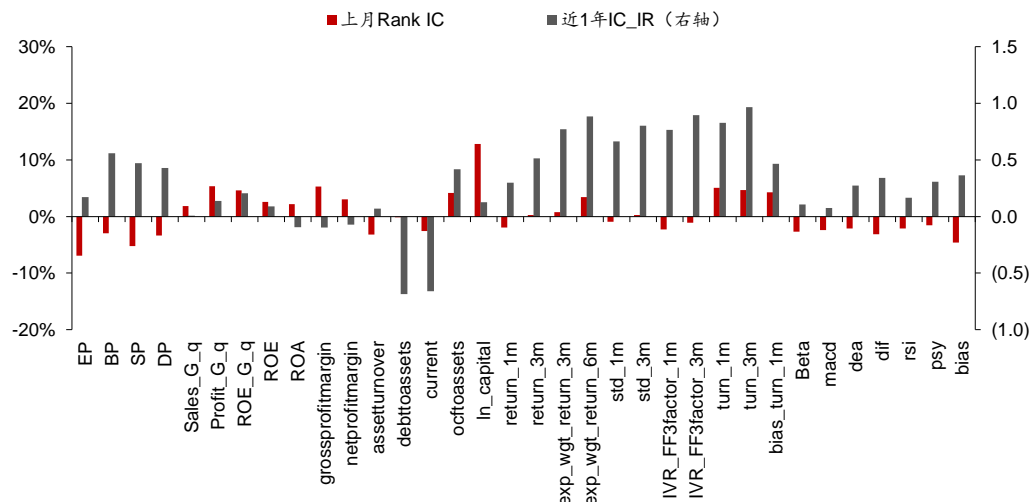


资料来源：Wind，华泰研究

近期细分因子表现

计算 34 个细分因子 6 月 Rank IC 值, 以及近 12 个月的月频 Rank IC 值。将近 12 个月 Rank IC 的均值除以标准差, 得到近 1 年 IC_IR。近期细分因子表现如下图所示。

图表12: 近 1 年细分因子表现



资料来源: Wind, 华泰研究

因子表现计算方法

因子 T 月表现的计算方法可以简要描述为:

1. 以 T-1 月最后一个交易日为截面期, 以全 A 股为股票池, 计算附录一中 10 大类共 34 个细分因子值。对每一大类风格因子内部各细分因子做去极值、标准化预处理及因子方向调整, 再等权合成, 得到截面期大类风格因子值。
2. 计算 T 月股票池内个股区间收益。
3. 对大类风格因子或细分因子做去极值、标准化和行业市值中性化预处理, 计算 T-1 月预处理后因子与 T 月个股区间收益的 Spearman 秩相关系数, 即 T 月因子 Rank IC 值。
4. 对大类风格因子值、个股总市值做去极值和标准化预处理, 以 T-1 月预处理后大类风格因子值、预处理后个股总市值、中信一级行业哑变量为自变量, 以 T 月个股区间收益为因变量, 以流通市值平方根为权重, 做加权线性回归, 大类风格因子值对应的回归系数即为 T 月大类风格因子收益率。

详细计算方法请参见附录二。

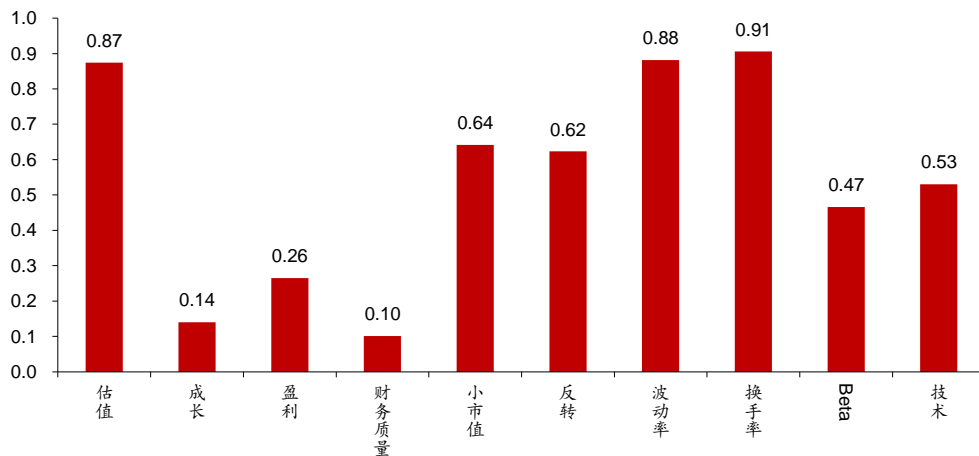
内生变量视角

内生变量视角基于因子动量、因子离散度、因子拥挤度三类指标，预测风格因子未来表现。因子动量反映因子过去一段时间内表现；因子离散度相当于因子估值，高因子离散度反映因子被过度忽视，低因子离散度反映因子被过度投资；因子拥挤度反映因子是否处于拥挤状态。一般看好高因子动量、高因子离散度、低因子拥挤度因子短期表现，三项指标在截面和时序上均有预测效果。

内生变量综合预测结果

基于因子动量、因子离散度、因子拥挤度三类指标，分别从截面和时序角度，对大类风格因子7月表现进行打分预测，综合预测结果如下表所示。**看好估值、波动率、换手率因子，不看好成长、盈利、财务质量因子。**其中，估值、波动率、换手率因子在动量指标上占优；估值、小市值、波动率、换手率因子在离散度指标上占优，处于过度忽视状态；估值、波动率、换手率因子在拥挤度指标上占优。

图表13： 内生变量预测大类风格因子7月表现综合预测得分



资料来源：Wind，华泰研究预测

综合预测得分是以下6项细分指标的均值：

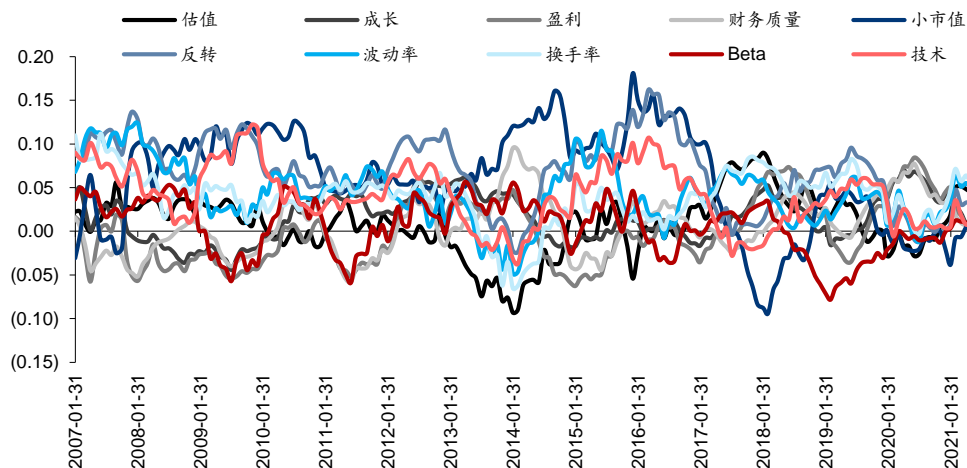
1. 归一化截面因子动量；
2. 因子动量历史分位数；
3. 归一化截面因子离散度；
4. 因子离散度历史分位数；
5. 1-归一化截面因子拥挤度；
6. 1-因子拥挤度历史分位数。

其中，归一化采用 min-max 归一化，根据截面上各类因子相应指标最小值和最大值，将其缩放至[0, 1]范围内；历史分位数取每个因子最新月末截面期相应指标，在其历史上（自2007年1月末截面期开始）所处分位数，本身在[0, 1]范围内；因子拥挤度为反向指标，故用1减去相应指标。

