

# 行业配置策略:中观景气视角(2)

## 华泰研究

2022年7月18日 | 中国内地

深度研究

### 本文优化了中观景气建模框架,构建了中观景气轮动策略

中观行业景气系列的定位是用高频、及时的行业基本面指标,对低频、滞后的行业财务数据的走势作出及时预判。本文是中观行业景气系列第四篇研报,主要工作是对前期报告提出的中观景气建模框架进行优化,生成中观行业景气指数,并应用于三类场景。优化后框架主要包括行业指标库构建、行业指标预处理、指标评价和筛选、景气指数生成等四个步骤。其中,行业指标预处理改用简化方法,指标评价和筛选改为滚动窗口内进行。中观行业景气指数的最重要应用是构建中观景气轮动策略。2016-04 至 2022-06,轮动策略相对于等权基准的超额年化收益是 13.32 pct。

# 最新的中观景气建模框架充分考虑了行业逻辑的动态变化

在最新的中观景气建模框架中:行业指标库构建从行业产业链结构出发,寻找与行业景气度存在逻辑联系的指标;行业指标预处理采用简化方法,得到当月同比口径的代理指标序列和财务指标序列;指标评价和筛选在每个财报季结束时,主要依据代理指标相对财务参照的相关性和领先滞后性,更新中观景气代理指标,非财报季则沿用最近筛选结果;景气指数生成将代理指标输入 Simple-Nowcasting 模型,合成中观行业景气指数。从指标预处理到景气指数生成均在 60 个月的滚动窗口内由程序自动执行,既保证了结果的公允性,也使得模型能够及时捕捉行业逻辑的变化。

### 中观行业景气指数应用之景气状态判断

中观行业景气指数的基础应用是判断行业景气状态。如在 2022-03-31,由于截至 2021Q4 和 2022Q1 的真实财报尚未披露完毕,只能获取 2021Q3 的财务数据,无法对 2022Q1 相对于 2021Q4 的景气边际变化做出判断;对此,可以沿用 2021-10-31 (即 2021Q3 财报披露完毕)选中的景气代理指标进行建模,得到截至 2022-03 的景气指数。本文以钢铁行业为例,得出其 2022Q1 的景气状况相对于 2021Q4 边际显著恶化。

# 中观行业景气指数应用之单行业择时

以净利/ROE/营收/毛利同比为财务参照,计算各行业的净利/ROE/营收/毛利景气指数。从交易景气相对位置和交易景气边际变化的逻辑出发,每条景气指数均可生成 orig/mom/moma3/qoq 四种模式因子,每个行业均有 16 个因子。基于因子正负性,对行业相对万得全 A 的超额净值指数开展择时,发现大多数因子在大多数行业上择时是有效的;其中,净利景气指数环比变化的择时效果最佳,跟投资者决策习惯有一定关联。考虑到景气指数择时更适合业绩呈现周期性波动的行业,本文进一步构建了顺周期风格择时策略。该策略成功抓到了顺周期风格最近两波大行情。

#### 中观行业景气指数应用之行业间轮动

进一步利用截面上各行业因子数值的相对大小信息,开展行业间轮动。首先开展单因子分层测试;根据分层测试结果,精选刻画景气相对位置逻辑的净利\_orig、ROE\_orig、营收\_orig和刻画景气边际变化逻辑的净利\_mom、营收\_qoq等5个因子,采用截面排序加权求和,计算中观景气得分。前三个因子权重之和为1,后两个因子权重之和均为1,保证两种逻辑对策略的贡献度相等。每月末选择得分最高的4个行业配置。2016-04至2022-06,轮动策略年化收益19.04%,夏普比率0.84,最大回撤-25.91%。最后,通过因子权重随机实验,验证了策略具有较强的鲁棒性。

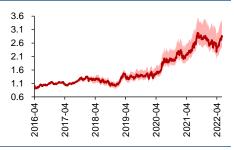
风险提示:模型根据历史规律总结,历史规律可能失效;报告中涉及到的具体行业不代表任何投资意见,请投资者谨慎、理性地看待。

研究员 SAC No. S0570516010001 SFC No. BPY421

联系人 SAC No. S0570121050032 林晓明 linxiaoming@htsc.com +(86) 755 8208 0134

> 徐特, PhD xute@htsc.com +(86) 10 6321 1166

## 中观景气轮动策略





# 正文目录

本文研究导读	3
优化后的中观景气建模框架	4
行业指标库构建:主观整理,保证景气联系	4
行业指标预处理:简化步骤,贴合财务特征	5
指标评价和筛选:滚动窗口,捕捉逻辑变化	6
时差拟合精度	7
DTW 算法	7
指标综合评分	9
景气指数生成: Simple-Nowcast, 使行业间可比	10
数学原理	10
求解过程	
建模过程中可能遇到的问题解答	
应用一:景气状态判断	14
应用二:单行业择时	15
应用三:行业间轮动	19
选行业因子分层测试结果	19
中观景气得分与行业轮动策略	22
改变因子权重对策略表现的影响	24
总结与思考	25
风险提示	26



# 本文研究导读

本文是中观行业景气系列研究的第四篇报告。中观视角是华泰金工-行业配置框架中时效性 较强、信息量较大的一环,但在使用中观基本面数据的过程中面临以下问题:

- 1) 指标品类繁杂、口径不一;
- 2) 指标对行业景气状况的解释力随着时间推移而变化;
- 3) 不同的行业指标不同,难以直接进行行业间比较。

本系列的前三篇报告属于探索阶段。彼时,我们尝试引入新的工具(如第一篇报告《中观行业景气度:Nowcasting 初探》(2021-09-26)评价了Nowcasting 模型在行业中观景气建模研究的适用性)、不断完善建模方法细节(如第三篇报告《行业配置策略:中观景气视角(1)》(2022-01-18)将程序化指标清洗、评价和筛选过程纳入研究框架)来应对上述棘手问题。本文基于前三篇报告的探索,进一步做了如下工作:

- 1) 中观景气建模覆盖的行业数量提升至25个,绝大部分板块均有行业覆盖;
- 2) 确定了最终的中观景气建模框架,各步骤之间的关联性和逻辑性显著提升;
- 3) 构建了第一版可跟踪的中观景气轮动策略,同时考虑了景气的位置和边际。

本文继续使用中信行业体系,其中,食品饮料行业和非银金融行业按照前期报告《确立研究对象:行业拆分与聚类》(2020-03-03)的方式进行拆分。本文的第一部分将详细介绍优化后的中观景气建模框架;第二部分到第四部分将介绍中观行业景气指数的三个应用——行业景气状态判断、单行业择时、行业间轮动。其中,最重要的是第三个应用。

上游资源 石油石化 有色金属 世岩 周期 中游材料 基础化工 钢铁 电力设备及新能源 中游制造 机械 国防军工 可选消费 家电 消费 必须消费 食品 饮料 医药 纺织服装 农林牧渔 商贸零售 轻工制造 消费者服务 证券|| 金融 大金融 保险|| **TMT** 成长 计算机 电子 传媒 通信 稳定 公共产业 交通运输 电力及公用事业 建筑

图表1: A股大类风格-主题板块-细分行业聚类结果

注:红色代表已覆盖,黑色代表将覆盖,紫色代表暂时无法构建行业指标库以支持本文提出的中观景气建模框架资料来源:Wind,华泰研究



# 优化后的中观景气建模框架

本文对前期报告中提出的建模框架进行了优化, 各步骤及其与前期报告的差异如下:

- 1. 行业指标库构建:与前期报告相比没有本质变化。
- 2. 行业指标预处理:本文采用较为简单的预处理方式,而前期报告采用较为复杂的预处理步骤。简化方法不仅在运算量上有优势,提升应用场景的可落地性,而且其处理结果跟预测目标即行业财务参照的口径更一致,反而有助于财务状况预测。
- 3. 指标评价和筛选:本文采用在滚动窗口期内筛选出相对于财务参照相关性较高、领先性较强的指标,而前期报告是在整个历史上开展指标评价和筛选。滚动评价和筛选有利于 捕捉行业逻辑的动态变化。
- 4. 景气指数生成: 本文依然采用 Simple-Nowcasting 模型对代理指标进行降维,得到中观行业景气指数。

除第一步行业指标库构建主要依托行业逻辑主观完成之外,其余步骤均由程序执行。从成果来说,本文覆盖的财务维度相比前期报告更丰富。从应用场景来说,本文基于景气指数构建策略,而不再像前期报告那样直接使用代理指标构建策略;这一做法在一定程度上提升了中观基本面信息在行业之间的可比性。

本研究考察的财务参照包括净利润 TTM 同比增速 (简称**净利同比**)、营业收入 TTM 同比增速 (简称**营收同比**)、ROE\_TTM 环比差分 (TTM 指标再做环比能够得到类似单季同比的口径,故简称 ROE 同比)、毛利率 TTM 环比差分 (简称毛利同比)。在众多财务指标中,这些财务指标更有可能被中观基本面指标预测,且其对行业配置的增益,在前期报告《行业配置策略:景气度视角》(2020-11-05)中已经得到检验。

