

基于企业营收相似性的选股策略

专题报告

“蓝海启航”系列研究之八

本报告采用秩鼎公司提供的营收拆解标准化数据，计算 A 股市场上市公司经营业务的相似性，构建了营收相似相对动量因子，并从信息来源和经营业务复杂度两个维度对该因子进行增强。进一步，我们创新性地提出不同经营业务之间可能也存在相关性，并构建了考虑业务相关性的营收相似相对动量因子。最后，基于因子测试的有效性，我们探索营收相似相对动量因子在不同股票样本池内的选股表现。本报告的主要结论如下：

- 秩鼎基于公司披露的分行业、分业务领域或分产品营业收入占比，结合公司年报、公告或官网对自身的经营情况描述、供应链交易信息、研报等信息，对公司营业收入进行经营领域拆解，并将拆解结果对照至国家统计局《统计用产品分类目录》的标准产品代码。
- 在一个非完全有效的市场，对于某个特定企业（即目标公司），其他具有相同经营业务或经营业务相似的企业信息可能不会被投资者及时地关注并处理，该目标公司的股票价格可能短时间的偏离其实际价值。其后，投资者逐渐充分接收、处理相关信息，股票价格逐渐反应该信息，这样就导致了该公司股票价格与其经营业务相似公司的股票价格出现了在时间维度上的联动效应。
- 基于上述逻辑，我们构建了营收相似相对动量因子（RevRelative_MOM），该因子在横截面上对股票收益的预测能力和稳定性均较强。从逻辑出发，我们从信息来源和经营业务复杂度两个维度对 RevRelative_MOM 因子进行增强。增强后的 IndComp_RevRelative_MOM 因子对股票收益的预测能力和稳定性有所提升。
- 进一步，我们创新性地引入标准产品之间的关联性，对企业之间的营收关联度进行修正，构建考虑业务关联性的营收相似相对动量因子。改进后因子效果有显著提升，IndComp_RevRelative_MOM_adj 因子 IC 均值达到 5.21%，ICIR 和 t 统计量分别达到 0.65 和 4.26。
- IndComp_RevRelative_MOM_adj 因子与反转因子的相关性较高。但是将 IndComp_RevRelative_MOM_adj 因子与反转因子进行中性化处理后，因子仍具有较强的预测能力和稳定性；通过双变量投资组合分析，我们发现控制了反转因子的影响后，IndComp_RevRelative_MOM_adj 因子仍具有较强的分组能力，且单调性较好。
- IndComp_RevRelative_MOM_adj 因子在不同的宽基指数中的有效性均较为显著，t 统计量均在 2 以上；此外，相比沪深 300 与中证 500，该因子在中证 1000 股票池中的有效性更强，其 IC 均值达到 6.15%，ICIR 为 0.77，t 统计量 5.06，IC 胜率达到了 79%。

风险提示：本报告结果通过历史数据统计、建模和测算完成，在政策、市场环境发生变化时模型存在失效的风险；本报告所提及个股或基金仅表示与相关主题有一定关联性，不构成任何投资建议。

任瞳 S1090519080004
rentong@cmschina.com.cn
周靖明 S1090519080007
zhoujingming@cmschina.com.cn
杨航 S1090523010004
yanghang4@cmschina.com.cn

正文目录

一、企业营收拆解标准化数据概述	4
1.1 为什么营收拆解数据需要标准化	4
1.2 标准化营收拆解数据	5
二、基于营收相似性的动量效应	7
2.1 基于营收相似性的绝对动量效应	7
2.2 基于营收相似性的相对动量效应	8
三、基于营收相似性的相对动量效应增强	10
3.1 信息来源	10
3.2 经营业务复杂度	11
3.3 同时考虑信息来源与经营业务复杂度	13
四、考虑业务关联性的营收相似相对动量因子	15
4.1 考虑业务关联性的营收相似相对动量因子	15
4.2 大类因子相关性分析	17
4.3 不同宽基指数中的因子表现	18
五、总结	22