因子动量

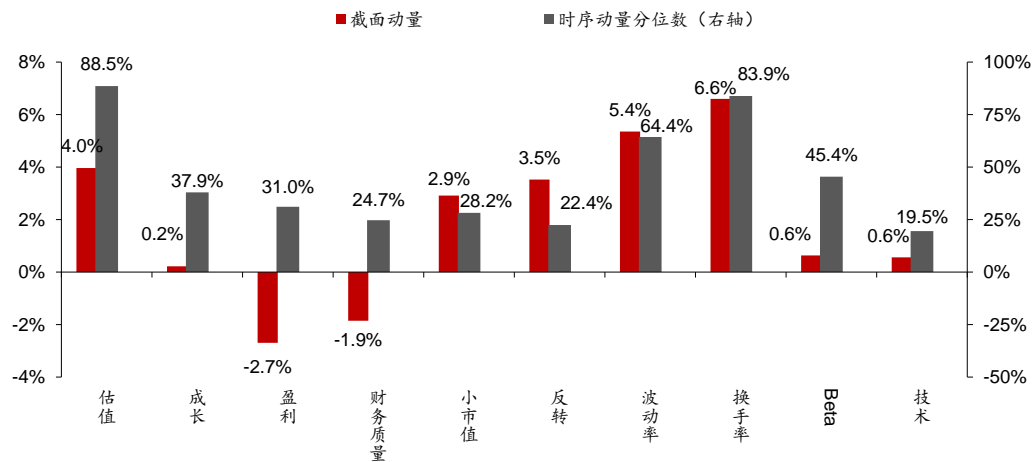
定义因子动量为因子在过去 12 个月末截面期的 21 日 Rank IC 均值。从截面角度看，波动率、换手率因子动量较强，盈利、财务质量因子动量较弱。从时序角度看，估值、换手率因子动量处于历史高位，超过 2007 年以来的 80% 分位数水平；技术因子动量处于历史低位，低于 2007 年以来的 20% 分位数水平。

图表14： 大类风格因子动量历史序列



资料来源：Wind，华泰研究

图表15： 大类风格因子 6 月底截面动量和 2007 年以来历史分位数



资料来源：Wind，华泰研究

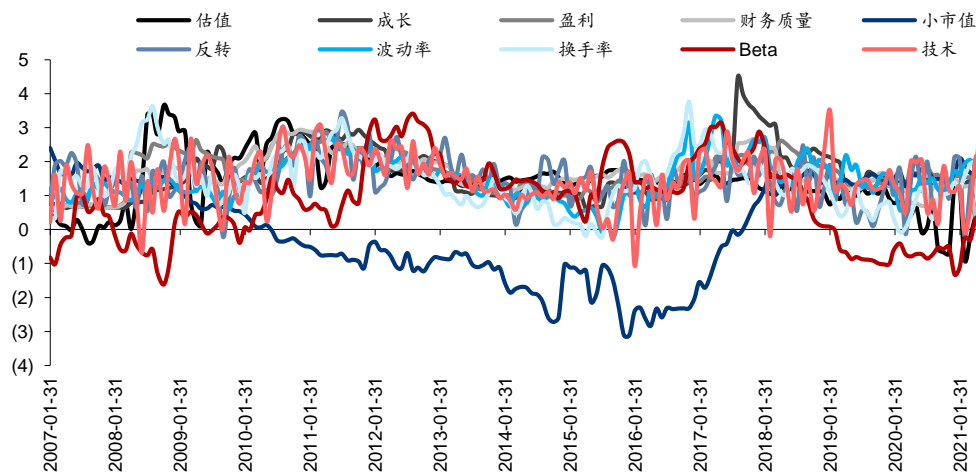
因子离散度

因子离散度计算方式为：

1. 在月末截面期，取每个中信一级行业因子值排名前 20% 个股的因子值中位数，以及因子值排名后 20% 个股的因子值中位数，两者求差值；
2. 30 个行业的差值求平均值，得到当前截面期该因子离散度原始值；
3. 对当前截面期因子离散度原始值做标准化，计算其在过去 72 个月末截面期的 Z 分数。

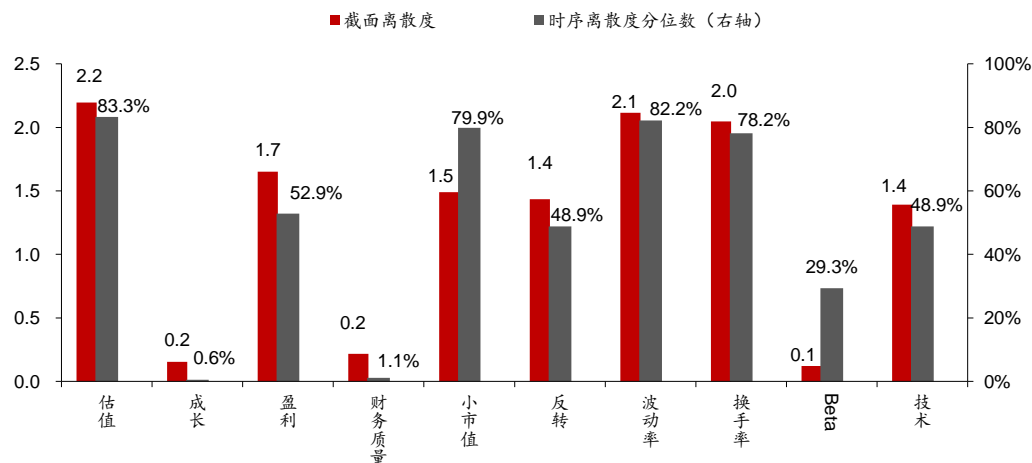
从截面角度看，估值、波动率、换手率因子离散度相对较高，成长、财务质量、Beta 因子离散度相对较低。从时序角度看，估值、小市值、波动率、换手率因子离散度处于历史高位，超过或接近 2007 年以来的 80% 分位数水平，可能处于过度忽视状态；成长、财务质量因子离散度处于历史低位，低于 2007 年以来的 5% 分位数水平，可能处于过度投资状态。

图表16： 大类风格因子离散度历史序列



资料来源：Wind，华泰研究

图表17： 大类风格因子 6 月底截面离散度和 2007 年以来历史分位数



资料来源：Wind，华泰研究

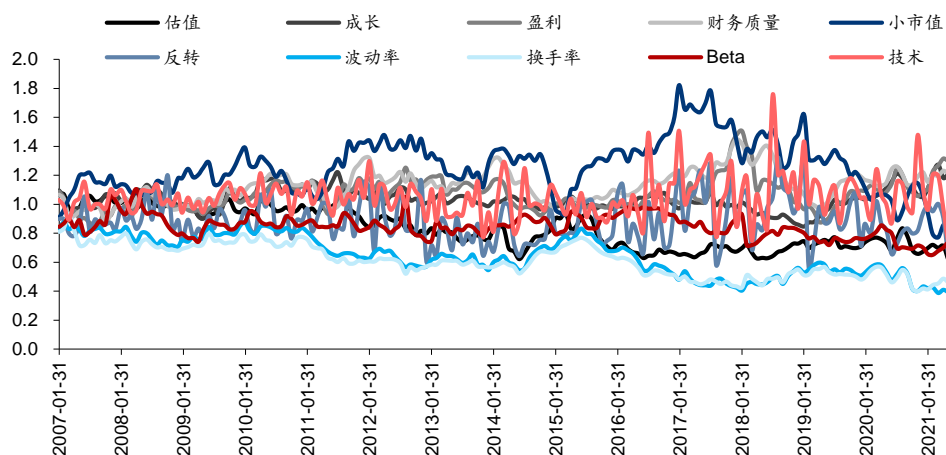
因子拥挤度

因子拥挤度计算方式为：

1. 在每个月末截面期，分别计算个股 a) 过去 63 个交易日的日度涨跌幅标准差作为波动率指标，b) 过去 63 个交易日的日均换手率作为换手率指标，c) 过去 63 个交易日个股收益与万得全 A 收益线性回归的回归系数作为 Beta 指标；
2. 取各个中信一级行业因子值排名前 20% 个股的波动率/换手率/Beta 均值，以及因子值排名后 20% 个股的波动率/换手率/Beta 均值，两者求比值；
3. 对波动率、换手率、Beta 的三个比值求均值，得到当前截面期该因子拥挤度。