1. 行业指标库构建 2. 指标预处理 财报季: 重新筛选 中观景气代理指标 上游原材料供需 下游商品的供需 统一频率至月频 内部缺失值处理 统一口径至同比 非财报季: 沿用最 行业内经营状况 行业外宏观环境 离群值处理 近财报季选中指标 增长、通胀等 4. 景气指数生成 6. 应用场景 3. 指标评价和筛选 5. 主要成果 净利景气指数 肘差拟会结磨 Simple-Nowcast 景气观点判断 DTW算法 ROE景气指数 尾部缺失值预测 单行业择时 支持指标共线性 营收景气指数 惩罚项 (序列过短/延迟发布 行业间轮动 /滞后指标) 指标信息无丢失 毛利景气指数

图表2: 宏观和中观数据预处理流程总结

资料来源:华泰研究

# 行业指标库构建: 主观整理, 保证景气联系

行业指标库构建依托行业逻辑主观完成:我们首先梳理了行业的产业链结构,然后从上游原材料供求、下游商品的供求、行业内经营状况、行业外宏观环境等四个方面入手,从各类数据库(以Wind为主)寻找与行业景气度存在逻辑联系的指标。对产业链结构感兴趣的读者可以参考前期报告《中观景气度之上游资源/中游材料》(2021-10-14),该报告对石油石化等行业的指标库构建过程有较为详细的解释,其他行业类似。这一步强调入库指标的行业逻辑,在一定程度上保证了后续步骤程序输出的可解释性。

因为行业财务状况是中观景气建模的目标,所以上市公司定期财务指标不应该出现在行业指标库中,否则存在自我解释之嫌。不过,**国家统计局发布的国民经济核算体系口径的制造业行业财务指标**,以及基于这些财务指标计算得到的毛利率等复合指标,可以被纳入行业指标库。这是因为国家统计局的行业财务指标月频发布,时效性优于上市公司定期财务报表,因而对整个行业的景气度判断依然具有价值。

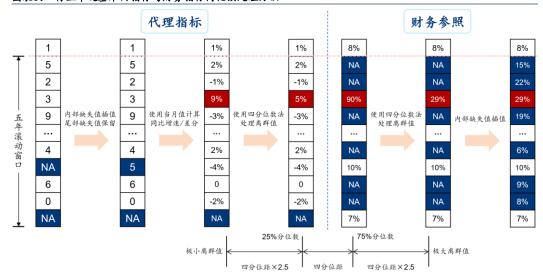


# 行业指标预处理: 简化步骤, 贴合财务特征

预处理的目标是得到**当月同比口径**的代理指标序列和财务指标序列。采用同比口径,一方面是考虑到市场投资者在进行决策时,更关注同比口径的财务指标,如净利润 TTM 同比增速;另一方面是因为做过一次同比变换后的指标,在平稳性上有所改善,可以减少指标评价步骤中伪回归的发生。

我们首先取 60 个月作为滚动窗口期。不过为了保证财务参照序列首尾都是有效值,便于通过线性插值得到月频序列,预处理多往前取了 1 个月数。对于财务参照,我们首先使用四分位数法识别并处理离群值(主要是净利同比数据,由于基数效应,离群值并不鲜见),然后再将季频指标插值成为月频指标。

对于代理指标,我们优先下载总量口径指标;如果没有总量口径指标,也可以考虑下载环比口径指标,因为环比口径指标可以通过累乘转换成总量口径指标。对于代理指标序列的内部缺失值(例如发电量指标每年1月都没有发布值),我们使用线性插值处理;对于代理指标序列的尾部缺失值(大多数时候由于指标延迟发布所致),为了不引入未来信息,暂时保留,在最后一个步骤中使用 Simple-Nowcasting 进行填充。



图表3: 行业中观基本面指标与财务指标简化预处理方法

资料来源: 华泰研究

在前期报告《行业配置策略:中观景气视角(1)》(2022-01-18)中,我们对总量类代理指标采用了较为复杂的预处理方法,主要包括季节性调整和 HP 滤波去趋势等两个步骤。但是经过我们的评估与试验发现,季节性调整和 HP 滤波更适用于宏观因子建模,而不太适用于中观行业景气建模,理由如下:

- 1) 季节性调整和 HP 滤波算法所需的运算量远大于上图所示的简化预处理方法。本研究涉及 20 多个行业,每个行业又有上百个指标需要在滚动窗口内进行预处理,简化方法能在较短的时间内(对于普通的笔记本电脑来说,通常在 1 小时以内)完成,使得后续的应用场景更具备可落地性。另外,建模使用同比口径的数据,同比变换也能够消除一部分季节效应。
- 2) HP 滤波能够将指标序列分解为趋势项和循环项。对于代理指标来说,趋势项刻画的是行业的成长性,循环项刻画的是行业的周期性。如果使用 HP 滤波消除趋势项,再用只有周期性的代理指标去预测既有周期性又有成长性的财务参照,财务参照中刻画成长性的部分就很难被预测。而事实上,成长性是投资者关注的重点之一。



# 指标评价和筛选:滚动窗口,捕捉逻辑变化

在每个财报秊结束时,我们基于口径统一的代理指标和财务参照序列,在 60 个月的滚动窗 口内评价两者之间的相关性和领先滞后性, 最终筛选出中观景气代理指标; 在非财报季月 份,我们将沿用最近财报季选中的代理指标。相比于前期报告中的人工筛选,程序对所有 的行业"一视同仁". 能够避免分析师对不同行业逻辑的认识和理解差异: 相比于全局使用 相同的代理指标建模,在滚动窗口内利用最新的财务数据更新中观景气代理指标,能及时 捕捉行业逻辑的变化。

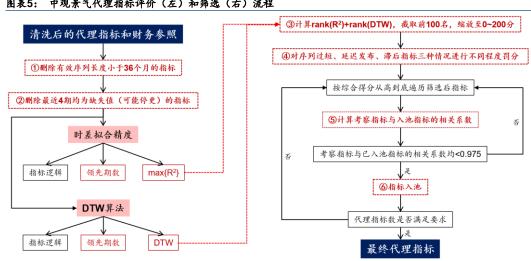
下载完整的指标库 下载完整的指标库 清洗、评价、筛选 清洗、评价、筛选 上年Q4财报 本年Q2财报 本年Q3财报 本年Q1财报 沿用最近财报季选中的代理指标

图表4: 中观景气代理指标更新频率说明

资料来源: 华泰研究

首先,我们删除有效序列长度小于36个月的指标和最近4期均为缺失值的指标。根据我们 的观察,如果一个指标连续 4 个月都没有发布值,该指标很可能停更,故丢弃该指标。然 后,基于时差拟合精度和动态时间弯曲 (DTW) 算法来同时评价代理指标相对于财务参照 的相关性和领先滞后性。接着,以相关性赋予指标底分,再结合有效序列长度、延迟发布 情况、领先滞后期数等罚分项,计算指标综合得分。最后,我们按照综合得分从高到低排 序,依次选出不重复的10个代理指标,用于中观景气建模。

虽然前期报告《行业配置策略:中观景气视角(1)》(2022-01-18)、《宏观因子指数回顾、改 进与应用》(2022-02-12)等已经对部分方法的细节进行介绍。但为了便于读者更好地理解 本文, 本文采用更通俗直观的方式对这些方法进行解释。



图表5: 中观景气代理指标评价(左)和筛选(右)流程

资料来源:华泰研究

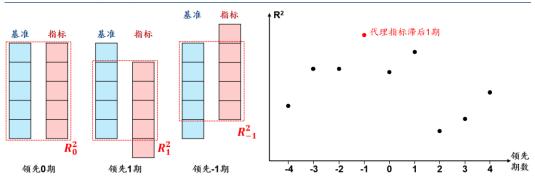


#### 时差拟合精度

我们在做相关性分析的时候,通常会把财报基准的 t 月和代理指标的 t 月进行对齐。这种对齐方式被称作"领先 O 期"。如果用财报基准的 t+k 月和代理指标的 t 月进行对齐,就称代理指标领先财报基准 k 期,因为我们使用早 k 月的代理指标去解释财报基准;类似地,如果用财报基准的 t-k 月和代理指标的 t 月进行对齐,就称代理指标滞后财报基准 k 期,或者说领先财报基准-k 期。

我们在代理指标领先财务参照±4期的范围内,分别对代理指标和财务参照开展一元线性回归,一共是9组试验。R²最大的那组试验,对应的领先期数就是滚动窗口内,代理指标相对于财务参照的**领先期数**; R²等于相关系数的平方,刻画了两者相关性;对应的回归系数的正负性就是滚动窗口内,代理指标对财务参照的驱动逻辑方向。