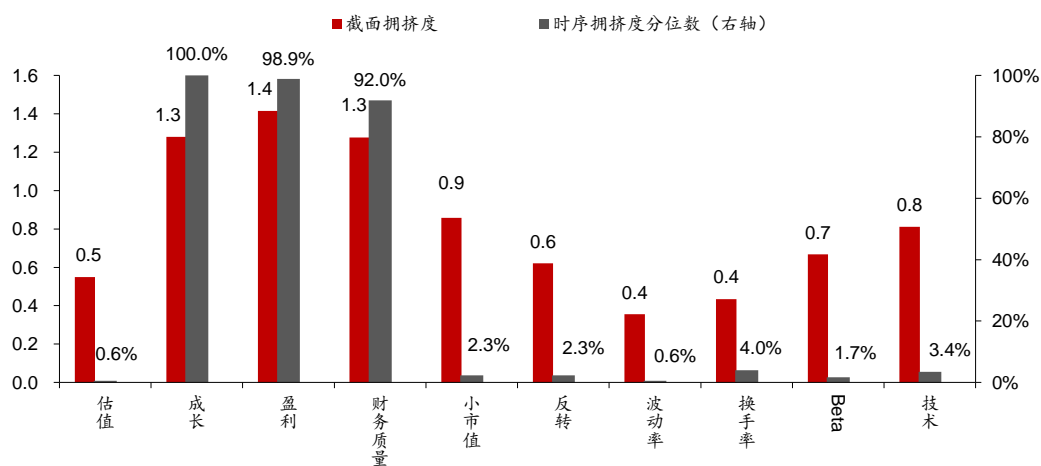
从截面角度看，成长、盈利、财务质量因子拥挤度相对较高，估值、波动率、换手率因子拥挤度相对较低。从时序角度看，成长、盈利、财务质量因子拥挤度超过 2007 年以来的 90% 历史分位数水平，可能处于交易拥挤状态；其余因子拥挤度低于 2007 年以来的 5% 历史分位数水平，交易相对不拥挤。

图表18： 大类风格因子拥挤度历史序列



资料来源：Wind，华泰研究

图表19： 大类风格因子 6 月底截面拥挤度和 2007 年以来历史分位数



资料来源：Wind，华泰研究

外生变量视角

外生变量视角假设风格因子表现受市场变量、宏观变量等外生变量驱动，对每个大类风格因子筛选出具备预测能力的外生变量，使用线性回归预测未来一个月的 Rank IC 值。以下分别展示最新一期预测方法及预测结果。

外生变量选取

选取 6 项市场指标和 13 项宏观指标作为外生变量。选择沪深 300 和中证 500 月涨跌幅、月波动率及月均换手率，用于表征市场行情的变化。从国民经济核算、工业、价格指数、银行与货币、利率与汇率 5 个方面，选择较常用的 13 项宏观指标。

时间序列回归模型的一个重要前提是自变量为平稳时间序列，而大部分宏观指标和市场指标在时间序列上的分布非平稳，不符合预测模型对自变量的假设，因此我们针对不同的外生变量，进行一定差分处理或计算变化率，使得外生变量在单位根检验下均为平稳时间序列。另外，部分宏观变量数据发布具有滞后性，我们在建模时将这部分宏观数据按其实际发布时间滞后使用。例如，T-1 月的 M1 同比数据通常在 T 月中旬公布，因此我们在 T 月末建模时用到的是 T-1 月的 M1 同比数据。

下表展示具体 19 项外生变量及其所属类别和变量处理方法。为了更清楚地展示部分宏观变量的滞后期，我们以 T 月末为截面期，最右列为外生变量能取到的最新数据时间。

图表20： 外生变量列表

变量类别	变量细分类别	变量名称	变量处理方法	最新数据时间
市场变量		沪深 300 月涨跌幅	不处理	第 T 月
市场变量		中证 500 月涨跌幅	不处理	第 T 月
市场变量		沪深 300 月波动率	变化率	第 T 月
市场变量		中证 500 月波动率	变化率	第 T 月
市场变量		沪深 300 月均换手率	一阶差分	第 T 月
市场变量		中证 500 月均换手率	一阶差分	第 T 月
宏观变量	国民经济核算	PMI	二阶差分	第 T 月
宏观变量	国民经济核算	宏观经济景气指数先行指数	变化率	第 T-3 月
宏观变量	工业	工业增加值当月同比	二阶差分	第 T-1 月
宏观变量	价格指数	CPI 当月同比	二阶差分	第 T-1 月
宏观变量	价格指数	PPI 全部工业品当月同比	一阶差分	第 T-1 月
宏观变量	价格指数	企业商品价格指数 CGPI 当月同比	二阶差分	第 T-1 月
宏观变量	银行与货币	M1 同比	一阶差分	第 T-1 月
宏观变量	银行与货币	M2 同比	一阶差分	第 T-1 月
宏观变量	银行与货币	大型存款机构存款准备金率	二阶差分	第 T 月
宏观变量	利率与汇率	中债国债到期收益率 3 个月	变化率	第 T 月
宏观变量	利率与汇率	国债收益率差 10 年期-3 月期	变化率	第 T 月
宏观变量	利率与汇率	美元兑人民币汇率	变化率	第 T 月
宏观变量	利率与汇率	SHIBOR1W	变化率	第 T 月

资料来源：Wind，华泰研究

外生变量预测方法

参考 Hua, Kantsyrev 和 Qian 于 2012 年发表论文 *Factor-Timing Model*，使用外生变量预测大类风格因子 Rank IC 值。首先使用条件期望模型和 AIC 准则，在 T-1 月最后一个交易日，对每个大类风格因子筛选出可能影响因子 Rank IC 值的条件变量；其次将大类风格因子与筛选出的条件变量做线性回归，预测 T 月的因子 Rank IC 值。

条件期望模型

假设单个风格因子 Rank IC 时序向量 Y 与外生变量时序向量 X 的联合分布服从多元正态分布, 即:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}\right),$$

则在给定 $X=x$ 的情况下, Y 的条件分布仍为正态分布, 条件数学期望和条件方差分别为:

$$\mu_{Y|X=x} = E(Y|X=x) = \mu_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(x - \mu_X)$$

$$\Sigma_{Y|X} = \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}$$

其中 $E(Y|X=x)$ 称为 Y 关于 X 的回归, 它是 x 的线性函数。

AIC 和 AICc 准则

使用 AIC 赤池信息量准则衡量统计模型拟合优良性。相比其它模型评价指标, AIC 的优势在于兼顾模型简洁性和准确性。在进行多个模型比较和选择时, 通常选择 AIC 最小的模型。

基于条件期望模型的 AIC 计算公式如下:

$$AIC = T \cdot \log(|\Sigma_{Y|X}|) + 2N \cdot K$$

其中 T 为样本窗口长度, N 为风格因子数量 (这里固定为 1), K 为外生变量数量。

当样本数量较小时, 基于 AIC 的方法可能会产生模型过拟合问题。为避免此问题, 研究者提出 AICc 指标, 在 AIC 上添加小样本误差修正量。基于条件期望模型的 AICc 计算公式如下:

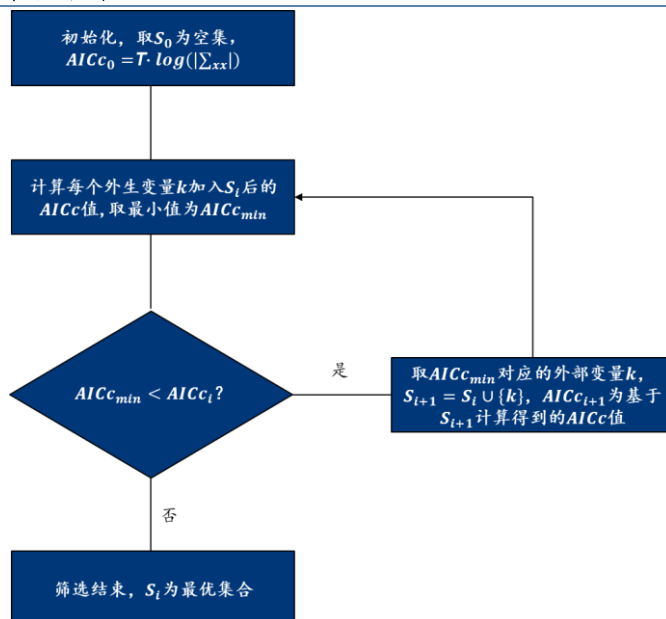
$$AICc = T \cdot \log(|\Sigma_{Y|X}|) + 2N \cdot K + \frac{2K \cdot (K+1)}{T-K-1}$$

外生变量筛选流程

样本区间内筛选外生变量流程如下图所示, 具体步骤为:

1. 最优外生变量集合 S_0 初始为空集, $AICc_0 = T \cdot \log(|\Sigma_{xx}|)$;
2. 遍历不在 S_i 中的每个外生变量 k , 计算将 k 加入到集合 S_i 后模型 $AICc$ 值, 取最小值 $AICc_{min}$;
3. 若 $AICc_{min}$ 小于 $AICc_i$, 则转向第 4 步; 否则筛选结束, 集合 S_i 为最优外生变量集合; 将 $AICc_{min}$ 对应的外生变量 k 加入 S_i , 即 $S_{i+1} = S_i \cup \{k\}$, $AICc_{i+1}$ 为基于外生变量集合 S_{i+1} 计算得到的 $AICc$, 并转至第 2 步。

图表21: 外生变量筛选流程图



资料来源: 华泰研究

回归预测模型构建

对于每个大类风格因子，取样本区间为过去 72 个月，分三步预测未来一个月 Rank IC 值：

1. 使用 AICc 在样本区间内筛选出对因子 Rank IC 值具有解释能力的外生变量；
2. 将筛选出来的全部外生变量纳入回归模型，在样本区间内构造因子 Rank IC 值序列关于外生变量的多元线性回归模型；

将最新一期的外生变量代入回归模型，得到未来一个月的因子 Rank IC 预测值。

外生变量预测结果

外生变量对大类风格因子 7 月 Rank IC 预测结果如下图所示。反转、波动率、换手率因子 Rank IC 预测值超过 5%，可能较为有效，建议短期看多上述因子表现；不看好财务质量、小市值因子未来一个月表现。

图表22： 外生变量对大类风格因子 7 月 Rank IC 预测值

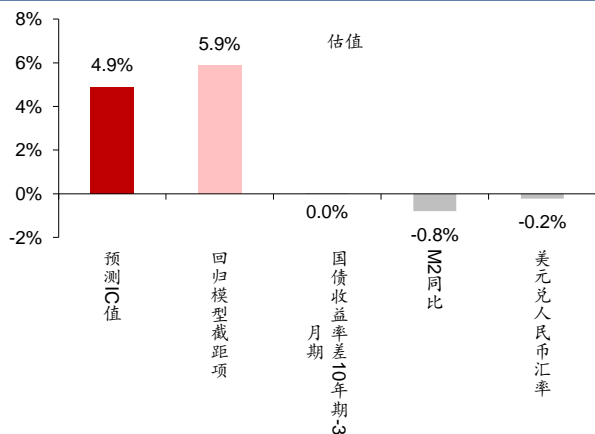


资料来源：Wind，华泰研究预测

因子预测结果解读

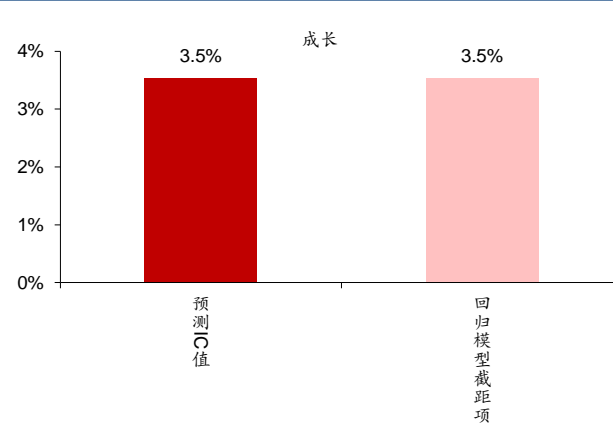
下面展示各个大类风格因子 Rank IC 预测值，以及进入回归模型的外生变量对预测结果的贡献情况，贡献度定义为自变量与回归系数的乘积。