图表6: 时差拟合精度计算原理示意图



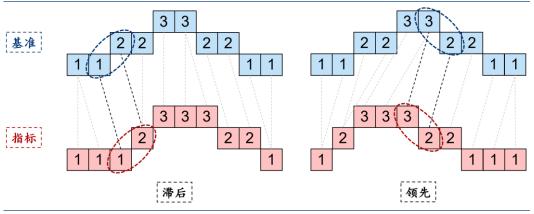
资料来源: 华泰研究

### DTW 算法

虽然时差拟合精度能够同时评价相关性和领先滞后性,但时差拟合精度的计算结果比较容易受到个别波动幅度比较大(如 zscore 在±2.0 附近)的样本的影响,而且时差拟合精度的计算建立在恒定的领先滞后关系的假设之上,事实上,滚动窗口内代理指标相对于财务参照的领先滞后关系可能是动态变化的。这是我们进一步引入 DTW 算法的原因。

下图展示了 DTW 算法的结果——虚线表示两条序列样本点的匹配结果。不难发现,与一元线性回归不同,两条序列样本点并非一一对应,而是根据**形状相似度**得到了一对多或者多对一的匹配关系。我们在图上用椭圆虚线标注了其中两组相同形状的匹配结果。

图表7: DTW 算法呈现的结果示意图



资料来源: 华泰研究



根据展示匹配结果的"虚线", 我们可以计算两个指标:

- 1) 月均 DTW 距离:每一条虚线的距离等于所连接的两个样本点的欧氏距离;所有虚线距离求和、再除以序列长度,就得到了月均 DTW 距离。月均 DTW 距离的数学含义是两条序列的差异相当于序列自身波动的多少倍。月均 DTW 距离数值越小,表明两条序列形状越相似;对于两条一模一样的序列,月均 DTW 距离等于 0。
- 2) 领先滞后期数:通过观察虚线倾斜的方向,还可以判断代理指标相对于财务基准的领先滞后关系。左图虚线整体往左倾斜,对应代理指标滞后于财务基准;右图虚线整体往右倾斜,对应代理指标领先于财务基准。每一条虚线的领先期数等于所连接的两个样本点时间刻度的差值;所有虚线领先期数的均值,就是滚动窗口内代理指标相对于财务基准的平均领先期数。实际上,DTW 算法给出的是领先滞后关系的动态变化。

在执行 DTW 算法之前,我们需要对两条序列进行 zscore 标准化。由于代理指标可能和财务参照正相关,也有可能和财务参照负相关,所以除了使用代理指标的原始序列执行 DTW 算法之外,还需要将代理指标序列乘以-1,再次执行 DTW 算法,并取两次计算结果中月均 DTW 距离更小的结果。如果更小的结果对应乘以-1 后的代理指标序列,则表明代理指标对财务参照的逻辑驱动方向为负;反之同理。

为了得到最优的"虚线",我们需要将 DTW 算法转换为求解等价问题——旅行商问题。如右图所示,有一个 T×T (T 等于代理指标或财务参照的序列长度) 的棋盘,每个格子(i,j)都有一个距离 d(i,j),d(i,j)等于财务参照的第 i 个样本点和代理指标的第 j 个样本点数值之差的绝对值。规定棋子只有三种移动方向"↑""→""J",棋子从棋盘左下角的起点(1,1)走到棋盘右上角的终点(T,T),有多条路径可选,每一条路径的距离等于所经过格子的 d(i,j)之和。旅行商问题的目标就是求最短路径。

一旦求出了最短路径,那么 DTW 算法就迎刃而解——左图中虚线连接的两个样本点的时间 刻度和右图中棋子所经格子的坐标是一一对应的。例如,棋子的最短路径经过了(i,j)这个点 (右图中已用白色×标注),于是,我们在左图中把财务参照的第 i 个样本点和代理指标的第 j 个样本点进行连接(左图中已用黑色虚线标注)。

图表8: DTW 算法与求解旅行商问题是等价的

资料来源:华泰研究

$$D(i,j) = d(i,j) + \min\{D(i-1,j), D(i,j-1), D(i-1,j-1)\}$$

有时候,为了限制代理指标和财务参照的单期领先滞后期数,或者为了降低运算量,我们可以引入|i-j|≤r的约束条件。例如,右图引入了|i-j|≤4的约束条件,意味着代理指标和财务参照的单期领先滞后期数被限制在了4期以内,棋盘的可达范围也相应缩小。由于本研究涉及上千个中观基本面指标,为了提升运算速度,我们令|i-j|≤12。

图表9: 用动态规划求解旅行商问题

资料来源: 华泰研究

## 指标综合评分

在删除有效序列长度小于 36 个月的指标和最近 4 期均为缺失值的指标之后,我们基于时差拟合精度和 DTW 算法评价了各代理指标的相关性和领先滞后性。接着,对 R<sup>2</sup>和月均 DTW 距离进行排序——R<sup>2</sup>是升序,值越大排名得分越高;月均 DTW 距离是降序,值越小排名得分越高。然后,将两者的排名得分相加,截取总排名得分最高的 100 个指标,并将总排名得分缩放至 0-200——排名第 1 的指标为 200 分,排名第 100 的指标为 0 分。 0-200 分衡量的是代理指标与财务参照的相关性。

以相关性得分为基础, 我们进一步考虑三个惩罚项:

- 1) 惩罚序列长度不足的指标:部分指标的起点晚于滚动窗口的起点。在滚动窗口内、指标起点之前,每存在1个前部缺失值,就扣1分。
- 2) 惩罚领先滞后期数偏离理想状况的指标:每个指标的领先滞后期数均被限制在了±4个月内。最理想的领先期数是3个月——领先4个月实则为至少领先4个月,指标领先性过强有可能导致配置信号过于左侧。各指标领先滞后期数与领先3个月之间,每相差1个月、扣3分。
- 3) 惩罚延迟发布的月频指标:对于延迟发布的月频代理指标(除 PMI 族以外的月频代理 指标几乎都有延迟发布的问题),会引入尾部缺失值,由于很难追溯历史上各指标的延 迟发布期数,对此一律扣 5 分(延迟发布既有滞后 1 期效果,又有缺失 1 期效果,而 且缺失值出现在更为重要的序列尾部,因此扣分比前两类惩罚叠加还要多)。

最后,对于每个行业、每个财务参照,我们将相关性最高的 100 个代理指标按综合得分从高到低排序,依次选取**不重复**(指标库中存在数值相同但名称、代码不同的重复指标,我们用 Pearson 相关系数 $\geq$ 0.975 即  $R^2\geq$ 95%"拦截"重复指标)的 10 个指标,作为中观景气代理指标。



# 景气指数生成: Simple-Nowcast, 使行业间可比

在得到中观景气代理指标的过程中,和财务参照之间的相关性是最重要的考虑因素,因此得到的中观景气代理指标必然含有能够解释财务参照的信息。于是,我们自然能够想到用主成分分析去提取这些指标的共同信息,即第一主成分,从而把低频滞后的财务参照"月频化"了。但是,由于代理指标可能存在尾部缺失值,如果直接使用主成分分析提取第一主成分,会导致第一主成分也会存在尾部缺失值,缺失的期数取决于尾部缺失最严重的代理指标。对此,我们改用"加强版"主成分分析——Nowcasting。

Nowcasting 的含义是对尚未发布的"非将来时"状态进行预测。如部分代理指标存在尾部缺失值,其对应的现实情况已成定数,只是数据没有及时披露。于是,可以通过 Nowcasting 填充这些尾部缺失值。前期报告《行业配置策略:中观景气视角(1)》(2022-01-18)提出了 Simple-Nowcasting 模型。 Simple-Nowcasting 模型的核心原理与海外原版 Nowcasting 模型相同,在结果准确性和模型收敛性方面没有显著削弱的基础上,在运算效率和可扩展性方面占优,对代理指标平稳性的要求也相对较低。

在前期报告《行业配置策略:中观景气视角(1)》(2022-01-18)中,我们直接使用评价筛选得到的代理指标构建中观景气打分体系,开展行业间的比较。然而,这种做法可能经不起推敲——不同行业的中观景气代理指标含义不同、共线性程度也不同,难以直接比较。而中观行业景气指数通过景气代理指标的降维,其含义又重新回归财务层面,大大提升了行业间的可比性。

#### 数学原理

模型的核心是**动态因子模型**,包括隐含状态方程、状态转移方程(又分为隐含因子状态转移方程和特质因子状态转移方程)等两类方程。

隐含状态方程左侧的 $y_t^i$ 是第 i 个景气代理指标的历史序列;方程右侧的 $f_t$ 是景气指数,因为无法观测,故被称作隐含因子; $b_i$ 是第 i 个代理指标在景气指数上的暴露; $c_i$ 是第 i 个代理指标的截距项,建模时如果对 $y_t^i$ 做了 zscore 标准化,则可以省略该项; $e_t^i$ 是第 i 个代理指标无法被景气指数解释的部分,故被称作特质因子。由于指标延迟发布,部分代理指标序列存在尾部缺失值,故用 0-1 变量 $w_t^i$ 标记尾部缺失情况,使方程在形式上成立:

$$y_t^i = \begin{cases} b_i f_t + c_i + e_t^i, & w_t^i = 1 \\ 0, & w_t^i = 0 \end{cases}, i = 1, 2, \dots, n$$