图表23： 估值因子 Rank IC 预测值及外生变量贡献情况



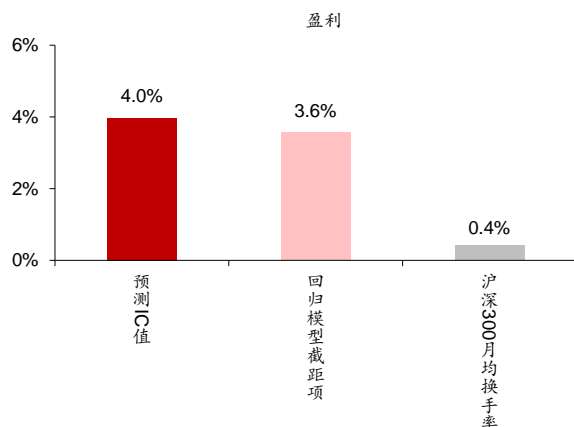
资料来源：Wind，华泰研究预测

图表24： 成长因子 Rank IC 预测值及外生变量贡献情况



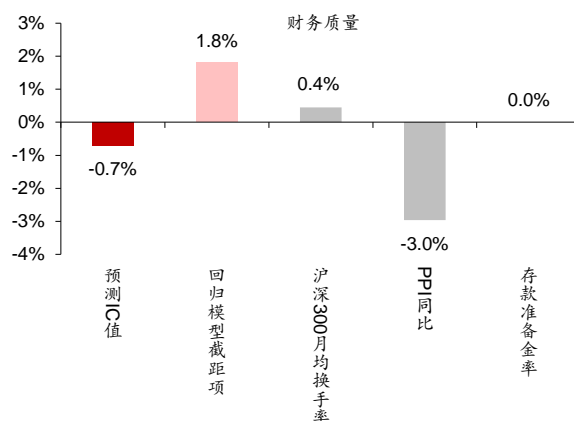
资料来源：Wind，华泰研究预测

图表25: 盈利因子 Rank IC 预测值及外生变量贡献情况



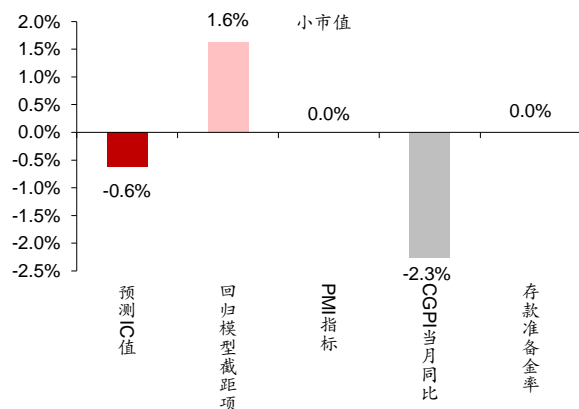
资料来源: Wind, 华泰研究预测

图表26: 财务质量因子 Rank IC 预测值及外生变量贡献情况



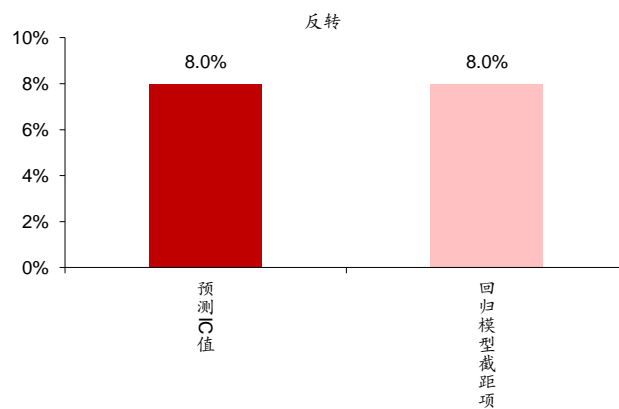
资料来源: Wind, 华泰研究预测

图表27: 小市值因子 Rank IC 预测值及外生变量贡献情况



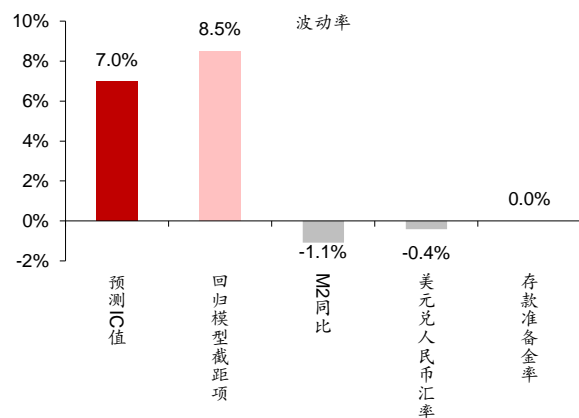
资料来源: Wind, 华泰研究预测

图表28: 反转因子 Rank IC 预测值及外生变量贡献情况



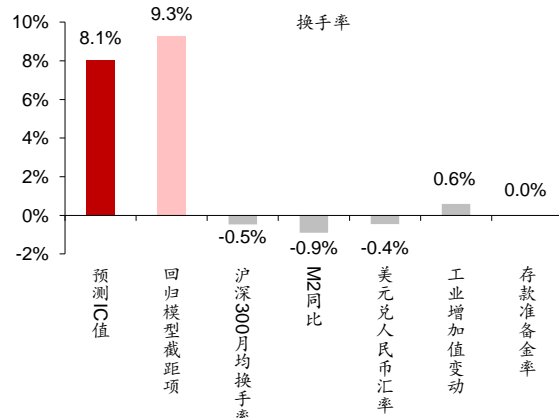
资料来源: Wind, 华泰研究预测

图表29: 波动率因子 Rank IC 预测值及外生变量贡献情况



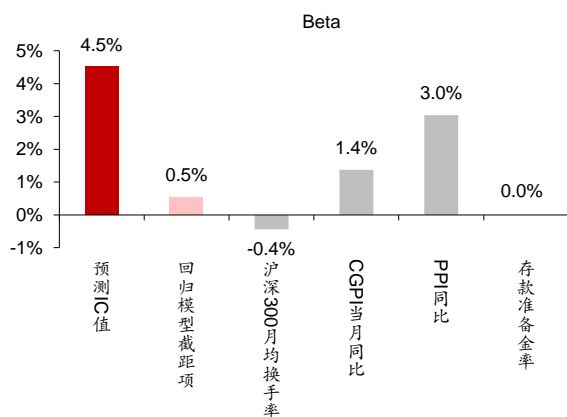
资料来源: Wind, 华泰研究预测

图表30: 换手率因子 Rank IC 预测值及外生变量贡献情况



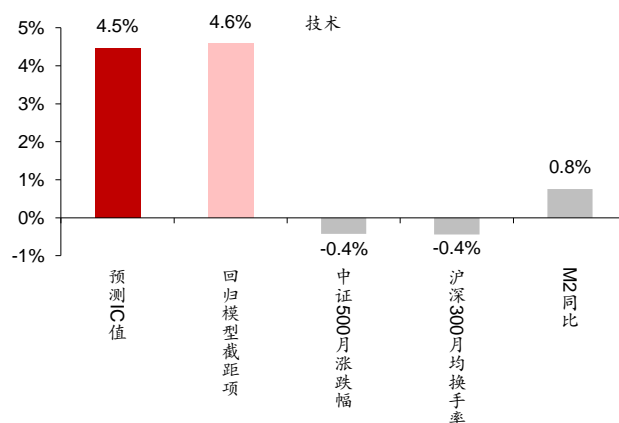
资料来源: Wind, 华泰研究预测

图表31: Beta因子 Rank IC 预测值及外生变量贡献情况



资料来源: Wind, 华泰研究预测

图表32: 技术因子 Rank IC 预测值及外生变量贡献情况



资料来源: Wind, 华泰研究预测

下面尝试对预测 Rank IC 绝对值超过 5% 的风格因子进行解读。在构建回归模型前，我们对部分外生变量进行了差分处理或计算变化率，在解读预测结果时，我们应关注这些外生变量取值的边际变化，而非变量值本身。

1. 反转因子：回归模型截距项为 8.0%，长期看反转因子存在正收益。其余外生变量对反转因子 Rank IC 预测值无影响。
2. 波动率因子：回归模型截距项为 8.5%，长期看波动率因子存在正收益。M2 同比新值（5 月，8.3%）较前值（4 月，8.1%）提升 0.2pct，为 Rank IC 预测值提供-1.1%的负向贡献。
3. 换手率因子：回归模型截距项为 9.3%，长期看换手率因子存在正收益。M2 同比新值较前值提升 0.2pct，为 Rank IC 预测值提供-0.9%的负向贡献。

另外，值得注意的是外生变量对财务质量和小市值因子表现的预测：

1. 财务质量因子：Rank IC 预测值为-0.7%。回归模型截距项为 1.8%，说明长期看波动率因子存在微弱的正收益。PPI 同比新值（5 月，9%）较前值（4 月，6.8%）提升 2.2pct，为 Rank IC 预测值提供-3.0%的负向贡献。
2. 小市值因子：Rank IC 预测值为-0.6%，7 月大小盘风格看均衡。回归模型截距项为 1.6%，长期看小市值因子存在微弱的正收益。CGPI 同比增速下滑，二阶差分新值较前值降低 0.5pct，为 Rank IC 预测值提供-2.3%的负向贡献。

总的来看，财务质量和小市值因子未来表现可能受宏观环境制约。

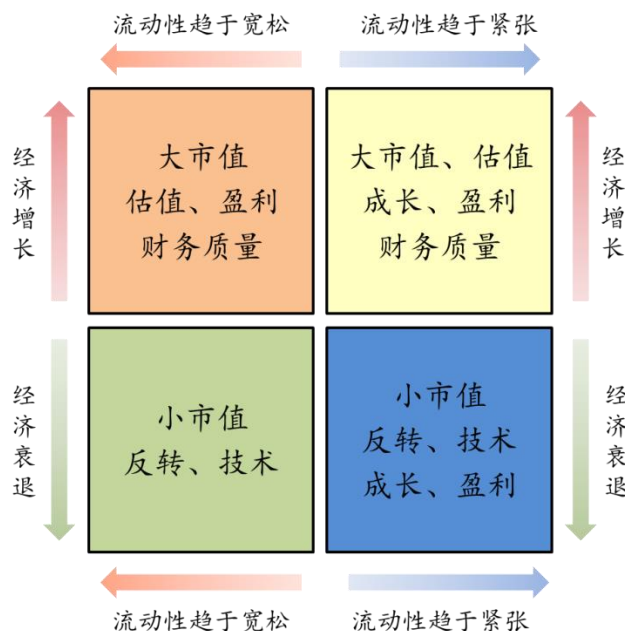
因子周期视角

因子周期视角将华泰金工经济周期研究成果应用于风格因子表现预测。预测分为两部分：自上而下构建经济系统周期运行状态与因子表现对应关系的“因子投资时钟”，根据当前周期状态预测因子未来表现的整体趋势；自下而上对风格因子去趋势累计收益率进行三周期回归拟合，外推因子未来表现的边际变化。与前述内生变量、外生变量视角不同，因子周期视角侧重于对因子中长期表现的预测。

自上而下预测整体趋势：因子投资时钟

对因子周期的研究发现，大部分因子的中长期走势与经济形势挂钩，影响较为明显的两个因素是经济增长和流动性，前者可使用库兹涅茨周期刻画，后者可使用朱格拉周期辅助判断。我们借助这两个变量将经济环境分为四种情况，不同经济周期下因子表现分类如下图所示，称之为因子投资时钟。因子投资时钟的详细构建方法请参考华泰金工研究报告《周期视角下的因子投资时钟》（20181011）。

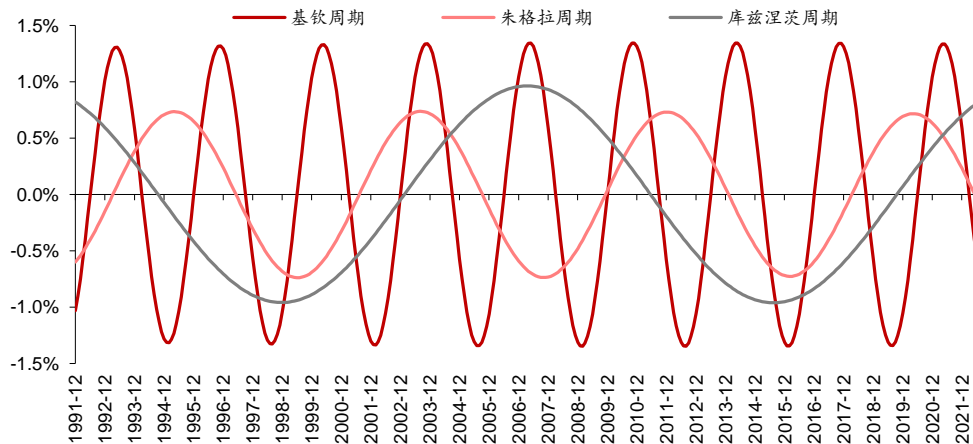
图表33：不同经济周期下的因子表现分类



资料来源：华泰研究

在不同资产类别中，商品是靠近实体经济的一种资产，商品价格长周期上的波动在一定程度上反映了实体经济的总需求，因此可借助商品识别经济的中长期波动。CRB 综合现货指数三周期分解结果如下图所示，预计未来一年处于库兹涅茨周期向上，经济长期趋势向上的环境，同时朱格拉周期向下，流动性趋于宽松。当前周期运行状态处于因子投资时钟的第四象限，建议配置大市值、估值、盈利、财务质量因子。另外，波动率、换手率因子在不同的宏观环境下表现都较好，这两个因子建议长期配置。综合来看，预测未来一年大市值、估值、盈利、财务质量、波动率、换手率因子有效。

图表34：实体经济三周期滤波结果



资料来源：Wind，华泰研究预测

自下而上预测边际变化：因子周期回归拟合

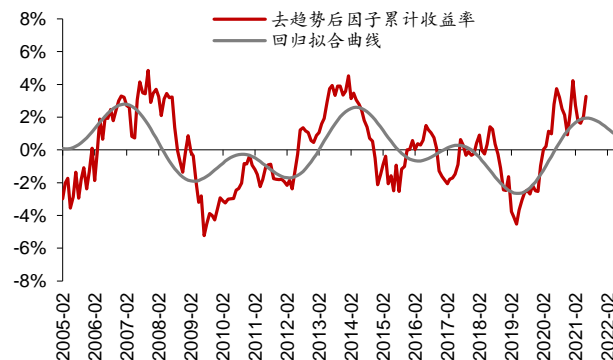
下面展示各个大类风格因子三周期回归拟合结果。三周期回归拟合详细方法请参考华泰金工研究报告《因子收益率的周期性研究初探》（20180605）。图中红线为去趋势后因子累计收益率，灰线为回归拟合曲线，灰线相对红线的延伸部分为因子未来一年表现边际变化情况预测。综合来看，波动率因子处于周期上行状态初期，Beta 因子处于周期上行状态，成长因子即将见顶回落，盈利、财务质量因子处于周期下行初期，估值、小市值、反转、换手率、技术因子即将触底回升。