隐含状态方程的本质是降维——n 个景气代理指标降维到 1 个景气指数。从形式上看,隐含状态方程和主成分分析取第一主成分的过程是相同的。不过正如前文所提及,由于代理指标存在尾部缺失值,直接开展主成分分析会使得景气指数也存在尾部缺失值。

对此,模型使用状态转移方程对代理指标尾部缺失值进行合理预测。模型将代理指标分为两部分——能够被景气指数所解释的部分和特质因子,使用 AR(p)过程(本研究 p=2)分别预测景气指数的尾部缺失值(隐含因子状态转移方程)和特质因子的尾部缺失值(特质因子状态转移方程)。这里需要注意,AR(p)描述零均值的随机过程,不应带常数项:

$$\begin{split} \hat{f}_t &= a_1 f_{t-1} + a_2 f_{t-2} \\ \hat{e}_t^i &= h_1^i e_{t-1}^i + h_2^i e_{t-2}^i, i = 1, 2, \dots, n \end{split}$$

基于 $\hat{f}_t$ 和 $\hat{e}_t^i$ ,可以进一步计算代理指标尾部缺失值的预测值:

$$\hat{y}_t^i = b_i \hat{f}_t + c_i + \hat{e}_t^i, i = 1, 2, ..., n$$



#### 求解过程

模型并不存在解析解——三个方程只有隐含状态方程左侧的代理指标序列是已知的,待估计参数 $\{f_t,b_i,a,h^i\}$ 都是未知的(这里假设建模前对所有代理指标序列均进行了 zscore 标准化,因而省略 $c_i$ )。Simple-Nowcasting 采用 **EM 算法**来估计这些参数:

- 1) Expectation 步: 固定 $\{b_i,a,h^i\}$ ,用代理指标的预测值填充尾部缺失值后,再用主成分分析估计 $f_i$ ;
- 2) Maximization 步: 固定 $f_t$ , 使用最小二乘法依次估计 $\{b_i,a,h^i\}$ 。

以 EM 算法为内核, Simple-Nowcasting 首先设计了"小循环", 其目的是为了填充尾部缺失值、得到完整的景气指数序列的估计; 为了提升模型收敛性, 我们又在"小循环"外套了"大循环"。下文举了含有 4 个代理指标的例子来阐述求解过程, 这 4 个代理指标分别存在 3 期、2 期、1 期、0 期尾部缺失值。

第一步: 我们用代理指标的第一主成分初始化景气指数。注意,主成分分析的样本不包括含有缺失值的截面。因此,在本文的示例中,初始化后的景气指数只有第 1 期到 t-3 期有数值,由序列长度最短的 y<sup>1</sup> 决定。

第二步: 我们带着初始化后的景气指数进入"大循环"。而"大循环"将会反复调用"小循环"。"小循环"则使用 EM 算法,在填充缺失值的同时,不断修正拟合结果:

- 1) 使用最小二乘法估计各代理指标在景气指数上的暴露,并计算残差;因为回归只能使用截至 t-3 期的样本,得到的特质因子序列也只有截至 t-3 期的数值。
- 2) 使用 AR(2)过程拟合景气指数和各指标特质因子,得到自回归系数。
- 3) 将 AR(2)过程线性外推一期,得到 t-2 期的景气指数和各指标特质因子的预测值,进一步计算得到代理指标 t-2 期的预测值。
- 4) 用代理指标 t-2 期的预测值填充 t-2 期的缺失值,于是 t-2 期全体指标都有数值,可以加入主成分分析的样本,取第一主成分,得到截至 t-2 期的景气指数。
- 5) 重复上述步骤,直到能够得到截至 t 期的完整的景气指数序列。

需要注意的是,"小循环"每执行一次,除了景气指数会减少 1 期尾部缺失值(即被 1 期 Nowcasting 值填充)外,景气指数更早期的有效值也会发生变化。这是因为"小循环"每执行一次,参与主成分分析的样本数多了 1 期,导致代理指标在第一主成分中的权重发生了变化。因此,只有"小循环"无法保证模型的收敛性。

第三步: 我们回到"大循环",将景气指数 t-2 至 t 期的 Nowcasting 值还原为缺失值,重新执行"小循环"。与上一轮"小循环"不同的是,由于上一段文字描述的原因,再次进入"小循环"前景气指数的有效值是不同的。不断执行"大循环",当进入"小循环"前景气指数的有效值和离开"小循环"后景气指数的有效值相等时,模型收敛。

不管是海外原版解法,还是我们提出的 Simple-Nowcasting,两者的核心思想都是在第一主成分的基础上对尾部缺失值进行合理估计。这就是前文把 Nowcasting 称作"加强版"主成分分析的原因。

最后还需要强调一点:模型收敛的一个必要条件是全体代理指标两两之间显著相关。我们先来考虑一种比较极端的情况。假设有两个指标 y1 和 y2,两者不相关。那么, y1 和 y2 的相关系数矩阵就是单位阵,其两个特征值都为 1。根据主成分的定义,第一主成分和第二主成分的贡献率均为 50%。这就让模型"困惑"了——第一主成分究竟应该刻画 y1,还是刻画 y2?于是,"大循环"的结果可能会在 y1 和 y2 之间反复跳跃,无法收敛。



所以,对于指标评价和筛选过程筛出高相关性的代理指标,读者无需过度忧虑。而且,每个代理指标只反映了行业景气度的一个侧面,会受到不同程度噪声的干扰。显著相关的指标数目越多,其公共部分——真正被行业景气度驱动的部分能够被提取得更纯粹。在本研究中,指标评价和筛选过程以财务参照为"锚",最终确定的代理指标和财务参照指标通常都有较高的相关性,使得代理指标之间存在显著的间接相关性。

不过,在极个别情况中,在筛选得到的 10 个指标中,排名靠后的指标可能与财务参照的相关性较弱,被选中是"被迫顶上"的结果,导致模型无法收敛。遇到这种特殊情况,我们依次删除排在第 10 名、第 9 名……的代理指标,直至模型可以收敛。

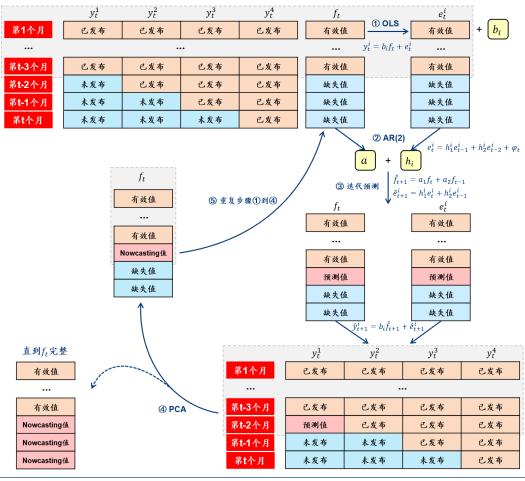
下文我们将应用中观行业景气指数开展景气观点判断、单行业择时、行业间轮动。

图表10: Simple-Nowcasting 之"大循环"

	$y_t^1$	$y_t^2$	$y_t^3$	$y_t^4$		$f_{t}$		$f_t$	
第1个月	已发布	已发布	已发布	已发布		有效值		有效值	
•••			••		_	•••	_	•••	
第t-3个月	已发布	已发布	已发布	已发布	① PCA	有效值	② 小循环	有效值	
第 <b>t-2</b> 个月	未发布	已发布	已发布	已发布		缺失值		Nowcasting值	
第 <b>t-1</b> 个月	未发布	未发布	已发布	已发布		缺失值		Nowcasting值	
第t个月	未发布	未发布	未发布	已发布	]	缺失值		Nowcasting值	
③ 缺失值还原									
				直到	f <sub>t</sub> 不再变化 ←				

资料来源:华泰研究

图表11: Simple-Nowcasting 之"小循环"



资料来源: 华泰研究



# 建模过程中可能遇到的问题解答

问题一:对于 Wind 上已经停止更新的指标,是否仍需要放入行业指标库?答:视研究目标而异。如果研究目标是检验模型在历史上的表现,那么当前停更的指标在历史的某一个时间段,有可能是行业景气状态的重要解释变量,此时仍然应该将其放入指标库。如果研究目的主要是为了对行业景气状态进行样本外跟踪,那么由于指标评价步骤会自动剔除停更指标,此时是否放入该变量对建模结果没有影响。

问题二:如果关心某些中信二级行业的景气指数,还需要重新构建指标库吗?答:无需重新构建指标库。在自上而下视角,二级行业指标库只是一级行业指标库的子集,而且二级行业之间可能也存在行业逻辑的联系。在开展二级行业中观景气建模时,我们可以沿用所属一级行业的指标库,额外计算二级行业的财务参照,再执行同样的建模步骤即可。

问题三:在计算行业财务指标时,如何处理新上市或者所属行业变更的成分股?答:不纳入计算范围。首先,本研究使用的行业财务参照都是同比口径的,为了保证当期财务指标和比较期财务指标可比,计算当期数值使用的成分股和计算比较期数值使用的成分股必须一致。其次,对于新上市或者所属行业变更的企业,其财务报表只追溯调整一年(其中股东权益表只有年报会追溯调整),不足以计算比较期 TTM 指标。

问题四:主成分分析时,如何应对代理指标权重的正负性与指标逻辑不一致?答:在主成分分析时,全体指标权重都乘以-1 依然是合理的结果,只不过权重的正负性与指标逻辑不一致(指标对行业景气状态是正的贡献,但权重却是负的),导致输出的景气指数与真实景气状态可能恰好相反。有两种方式可以控制指标权重的正负性。第一种方法,我们可以计算输出结果与可得财务参照的相关系数,如果相关系数为负,就需要对全体指标权重都乘以-1;第二种方法,指标评价过程能够同时获得滚动窗口内的指标逻辑,当大多数指标逻辑和指标权重正负性不一致时,就需要对全体指标权重都乘以-1。本研究采用的是第二种方法。



# 应用一:景气状态判断

中观行业景气系列的定位是使用高频、及时的行业基本面指标,对低频、滞后的行业财务数据的走势作出及时预判。所以,中观行业景气指数最基础的应用就是用于行业景气状态判断。例如,站在2022-03-31,2021Q4和2022Q1的真实财报尚未披露完毕。我们可以沿用2021-10-31(对应2021Q3财报披露完毕)选中的景气代理指标进行建模,得到截至2022-03的景气指数。以钢铁行业为例,滚动区间2016-10-31至2021-09-30筛选得到的中观景气代理指标如下表所示:

图表12: 钢铁行业滚动窗口(2021-10-31)筛选得到的中观景气代理指标

参照:净利同比	参照: ROE 同比	参照:营收同比	参照:毛利同比
期货结算价(连续):螺纹钢	MySpic 指数:热卷	MySpic 指数:长材	黑色金属冶炼及压延加工业:毛利率
出口金额:铁矿砂及其精矿:当月值	Myspic 综合钢价指数	PPI:黑色金属冶炼及压延加工业:环比	PMI:产成品库存
PMI:产成品库存	PMI:生产经营活动预期	PPIRM:黑色金属材料类:环比	富宝:废钢价格指数:全国
MySpic 指数:长材	黑色金属冶炼及压延加工业:毛利率	期货结算价(连续):螺纹钢	钢铁 PMI:生产
钢铁 PMI:新订单	PMI:进口	MySpic 指数:冷板	MySpic 指数:热卷
库存:铁矿石:港口合计	PPI:黑色金属冶炼及压延加工业:环比	钢铁 PMI:原材料购进价格	Myspic 综合钢价指数
PMI:进口	PMI:在手订单	南华热轧卷板指数	PMI:生产经营活动预期
钢铁 PMI:生产	黑色金属冶炼及压延加工业:亏损企业单位数	南华螺纹钢指数	PMI:进口
南华螺纹钢指数	南华热轧卷板指数	出口价格指数(SITC2):钢铁	产量:铁矿石原矿量:累计值
出口价格指数:黑色金属冶炼业及压延加工	螺纹价格指数	期货结算价(连续):热轧卷板	PMI:采购量

资料来源: Wind, 华泰研究

下图中,深蓝色线是钢铁行业真实财务指标,站在 2022-03-31,只能获取截至 2021-09-30 (对应 2021Q3)的数据;深红色线是钢铁行业景气指数,更新到了 2022-03-31。从相对位置来看,除营收景气指数外,zscore 值均下降至零轴以下;从边际变化看,可以判断钢铁行业 2022Q1 的景气状况相对于 2021Q4 边际显著恶化。

图表13: 钢铁行业净利景气指数与净利同比对比(2022-03-31)



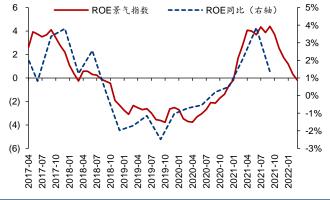
资料来源: Wind, 华泰研究

图表14: 钢铁行业营收景气指数与营收同比对比 (2022-03-31)



资料来源: Wind, 华泰研究

图表15: 钢铁行业 ROE 景气指数与 ROE 同比对比(2022-03-31)



资料来源: Wind, 华泰研究

图表16: 钢铁行业毛利景气指数与毛利同比对比(2022-03-31)





# 应用二:单行业择时

市场在交易景气的时候有两种逻辑: (1) 景气相对位置是高位还是低位; (2) 景气边际变化是正向还是负向。据此,我们基于中观行业景气指数生成了 4×4=16 个因子 (4 种财务参照和 4 种信号模式两两组合)。其中,模式 orig 对应第一种逻辑,称作第一类因子;模式 mom/moma3/qoq 对应第二种逻辑,称作第二类因子。

#### 图表17: 基于中观行业景气指数构建行业因子

财务参照

信号模式

净利景气 ROE景气 营收景气

毛利景气

模式orig: 行业zscore景气指数

模式mom: 行业zscore景气指数环比变化

模式moma3: 行业zscore景气指数减去过去三月均值

模式qoq: 行业zscore景气指数近三月均值减去再往前三月均值

资料来源: 华泰研究

中观行业景气指数的第二项应用是用上述信号对行业相对万得超额净值指数开展择时。使用超额净值指数是为了排除大盘 beta 的影响。下图以酒类和建材行业为例,对比了原指数与超额净值指数的趋势差异。

图表18: 酒类行业相对万得全 A 超额净值指数



资料来源: Wind, 华泰研究

图表19: 建材行业相对万得全 A 超额净值指数



资料来源: Wind, 华泰研究

对于已覆盖的每个行业,依次使用 16 个因子,对行业超额指数开展 16 次择时:若因子大于 0,满仓配置;反之,持有现金。每一次择时都记录择时策略相对于行业超额指数的超额年化收益,并开展统计。回测区间为 2016-04 至 2022-06,不考虑手续费。

结果显示,大多数因子在大多数行业上择时有效。从因子视角来看,有 9 个因子在 2/3 以上的行业中通过择时获得了正的超额收益,有 15 个因子在 1/2 以上的行业中通过择时获得了正的超额收益,各因子在全体行业上的平均超额年化收益均大于 0。其中,无论是从超额为正的比例,还是从平均超额年化收益来看,净利景气指数的第二类因子都位居前列,可能和主动投资者主要根据净利同比的环比变化进行决策有一定关联。毛利景气指数的两类因子总体表现靠后,这可能和毛利同比预测难度较大有关——其他三个维度跟量和价均有关系,可供选择的基本面指标数量大幅减少。下文将呈现毛利同比的预测效果,能够佐证上述结论。

图表20: 使用中观行业景气指数开展单行业择时分因子表现统计

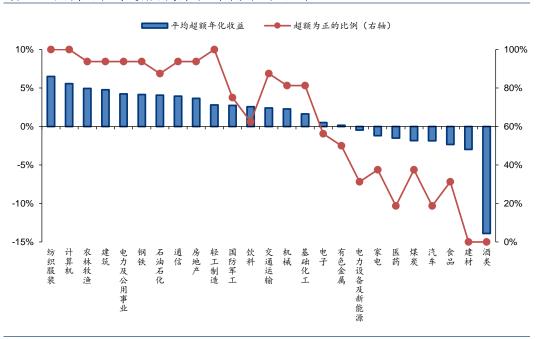


注:标的是中信行业指数相对于万得全 A 的超额净值指数

资料来源: Wind, 华泰研究

从行业视角来看,16个信号平均超额年化收益录得正值的比例为17/25。排名靠前的主要是业绩呈现周期性波动的行业,如石油石化、钢铁、建筑、农林牧渔等;录得负值的更多是业绩具有一定成长性的行业,如酒类、电力设备及新能源、医药、汽车等。其中,有两个行业16个信号均未能取得正的超额收益——建材和酒类。酒类行业可能得益于其业绩的稳定性和成长的确定性,在回测区间内能够"穿越牛熊"、持续获得超额收益,如图表18所示,空仓一段时间有较大可能会错失一段超额收益。而建材行业的超额净值在很窄的范围内波动,如图表19所示,并没有呈现出典型顺周期行业持续半年到数年尺度的大波段特征,给择时增加了难度。