图表35：估值因子去趋势后累计收益率与三周期回归拟合



资料来源：Wind，华泰研究预测

图表36：成长因子去趋势后累计收益率与三周期回归拟合



资料来源：Wind，华泰研究预测

图表37：盈利因子去趋势后累计收益率与三周期回归拟合



资料来源：Wind，华泰研究预测

图表38：财务质量因子去趋势后累计收益率与三周期回归拟合



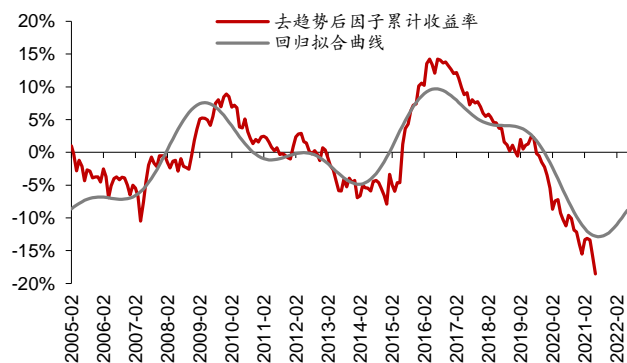
资料来源：Wind，华泰研究预测

图39: 小市值因子去趋势后累计收益率与三周期回归拟合



资料来源: Wind, 华泰研究预测

图40: 反转因子去趋势后累计收益率与三周期回归拟合



资料来源: Wind, 华泰研究预测

图41: 波动率因子去趋势后累计收益率与三周期回归拟合



资料来源: Wind, 华泰研究预测

图42: 换手率因子去趋势后累计收益率与三周期回归拟合



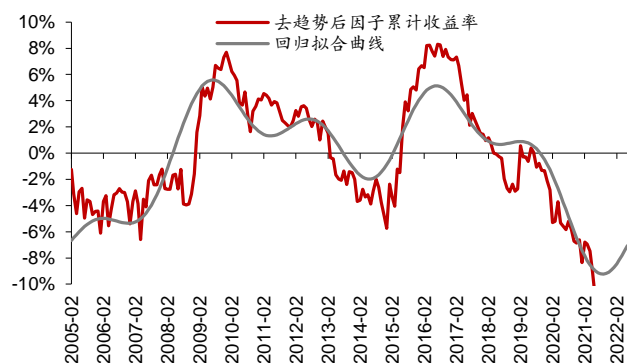
资料来源: Wind, 华泰研究预测

图43: Beta 因子去趋势后累计收益率与三周期回归拟合



资料来源: Wind, 华泰研究预测

图44: 技术因子去趋势后累计收益率与三周期回归拟合



资料来源: Wind, 华泰研究预测

风险提示

风格因子的效果与宏观环境和大盘走势密切相关，历史结果不能预测未来。金融周期规律被打破。市场出现超预期波动，导致拥挤交易。

附录

附录一：所有细分因子计算方法

图表45： 报告中涉及的所有细分因子及其计算方式

因子类别	因子简称	因子简称	因子方向	因子计算方式	计算说明
估值	估值-EP	EP	1	Wind 因子 pe_ttm 的倒数	估值类计算过程中涉及到的四个 Wind 因子均可以按交易日提取。
估值	估值-BP	BP	1	Wind 因子 pb_lf 的倒数	
估值	估值-SP	SP	1	Wind 因子 ps_ttm 的倒数	
估值	股息率	DP	1	Wind 因子 dividendyield2	
成长	营业收入增长率	Sales_G_q	1	由 Wind 因子 oper_rev 自行计算同比增长率	成长、盈利、财务质量、市值类计算过程中涉及到的 11 个 Wind 因子均按季度(报告期)提取,使用方式为:截面日在 4 月 30 日及以后的可使用当年一季报、上一年年报信息,截面日在 8 月 31 日及以后的可使用当年半年报信息,截面日在 10 月 31 日及以后的可使用当年三季报信息。
成长	净利润增长率	Profit_G_q	1	由 Wind 因子 net_profit_is 自行计算同比增长率	
成长	ROE 增长率	ROE_G_q	1	由 Wind 因子 roe_avg 自行计算同比增长率	
盈利	ROE	ROE	1	Wind 因子 roe_avg	
盈利	ROA	ROA	1	Wind 因子 roa	成长、盈利、财务质量、市值类计算过程中涉及到的 11 个 Wind 因子均按季度(报告期)提取,使用方式为:截面日在 4 月 30 日及以后的可使用当年一季报、上一年年报信息,截面日在 8 月 31 日及以后的可使用当年半年报信息,截面日在 10 月 31 日及以后的可使用当年三季报信息。
盈利	毛利率	grossprofitmargin	1	Wind 因子 grossprofitmargin	
盈利	净利率	netprofitmargin	1	Wind 因子 netprofitmargin	
财务质量	总资产周转率	assetturnover	1	Wind 因子 assetturn	
财务质量	资产负债率	debttoassets	-1	Wind 因子 debttoassets	成长、盈利、财务质量、市值类计算过程中涉及到的 11 个 Wind 因子均按季度(报告期)提取,使用方式为:截面日在 4 月 30 日及以后的可使用当年一季报、上一年年报信息,截面日在 8 月 31 日及以后的可使用当年半年报信息,截面日在 10 月 31 日及以后的可使用当年三季报信息。
财务质量	流动比率	current	1	Wind 因子 current (*部分金融企业不适用)	
财务质量	资产的经营现金流量回报率	ocftoassets	1	Wind 因子 ocftoassets	
小市值	对数总市值	ln_capital	-1	Wind 因子 mkt_cap_ard 取对数	
反转	1 个月反转	return_1m	-1	个股最近 N 个月收益率 (N=1, 3)	个股每日收益率均采用 Wind 后复权收盘价自行计算;个股每日换手率采用 Wind 因子 turn, 将个股停牌、涨跌停的交易日的换手率设为 nan(计算均值、标准差时不取用);最近 1 个月的定义:若当前截面日为月底最后一个交易日,则最近 1 个月指代该自然月内所有交易日,若当前截面日不是月底最后一个交易日,例如截面日为 9 月 20 日,则指代 8 月 21 日(含)至 9 月 20 日(含)之间的所有交易日,最近 3 个月的定义依此类推;计算技术指标时均使用后复权收盘价。
反转	3 个月反转	return_3m	-1	个股最近 N 个月收益率 (N=1, 3)	
反转	衰减换手率加权 3 个月反转	exp_wgt_return_3m	-1	最近 N 个月内,在每个交易日对个股计算如下公式的值,最终求和:日换手率乘以函数 $\exp(-x_i/N/4)$ 乘以日收益率。其中 x_i 为该日距离截面日的交易日个数 (N=1, 3)	
反转	衰减换手率加权 6 个月反转	exp_wgt_return_6m	-1	最近 N 个月内,在每个交易日对个股计算如下公式的值,最终求和:日换手率乘以函数 $\exp(-x_i/N/4)$ 乘以日收益率。其中 x_i 为该日距离截面日的交易日个数 (N=1, 3)	
波动率	1 个月波动率	std_1m	-1	个股最近 N 个月内日收益率序列的标准差 (N=1, 3)	成长、盈利、财务质量、市值类计算过程中涉及到的 11 个 Wind 因子均按季度(报告期)提取,使用方式为:截面日在 4 月 30 日及以后的可使用当年一季报、上一年年报信息,截面日在 8 月 31 日及以后的可使用当年半年报信息,截面日在 10 月 31 日及以后的可使用当年三季报信息。
波动率	3 个月波动率	std_3m	-1	个股最近 N 个月内日收益率序列的标准差 (N=1, 3)	
波动率	FF 三因子残差 1 个月波动率	IVR_FF3factor_1m	-1	个股最近 N 个月内日收益率序列对 Fama-French 三因子(中证全指日收益率、总市值因子日收益率、BP 因子日收益率)进行多元线性回归的残差的标准差 (N=1, 3)	
波动率	FF 三因子残差 3 个月波动率	IVR_FF3factor_3m	-1	个股最近 N 个月内日收益率序列对 Fama-French 三因子(中证全指日收益率、总市值因子日收益率、BP 因子日收益率)进行多元线性回归的残差的标准差 (N=1, 3)	
换手率	1 个月日均换手率	turn_1m	-1	个股最近 N 个月内日均换手率 (N=1, 3)	成长、盈利、财务质量、市值类计算过程中涉及到的 11 个 Wind 因子均按季度(报告期)提取,使用方式为:截面日在 4 月 30 日及以后的可使用当年一季报、上一年年报信息,截面日在 8 月 31 日及以后的可使用当年半年报信息,截面日在 10 月 31 日及以后的可使用当年三季报信息。
换手率	3 个月日均换手率	turn_3m	-1	个股最近 N 个月内日均换手率 (N=1, 3)	
换手率	近 1 个月/近 2 年的日均换手率	bias_turn_1m	-1	个股最近 1 个月内日均换手率除以最近 2 年内日均换手率	
beta	近 100 周相对上证综指的 beta	Beta	-1	Wind 因子 beta_100w	
技术	技术-MACD	macd	-1	经典技术指标,参数设置:长周期取 30 日,短周期取 10 日,计算 DEA 均线时的周期(也称为中周期)取 15 日	成长、盈利、财务质量、市值类计算过程中涉及到的 11 个 Wind 因子均按季度(报告期)提取,使用方式为:截面日在 4 月 30 日及以后的可使用当年一季报、上一年年报信息,截面日在 8 月 31 日及以后的可使用当年半年报信息,截面日在 10 月 31 日及以后的可使用当年三季报信息。
技术	技术-DEA	dea	-1	经典技术指标,参数设置:长周期取 30 日,短周期取 10 日,计算 DEA 均线时的周期(也称为中周期)取 15 日	
技术	技术-DIF	dif	-1	经典技术指标,参数设置:长周期取 30 日,短周期取 10 日,计算 DEA 均线时的周期(也称为中周期)取 15 日	
技术	技术-RSI	rsi	-1	RSI、PSY、BIAS 为经典技术指标,计算周期均取 20 日	
技术	技术-PSY	psy	-1	RSI、PSY、BIAS 为经典技术指标,计算周期均取 20 日	成长、盈利、财务质量、市值类计算过程中涉及到的 11 个 Wind 因子均按季度(报告期)提取,使用方式为:截面日在 4 月 30 日及以后的可使用当年一季报、上一年年报信息,截面日在 8 月 31 日及以后的可使用当年半年报信息,截面日在 10 月 31 日及以后的可使用当年三季报信息。
技术	技术-BIAS	bias	-1	RSI、PSY、BIAS 为经典技术指标,计算周期均取 20 日	

资料来源: Wind, 华泰研究

上表第四列因子方向解释:取值为 1 代表因子值越大越好, -1 代表因子值越小越好。当采用等权法合成风格因子时,需将因子值乘以因子方向之后再相加。

附录二：风格因子合成、Rank IC 值计算、因子收益率计算方法

我们以某一截面期 T 估值类风格因子的合成为例，具体说明风格因子合成过程：

1. 因子原始值计算：设第 T 期指定股票池（全 A 股）内共存在 N 只股票，首先根据附录一表格计算细分因子 EP、BP、SP、股息率在第 T 期的值，均为 $N \times 1$ 维向量；
2. 因子的行业与市值中性化处理：用 OLS 线性回归拟合

$$X^T = \log(\text{mktcap}^T) \cdot \beta_0 + \sum_j \text{Indus}_j^T \cdot \beta_j + \varepsilon$$

X^T ：第 T 期因子 X 的值， $N \times 1$ 维向量

$\log(\text{mktcap}^T)$ ：第 T 期‘对数总市值’因子的值， $N \times 1$ 维向量

Indus_j^T ：第 T 期第 j 个行业因子哑变量（属于该行业为 1，否则为 0）， $N \times 1$ 维向量

β_j, ε ：回归待拟合系数和残差， β_j 为数值， ε 为 $N \times 1$ 维向量

取上述回归方程的残差向量替代原始计算的因子值 X^T 作为第 T 期因子 X 的新值，仍记为 X^T ，这里的 X 指代 EP、BP、SP、股息率等细分因子（特别地，若因子 X 为‘对数总市值’因子，则它只对行业哑变量回归，亦即对该因子只进行行业中性化处理）；