图表21: 使用中观行业景气指数开展单行业择时分行业表现统计



注: 标的是中信行业指数相对于万得全 A 的超额净值指数

资料来源: Wind, 华泰研究

使用中观行业景气指数择时更适合业绩呈现周期性波动的行业,可能跟景气指数与财务参照的拟合优度较高有一定关系。在回测区间内的每个 3/6/9/12 月月末,我们都计算了滚动窗口内景气指数与 Q1/Q2/Q3/Q4 真实财务参照的  $R^2$ ; 实际上,因为 3/6/9/12 月月末无法获取最新的 Q1/Q2/Q3/Q4 财务数据,所以  $R^2$ 一定程度上反映了预测准确性; 然后,计算回测区间内  $R^2$ 的均值。



从财务维度来看,毛利同比维度的预测效果最差,一定程度上导致基于毛利景气指数生成的因子择时效果较差。从行业维度来看,诸如电力设备及新能源、国防军工等偏成长风格的行业拟合效果较差,而偏周期风格的行业拟合优度普遍在50%以上。可能原因包括:

- 1) 偏周期风格的行业业绩呈现周期性波动,使得 Nowcasting 模型中的状态转移方程更容易拟合;
- 2) 偏周期风格的行业产业链结构相对稳定,而偏成长风格的行业产业链结构更容易发生变化,很难保证景气代理指标在整个滚动窗口内都和财务参照高相关。
- 3) 偏周期风格的行业更易获得序列较长、披露及时、逻辑相关的基本面指标,偏成长风格的行业在 Wind 上的基本面指标数量较少。

在上述三个原因中, 第三个原因有可能通过切换高质量数据源得到改善, 但第一个和第二个原因在本文提出的中观景气建模框架中暂时无法得到解决。

图表22: 中观景气指数与真实财务参照的平均拟合优度(R2)

行业	净利同比维度	ROE 同比维度	营收同比维度	毛利同比维度
石油石化	76.51%	56.57%	85.06%	45.92%
煤炭	62.43%	40.66%	76.29%	73.10%
有色金属	86.23%	88.70%	67.40%	60.21%
钢铁	72.41%	70.63%	83.92%	56.25%
基础化工	52.04%	33.83%	64.23%	19.98%
建材	61.24%	66.48%	86.34%	40.68%
电力及公用事业	50.30%	45.55%	78.88%	77.22%
交通运输	67.62%	55.61%	80.55%	50.53%
建筑	57.86%	33.79%	54.18%	33.70%
机械	65.26%	32.08%	66.76%	19.98%
电力设备及新能源	46.76%	36.94%	31.23%	24.09%
国防军工	26.10%	16.44%	30.11%	23.79%
汽车	70.65%	41.10%	59.82%	26.28%
家电	50.77%	38.98%	71.68%	25.59%
酒类	76.31%	39.00%	75.76%	40.39%
饮料	47.95%	30.42%	63.07%	6.84%
食品	38.19%	10.29%	66.86%	49.91%
纺织服装	73.51%	30.71%	64.10%	34.58%
轻工制造	63.37%	30.55%	80.90%	30.69%
农林牧渔	63.49%	54.57%	77.45%	47.30%
医药	61.75%	48.69%	36.40%	43.07%
房地产	69.19%	47.10%	62.99%	58.17%
电子	62.18%	39.60%	75.11%	49.43%
计算机	53.08%	26.44%	50.54%	22.39%
通信	34.37%	27.77%	60.24%	26.60%

注:红色标注了平均 R<sup>2</sup>大于 50%的结果

资料来源: Wind, 华泰研究

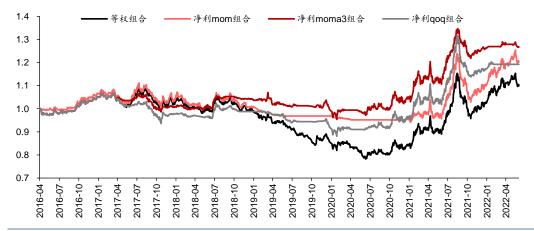
既然中观景气建模框架更适合偏周期风格的行业, 我们不妨构建**顺周期风格择时策略**,进一步考察中观行业景气指数对顺周期风格的择时效果。择时方案如下:

- 1) 择时对象包括石油石化、煤炭、有色金属、钢铁、基础化工、建材、机械、建筑等 8 个顺周期风格行业。
- 2) 择时信号采用单行业择时效果较优的净利景气指数的第二类因子:当≥5个行业发出看多信号时,顺周期风格行业可能景气共振、存在风格层面 beta 机会;于是,全仓配置即8个行业各配置12.5%。当发出看多信号的行业数不足5个时,顺周期风格可能不存在风格层面 beta 机会;于是,将仓位控制在半仓以下、只对每个发出看多信号的行业配置12.5%。
- 3) 回测区间为 2016-04 至 2022-06, 不考虑手续费。



结果显示,在 3 个净利景气指数的第二类因子中,基于净利\_moma3 因子构建的策略效果最佳。该策略基本抓到了顺周期风格最近的两波大行情,在 2016-04 至 2017-02 和 2020-06 至 2021-06 均选择了满仓。最近一波行情于 2021 年 9 月中旬见顶,策略在 2021 年 7 月开始减仓,可见净利\_moma3 信号略偏左侧。遗憾的是,该策略没有抓到 2022 年初的煤炭和石油石化行业行情,仅在 2022-04 和 2022-06 分别配置了煤炭和石油石化行业,其余月份空仓。可能原因是在 2021-11 至 2022-04 财报真空期中,俄乌战争、停止进口澳煤等事件改变了部分顺周期行业的逻辑,而中观景气建模所需指标沿用 2021Q3 指标评价筛选的结果,模型未能及时学习到行业逻辑变化。

图表23: 顺周期风格择时策略净值曲线



注: 择时标的是中信行业指数相对于万得全 A 的超额净值指数

资料来源: Wind, 华泰研究

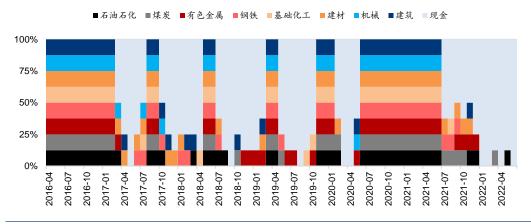
图表24: 顺周期风格择时策略业绩表现

组合名称	年化收益	年化波动	夏普比率	最大回撤	卡玛比率
顺周期风格行业等权组合	1.58%	9.74%	0.16	-28.65%	0.06
顺周期风格行业择时策略(净利_mom)	3.11%	8.07%	0.39	-16.98%	0.18
顺周期风格行业择时策略(净利_moma3)	4.01%	7.19%	0.56	-10.86%	0.37
顺周期风格行业择时策略(净利_qoq)	2.97%	7.68%	0.39	-17.65%	0.17

注:择时标的是中信行业指数相对于万得全 A 的超额净值指数

资料来源: Wind, 华泰研究

图表25: 顺周期风格行业择时策略(净利\_moma3)持仓变化



注: 择时标的是中信行业指数相对于万得全 A 的超额净值指数



# 应用三: 行业间轮动

单行业择时的应用只用到了因子正负性信息,没有用到因子数值大小信息。本节将利用截面上各行业因子数值的相对大小信息,开展行业间轮动。首先,我们对 16 个因子开展分层测试,按因子数值相对大小将已覆盖的 25 个行业分为 8/8/9 三层,对每一层开展回测。然后,根据分层测试结果,精选有效因子计算中观景气得分。因子精选规则如下:

- 1) 第一层组合相对于行业等权组合的净值比长期向上——能持续获得超额;
- 2) 第一层组合和第三层组合净值走势的区分度显著——超额不是源于运气;
- 3) 第一类因子和第二类因子对策略的贡献度相同——交易景气相对位置和交易景气边际变化的逻辑对策略的贡献度相同,我们暂时未找到合适的方法去预测市场偏好哪一种逻辑,故不希望策略在某一种逻辑上存在净暴露。

# 选行业因子分层测试结果

由净利景气指数生成的 4 个因子都能满足因子精选规则中的第一条和第二条。考虑到同一个财务维度的 3 个第二类因子可能存在高度共线性,我们只能选择其中 1 个进入精选因子池。考虑到净利\_mom 第一层的超额年化收益最高,多空年化收益高达 13.68%,净利同比维度最终选择了净利\_orig 和净利\_mom 等两个因子。

图表26: 净利景气指数四种模式因子分层测试结果汇总

信号模式		第一层第			第二层		
	超额年化	超额夏普	超额年化	超额夏普	超额年化	超额夏普	
orig	5.01%	0.78	0.16%	0.03	-4.87%	-0.93	
mom	8.33%	1.56	-2.19%	-0.47	-5.35%	-1.17	
moma3	5.59%	1.06	2.19%	0.46	-6.83%	-1.45	
qoq	3.71%	0.67	0.67%	0.13	-4.12%	-0.85	

资料来源: Wind, 华泰研究

图表27: 净利景气指数模式 orig 分层测试结果



资料来源: Wind, 华泰研究

图表28: 净利景气指数模式 mom 分层测试结果





#### 图表29: 净利景气指数模式 moma3 分层测试结果



资料来源: Wind, 华泰研究

## 图表30: 净利景气指数模式 qoq 分层测试结果



资料来源: Wind, 华泰研究

由 ROE 景气指数生成的 4 个因子中,只有 ROE\_orig 能够满足因子精选规则中的第一条和第二条,3 个第二类因子效果差强人意。例如,表现最优的 ROE\_qoq 多空年化收益仅有6.86%,比净利同比维度表现最差的第二类因子——净利\_qoq 的多空年化收益 7.83%还要低。这与真实市场中投资者决策习惯较为一致——通常更关注净利同比的环比变化而较少关心 ROE 同比的环比变化。ROE 同比维度只选择 ROE\_orig 进入精选因子池。

图表31: ROE 景气指数四种模式因子分层测试结果汇总

信号模式	第一层		第一层第二层			第一层第二层			第三层
	超额年化	超额夏普	超额年化	超额夏普	超额年化	超额夏普			
orig	7.41%	1.29	-4.74%	-0.90	-2.47%	-0.51			
mom	2.76%	0.48	-1.18%	-0.22	-1.72%	-0.33			
moma3	3.10%	0.56	-0.42%	-0.08	-2.67%	-0.55			
qoq	3.16%	0.58	0.76%	0.14	-3.70%	-0.72			

资料来源: Wind, 华泰研究

图表32: ROE 景气指数模式 orig 分层测试结果



资料来源: Wind, 华泰研究

图表33: ROE 景气指数模式 mom 分层测试结果





#### 图表34: ROE 景气指数模式 moma3 分层测试结果



图表35: ROE 景气指数模式 qoq 分层测试结果



资料来源: Wind, 华泰研究

由营收景气指数生成的 4 个因子中, 属于第一类因子的营收\_orig 和属于第二类因子的营收\_qoq 均满足因子精选规则中的第一条和第二条,故均进入精选因子池。

图表36: 营收景气指数四种模式因子分层测试结果汇总

信号模式		第一层		第二层	第三层	
	超额年化	超额夏普	超额年化	超额夏普	超额年化	超额夏普
orig	6.58%	1.05	-3.55%	-0.59	-2.96%	-0.54
mom	3.27%	0.56	-1.28%	-0.24	-1.97%	-0.39
moma3	4.98%	0.80	-3.24%	-0.61	-1.74%	-0.36
qoq	5.57%	0.88	-0.94%	-0.18	-4.38%	-0.80

资料来源: Wind, 华泰研究

图表37: 营收景气指数模式 orig 分层测试结果



资料来源: Wind, 华泰研究

图表38: 营收景气指数模式 mom 分层测试结果



资料来源: Wind, 华泰研究

图表39: 营收景气指数模式 moma3 分层测试结果



资料来源: Wind, 华泰研究

图表40: 营收景气指数模式 qoq 分层测试结果





由毛利景气指数生成的 4 个因子,在超额持续性和多空区分度上,均不如前文已经选中的因子。这有可能是因为真实市场中,投资者通常更关注净利/营收同比的环比变化,而较少关心毛利同比的环比变化;也可能跟前文提及的毛利同比预测难度较大有关。

图表41: 毛利景气指数四种模式因子分层测试结果汇总

信号模式		第一层		第二层		第三层
	超额年化	超额夏普	超额年化	超额夏普	超额年化	超额夏普
orig	2.13%	0.38	-0.18%	-0.03	-2.03%	-0.46
mom	-0.44%	-0.08	1.89%	0.36	-1.68%	-0.35
moma3	0.31%	0.06	3.33%	0.65	-3.43%	-0.74
qoq	4.73%	0.91	-2.60%	-0.49	-2.17%	-0.47

资料来源: Wind, 华泰研究

#### 图表42: 毛利景气指数模式 orig 分层测试结果



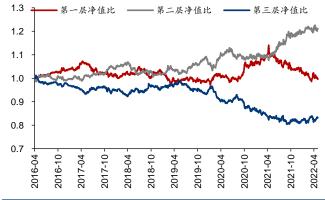
资料来源: Wind, 华泰研究

### 图表43: 毛利景气指数模式 mom 分层测试结果



资料来源: Wind, 华泰研究

### 图表44: 毛利景气指数模式 moma3 分层测试结果



资料来源: Wind, 华泰研究

图表45: 毛利景气指数模式 qoq 分层测试结果



资料来源: Wind, 华泰研究

# 中观景气得分与行业轮动策略

根据因子分层测试结果,我们最终选取了3个第一类因子——净利\_orig、ROE\_orig、营收\_orig和2个第二类因子——净利\_mom、营收\_qoq,采用**截面排序加权求和**,计算中观景气得分。本小节暂时不考虑因子权重优化,采用最简单的等权设置——3个第一类因子刻画了景气相对位置逻辑,权重均为1/3;2个第二类因子刻画了景气边际变化逻辑,权重均为1/2。这样做使得第一类因子和第二类因子的权重之和均为1,保证两种交易逻辑对策略的贡献度相等。行业i中观景气得分计算如下:

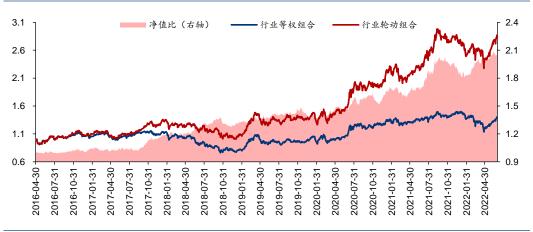
 $score_i = [rank(净利\_orig_i) + rank(ROE\_orig_i) + rank(营收\_orig_i)]/3 + [rank(净利\_mom_i) + rank(ROE\_qoq_i)]/2$ 

回测区间为 2016-04 至 2022-06, 月末选择中观景气得分最高的 4 个行业,次月第一个交易日调仓,暂不考虑手续费。



回测区间内,中观景气轮动策略每年都能够跑赢行业等权组合,相对于行业等权组合的净值比呈上升趋势;策略年化收益 19.04%,相比行业等权组合提升了 13.32 pct,最大回撤为-25.91%,相比行业等权组合下降了 8.15 pct; 月单边换手率 36%。

图表46: 中观景气轮动策略 (覆盖 25 个行业) 净值曲线



资料来源: Wind, 华泰研究

图表47: 中观景气轮动策略(覆盖25个行业)业绩表现

组合名称	年化收益	年化波动	夏普比率	最大回撤	卡玛比率
行业等权组合	5.72%	19.45%	0.29	-34.06%	0.17
中观景气轮动策略	19.04%	22.55%	0.84	-25.91%	0.73

资料来源: Wind, 华泰研究

图表48: 中观景气轮动策略 (覆盖 25 个行业) 分年度业绩对比

统计区间	ŧ	<b>亍业等权组合</b>		中观景气轮动策略		
	<b>年化收益</b>	夏普比率	最大回撤	年化收益	夏普比率	最大回撤
2016-04 至 2016-12	9.02%	0.50	-9.43%	10.62%	0.57	-9.32%
2017年	2.63%	0.21	-11.50%	18.57%	1.26	-12.43%
2018年	-30.45%	-1.39	-33.20%	-22.59%	-0.97	-25.91%
2019年	31.84%	1.54	-16.77%	40.02%	1.73	-15.77%
2020年	27.20%	1.18	-14.30%	51.12%	1.83	-16.34%
2021 年	12.38%	0.80	-9.55%	26.59%	1.18	-14.55%
2022-01 至 2022-06	-14.36%	-0.60	-24.33%	8.36%	0.34	-17.06%

资料来源: Wind, 华泰研究

如前文所述,对于市场偏好交易景气相对位置还是偏好交易景气边际变化,我们暂时无法给出答案,但我们可以从两类因子对策略的贡献度上窥知一二。我们只用第一类因子和第二类因子分别构建了一个轮动策略。从净值曲线来看,回测区间前半段第一类因子表现更优,回测区间后半段第二类因子表现更优。从业绩表现来看,第二类因子取得的长期业绩更优、年化收益高出 2.17 pct,表明市场大多数时候青睐景气边际变化逻辑。

我们还发现,第一类因子和第二类因子叠加之后,回测区间前半段策略的净值曲线与前半段表现更优的第一类因子的净值曲线更贴合,回测区间后半段策略的净值曲线与后半段表现更优的第二类因子的净值曲线更贴合。这说明两类因子叠加具有增强效果。