3. 因子去极值、标准化：
 - 1) 中位数去极值：对于经过中性化处理之后的因子值 X^T （ $N \times 1$ 维向量），先计算其中位数 D_M ，再计算 $|X^T - D_M|$ 的中位数 D_{M1} ，接下来将向量 X^T 中所有大于 $D_M + 5D_{M1}$ 的元素重设为 $D_M + 5D_{M1}$ ，将向量 X^T 中所有小于 $D_M - 5D_{M1}$ 的元素重设为 $D_M - 5D_{M1}$ ；
 - 2) 标准化：将去极值处理后的因子值 X^T （ $N \times 1$ 维向量）减去其现在的均值、再除以其现在的标准差，就得到了标准化后的新因子值，仍记为 X^T ；
4. 风格因子合成与 Rank IC 计算：将属于估值类的四个细分因子值（经过以上所有处理之后）乘以各自的因子方向（1 或 -1）之后直接等权相加，就得到了估值风格因子的值，与个股下期收益率向量 r^{T+1} 计算 Spearman 秩相关系数，即得到估值风格因子的 Rank IC 值；
5. 细分因子的 Rank IC 值即用第 2 步处理结束后得到的因子值 X^T 乘以因子方向（1 或 -1）再与个股下期收益率向量 r^{T+1} 计算 Spearman 秩相关系数。
6. 将 T 期因子值与 T+1 期个股收益率进行线性回归，得到的回归系数即为因子收益率。回归模型为：

$$r_i^{T+1} = \sum_j X_j^T f_{ji}^T + X_d^T d_i^T + \mu_i^T$$

r_i^{T+1} ：股票 i 在第 T+1 期的收益率

d_i^T ：股票 i 在第 T 期因子 d 上的暴露度

f_{ji}^T ：股票 i 在第 T 期第 j 个行业因子上的暴露度（属于该行业则为 1，否则为 0）

X_j^T （或 X_d^T ）：第 T 期第 j 个行业因子（或因子 d）的因子收益率，需回归拟合

μ_i^T ：股票 i 在第 T 期的残差收益率

在截面期对因子 d 进行回归测试，能够得到该因子的因子收益率。

免责声明

分析师声明

本人，林晓明、何康，兹证明本报告所表达的观点准确地反映了分析师对标的证券或发行人的个人意见；彼以往、现在或未来并无就其研究报告所提供的具体建议或所表达的意见直接或间接收取任何报酬。

一般声明及披露

本报告由华泰证券股份有限公司（已具备中国证监会批准的证券投资咨询业务资格，以下简称“本公司”）制作。本报告所载资料是仅供接收人的严格保密资料。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司及其关联机构（以下统称为“华泰”）对该等信息的准确性及完整性不作任何保证。

本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，华泰可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。以往表现并不能指引未来，未来回报并不能得到保证，并存在损失本金的可能。华泰不保证本报告所含信息保持在最新状态。华泰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。

本公司不是 FINRA 的注册会员，其研究分析师亦没有注册为 FINRA 的研究分析师/不具有 FINRA 分析师的注册资格。

华泰力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，华泰及作者均不承担任何法律责任。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。

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，过往的业绩表现不应作为日后回报的预示。华泰不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。

华泰及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，华泰可能会持有报告中提到的公司所发行的证券头寸并进行交易，为该公司提供投资银行、财务顾问或者金融产品等相关服务或向该公司招揽业务。

华泰的销售人员、交易人员或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。华泰没有将此意见及建议向报告所有接收者进行更新的义务。华泰的资产管理部门、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。投资者应当考虑到华泰及/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。有关该方面的具体披露请参照本报告尾部。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布的机构或人员，也并非意图发送、发布给因可得到、使用本报告的行为而使华泰违反或受制于当地法律或监管规则的机构或人员。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人（无论整份或部分）等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并需在使用前获取独立的法律意见，以确定该引用、刊发符合当地适用法规的要求，同时注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

中国香港

本报告由华泰证券股份有限公司制作，在香港由华泰金融控股（香港）有限公司向符合《证券及期货条例》及其附属法律规定的机构投资者和专业投资者的客户进行分发。华泰金融控股（香港）有限公司受香港证券及期货事务监察委员会监管，是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。在香港获得本报告的人员若有任何有关本报告的问题，请与华泰金融控股（香港）有限公司联系。

香港-重要监管披露

- 华泰金融控股（香港）有限公司的雇员或其关联人士没有担任本报告中提及的公司或发行人的高级人员。更多信息请参见下方“美国-重要监管披露”。

美国

在美国本报告由华泰证券（美国）有限公司向符合美国监管规定的机构投资者进行发表与分发。华泰证券（美国）有限公司是美国注册经纪商和美国金融业监管局（FINRA）的注册会员。对于其在美国分发的研究报告，华泰证券（美国）有限公司根据《1934年证券交易法》（修订版）第15a-6条规定以及美国证券交易委员会人员解释，对本研究报告内容负责。华泰证券（美国）有限公司联营公司的分析师不具有美国金融监管（FINRA）分析师的注册资格，可能不属于华泰证券（美国）有限公司的关联人员，因此可能不受FINRA关于分析师与标的公司沟通、公开露面和所持交易证券的限制。华泰证券（美国）有限公司是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。任何直接从华泰证券（美国）有限公司收到此报告并希望就本报告所述任何证券进行交易的人士，应通过华泰证券（美国）有限公司进行交易。

美国-重要监管披露

- 分析师林晓明、何康本人及相关人士并不担任本报告所提及的标的证券或发行人的高级人员、董事或顾问。分析师及相关人士与本报告所提及的标的证券或发行人并无任何相关财务利益。本披露中所提及的“相关人士”包括FINRA定义下分析师的家庭成员。分析师根据华泰证券的整体收入和盈利能力获得薪酬，包括源自公司投资银行业务的收入。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或不时会以自身或代理形式向客户出售及购买华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或其高级管理层、董事和雇员可能会持有本报告中所提到的任何证券（或任何相关投资）头寸，并可能不时进行增持或减持该证券（或投资）。因此，投资者应该意识到可能存在利益冲突。

评级说明

投资评级基于分析师对报告发布日后6至12个月内行业或公司回报潜力（含此期间的股息回报）相对基准表现的预期（A股市场基准为沪深300指数，香港市场基准为恒生指数，美国市场基准为标普500指数），具体如下：

行业评级

- 增持：**预计行业股票指数超越基准
中性：预计行业股票指数基本与基准持平
减持：预计行业股票指数明显弱于基准

公司评级

- 买入：**预计股价超越基准15%以上
增持：预计股价超越基准5%~15%
持有：预计股价相对基准波动在-15%~5%之间
卖出：预计股价弱于基准15%以上
暂停评级：已暂停评级、目标价及预测，以遵守适用法规及/或公司政策
无评级：股票不在常规研究覆盖范围内。投资者不应期待华泰提供该等证券及/或公司相关的持续或补充信息

法律实体披露

中国: 华泰证券股份有限公司具有中国证监会核准的“证券投资咨询”业务资格, 经营许可证编号为: 91320000704041011J

香港: 华泰金融控股(香港)有限公司具有香港证监会核准的“就证券提供意见”业务资格, 经营许可证编号为: AOK809

美国: 华泰证券(美国)有限公司为美国金融业监管局(FINRA)成员, 具有在美国开展经纪交易商业业务的资格, 经营业务许可编号为: CRD#:298809/SEC#:8-70231

华泰证券股份有限公司**南京**

南京市建邺区江东中路228号华泰证券广场1号楼/邮政编码: 210019

电话: 86 25 83389999/传真: 86 25 83387521

电子邮件: ht-rd@htsc.com

深圳

深圳市福田区益田路5999号基金大厦10楼/邮政编码: 518017

电话: 86 755 82493932/传真: 86 755 82492062

电子邮件: ht-rd@htsc.com

北京

北京市西城区太平桥大街丰盛胡同28号太平洋保险大厦A座18层/
邮政编码: 100032

电话: 86 10 63211166/传真: 86 10 63211275

电子邮件: ht-rd@htsc.com

上海

上海市浦东新区东方路18号保利广场E栋23楼/邮政编码: 200120

电话: 86 21 28972098/传真: 86 21 28972068

电子邮件: ht-rd@htsc.com

华泰金融控股(香港)有限公司

香港中环皇后大道中99号中环中心58楼5808-12室

电话: +852-3658-6000/传真: +852-2169-0770

电子邮件: research@htsc.com

<http://www.htsc.com.hk>

华泰证券(美国)有限公司

美国纽约哈德逊城市广场10号41楼(纽约10001)

电话: +212-763-8160/传真: +917-725-9702

电子邮件: Huatai@htsc-us.com

<http://www.htsc-us.com>

©版权所有2021年华泰证券股份有限公司