### 图表49: 只有第一类因子或第二类因子的中观景气轮动策略净值曲线



资料来源: Wind, 华泰研究

图表50: 只有第一类因子或第二类因子的中观景气轮动策略业绩表现

组合名称	年化收益	年化波动	夏普比率	最大回撤	卡玛比率
中观景气轮动策略(5因子)	19.04%	22.55%	0.84	-25.91%	0.73
中观景气轮动策略 (第一类因子)	16.63%	22.84%	0.73	-28.87%	0.58
中观景气轮动策略 (第二类因子)	18.80%	21.92%	0.86	-25.94%	0.72

资料来源: Wind, 华泰研究

# 改变因子权重对策略表现的影响

回顾一下中观景气得分的计算方式:

$$score_i =$$

$$w_1 \times rank$$
(净利\_orig<sub>i</sub>) +  $w_2 \times rank$ (ROE\_orig<sub>i</sub>) +  $w_3 \times rank$ (营收\_orig<sub>i</sub>) +  $w_4 \times rank$ (净利\_mom<sub>i</sub>) +  $w_5 \times rank$ (ROE\_qoq<sub>i</sub>)

前文策略为了保证交易景气相对位置和交易景气边际变化的逻辑对策略的贡献度相同,简单地采用等权方式设置了因子权重,即 W1、W2、W3 取 1/3, W4、W5 取 1/2。部分读者可能会担心这样的权重设置方式取得不错的业绩表现纯粹是因为运气,还是因为策略的鲁棒性较强。对此,我们通过随机实验,不断改变因子权重,考察策略表现的变化。

随机实验开展 100 组。权重的抽样空间是: W1、W2、W3、W4、W5 均≥25%, 且仍需满足两种交易逻辑对策略贡献度相同的约束,即 W1+W2+W3=100%, W4+W5=100%。回测条件与前文策略保持一致。

结果显示,表现最优的组合年化收益 22.72%,夏普比率 1.00,相对于行业等权组合的超额年化收益高达 17.00 pct;表现最差的组合年化收益 16.07%,夏普比率 0.74,相对于行业等权组合的超额年化收益也有 10.35 pct。而因子权重等权设置的轮动策略,在 100 组随机实验中仅排在第 60 名的位置。

图表51: 100 组因子权重随机实验的净值曲线范围



资料来源: Wind, 华泰研究

图表52: 100 组因子权重随机实验的业绩表现统计

业绩指标	最优组合	最差组合	中位数组合	因子等权组合
年化收益	22.72%	16.07%	19.30%	19.04%
夏普比率	1.00	0.74	0.86	0.84

资料来源: Wind, 华泰研究

可见**中观景气轮动策略的鲁棒性较强**。虽然通过权重优化,能够得到表现更优的策略,但 权重优化不是本文的重点,而且易增加策略的过拟合风险,所以本文最终依然采用因子权 重等权设置的轮动策略。

文末给出了自 2022-01 以来策略的持仓情况:

图表53: 中观景气轮动策略 (覆盖 25 个行业) 2022-01 以来持仓

月份	行业一	行业二	行业三	行业四
2022-01	有色金属	石油石化	电力设备及新能源	医药
2022-02	有色金属	电力设备及新能源	电力及公用事业	石油石化
2022-03	电力设备及新能源	饮料	煤炭	医药
2022-04	电力设备及新能源	汽车	煤炭	饮料
2022-05	医药	国防军工	煤炭	石油石化
2022-06	国防军工	石油石化	基础化工	医药
2022-07	国防军工	医药	农林牧渔	石油石化

资料来源: Wind, 华泰研究

# 总结与思考

本文从行业指标库出发,通过指标预处理、评价和筛选,得到各行业各财务维度不同时期的景气代理指标,通过 Simple-Nowcasting 生成中观行业景气指数,再根据景气指数生成两类选行业因子,然后精选分层测试效果理想的因子计算中观景气得分,据此构建行业轮动策略。这种思路是我们在权衡了数据挖掘和产业链建模的利弊之后提出的。

从中观基本面指标做行业轮动,有两大分支可循。第一大分支是数据挖掘,通过构建从基本面指标到股价的端到端模型,直接寻找驱动股价的基本面因素。第二大分支是产业链建模,通过梳理产业链结构和利润分配关键环节,寻找每一个环节最具有代表性的基本面指标,再综合得到景气分析框架。

**数据挖掘**的优势是可验证性较好,能够通过滚动窗口的训练和预测,根据策略跟踪结果来 判断模型的有效性;缺点是可解释性较弱,一旦策略失效,我们难以区分源于模型结构的 永久性失效,还是源于近期股价不受行业景气状况驱动的暂时性失效。



产业链建模的优势是可解释性较好,行业景气状况的变化可以穿透至每一个关键环节指标的变动,为信号的来龙去脉提供直接证据;缺点是可验证性较弱,模型可能只能阶段性有效,一旦产业链发生变化,模型需要人为进行调整,难以进行长期地回测与跟踪。

数据挖掘和产业链建模看似是"鱼和熊掌不可兼得"的关系,但是通过融入有逻辑可解释的模型结构,将数据挖掘的一部分"黑箱"转化成"白箱",有可能构建出兼具可解释性和可验证性的模型。本研究提出的框架是数据挖掘和产业链建模的折中,是一种兼具可解释性和可验证性的尝试。在生成景气指数之前,模型存在数据挖掘成分——筛选相对于财务参照相关性较高、领先性较强的指标;在应用于行业轮动时,模型又融入了产业链建模元素——用市场公认的财务指标去解释股价。这个框架的优势是,模型看好某个行业,我们容易得知是哪个财务维度、哪种交易逻辑发出了看多信号;模型的"黑箱"成分被压缩至了指标评价和筛选环节。

当然,只要模型有数据挖掘环节,不可解释的成分难以避免。以国防军工行业为例,我们统计了回测区间内各基本面指标被选中次数,发现模型偶尔会选出一些关注度较低、跟行业景气状况直接逻辑不清晰的指标,如"进口平均单价:自动驾驶仪:当月值",不排除伪相关的可能性。对此,模型通过提升数据挖掘的可靠性来尽可能排除伪相关,例如在评价相关性和领先滞后性时,同时应用了时差拟合精度和DTW 算法。

图表54: 回测区间内国防军工行业基本面指标被选中次数

编号	指标名称	选中次数	wind 代码	类别	使用口径
1	PMI	44	M0017126	扩散	同比增速
2	PPI:航空、航天器及设备制造:当月同比	39	M6409666	同比	原始
3	PPI:铁路、船舶、航空航天和其他运输设备制造业:环比	36	M0096821	环比	同比增速
4	中国船舶交易价格指数:综合	32	S6010821	价格	同比增速
5	工业增加值:铁路、船舶、航空航天和其他运输设备制造业:当月同比	31	M0000080	同比	原始
6	进口价格指数(SITC2):陆路车辆(包括气垫式)	29	M0045001	同比	原始
7	进口均价:电容器:当月值	29	T8577733	价格	同比增速
8	产量:发动机:当月值	27	S0027858	总量	同比增速
9	PMI:装备制造业	25	M6642295	扩散	同比增速
10	出口交货值:铁路、船舶、航空航天和其他运输设备制造业:累计值	23	M0059073	总量	同比增速
64	进口金额:片式钽电容器(85322110):当月值	1	K9022279	总量	同比增速
65	进口金额:碳纤维(68159920):当月值	1	S5917147	总量	同比增速
66	进口数量:片式多层瓷介电容器(85322410):当月值	1	O4333519	总量	同比增速
67	造船完工量:中国:累计值	1	S6000029	总量	同比增速
68	进口平均单价:自动驾驶仪(90142010):当月值	1	K4368259	价格	同比增速
69	新接船舶订单量:中国:累计值	1	S6000031	总量	同比增速
70	特种陶瓷制品制造:利润总额:累计值	1	S0013762	总量	同比增速
71	进口数量:碳纤维(68159920):当月值	1	S5917143	总量	同比增速
72	进口价格指数:(HS4):雷达设备、无线电导航设备及无线电遥控设备	1	S0058470	同比	原始
73	出口金额:航空或航天导航仪器及装置(罗盘除外)(901420):当月值	1	U7519356	总量	同比增速
	IAP I DETERMINE				

资料来源: Wind, 华泰研究

#### 风险提示

- 1. 模型根据历史规律总结,历史规律可能失效;尤其是财报真空期,即每年的 11 月至次年 4 月,如果行业逻辑在这段时期发生显著变化,选中的中观景气代理指标对于财务状况的相关性和领先性可能会被削弱,从而对模型表现产负面生影响。
- 2. 报告中涉及到的具体行业只是为了便于阐述模型的特点及其应用场景,不代表任何投资意见,请投资者谨慎、理性地看待。



# 免责声明

#### 分析师声明

本人,林晓明,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。本报告仅供本公司及其客户和其关联机构使用。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。
- 有关重要的披露信息,请参华泰金融控股(香港)有限公司的网页 https://www.htsc.com.hk/stock disclosure 其他信息请参见下方 "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

# 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2022年华泰证券股份有限公司

#### 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