图表目录

图 1: 营收数据标准化	5
图 2: 营收数据标准化 (以南玻 A (000012.SZ) 2021 年年报数据为例)	6
图 3: 营收相似相对动量因子 (RevRelative_MOM) 因子分组测试	9
图 4: 不同行业的营收相似相对动量因子 (Ind_RevRelative_MOM) 因子分组测试	11
图 5: 考虑业务复杂度的营收相似相对动量因子 (Comp_RevRelative_MOM) 因子分组测试	12
图 6: IndComp_RevRelative_MOM 因子分组测试	14
图 7: IndComp_RevRelative_MOM_adj 因子分组测试	17
图 8: IndComp_RevRelative_MOM_adj 因子与反转因子双变量投资组合分析	18
图 9: IndComp_RevRelative_MOM_adj 因子在沪深 300 中的表现	19
图 10: IndComp_RevRelative_MOM_adj 因子在中证 500 中的表现	20
图 11: IndComp_RevRelative_MOM_adj 因子在中证 1000 中的表现	20

图 12: 指数成分股 180 天平均覆盖机构数	21
表 1: 相同业务不同名称示例	4
表 2: 营收数据标准化结果 (以南玻 A (000012.SZ) 2021 年年报数据为例)	6
表 3: 因子回测框架	8
表 4: 营收相似绝对动量因子 (Rev_MOM) IC 测试结果)	8
表 5: 营收相似相对动量因子 (RevRelative_MOM) IC 测试结果)	9
表 6: 营收相似相对动量因子 (RevRelative_MOM) 因子分组测试结果展示	9
表 7: IC 测试结果比较)	10
表 8: 不同行业的营收相似相对动量因子 (Ind_RevRelative_MOM) 因子分组测试结果展示	11
表 9: IC 测试结果比较)	12
表 10: 考虑业务复杂度的营收相似相对动量因子 (Comp_RevRelative_MOM) 因子分组测试结果展示	13
表 11: IC 测试结果比较	13
表 12: IndComp_RevRelative_MOM 因子分组测试结果展示	14
表 13: IC 测试结果比较)	16
表 14: IndComp_RevRelative_MOM_adj 与 IndComp_RevRelative_MOM 因子分组测试对比	17
表 15: IndComp_RevRelative_MOM_adj 因子与常见风格因子相关性 (%)	17
表 16: IC 测试结果比较	18
表 17: 不同股票池的 IndComp_RevRelative_MOM_adj 因子 IC 测试结果	19
表 18: 不同股票池下的 IndComp_RevRelative_MOM_adj 因子分组测试	19

在一个完全有效的市场，在任意时刻，针对某项经营业务的信息，将会被充分及时的反应到所有经营该业务的上市公司的股价中。然而，行为金融领域的相关研究表明，投资者由于存在对这些公开信息的有限关注和信息处理能力等认知资源的约束，将会导致这些信息在市场传播中出现“摩擦”，影响投资者对公司未来的预期，进一步影响了资产定价。因此，在一个非完全有效的市场，对于某个特定企业（即目标公司），其他具有相同经营业务或经营业务相似的企业的信息可能不会被投资者及时地关注并处理，该目标公司的股票价格可能短时间的偏离其实际价值。其后，投资者逐渐充分接收、处理相关信息，股票价格逐渐反应该信息，这样就导致了该公司股票价格与其经营业务相似公司的股票价格出现了在时间维度上的联动效应。

作为“蓝海启航”系列报告的第八篇，本报告采用秩鼎公司提供的营收拆解标准化数据，围绕 A 股市场上市公司经营业务的相似性，构建了营收相似相对动量因子（RevRelative MOM），并从信息来源和经营业务复杂度两个维度对该因子进行增强。进一步，我们创新性地提出不同经营业务之间可能也存在相关性，并构建了考虑业务相关性的营收相似相对动量因子。最后，基于因子测试的有效性，我们探索营收相似相对动量因子在不同股票样本池内的选股表现。

一、企业营收拆解标准化数据概述

1.1 为什么营收拆解数据需要标准化

上市公司在年度报告中需要按规定披露公司所从事和生产的主要业务、主要产品等公司经营业务信息。这些信息是投资者整体把握公司经营概况和业务发展的重要依据。然而，上市公司披露的营收数据存在两个问题：一方面，不同公司对相同业务可能有不同的命名，这就导致当分析的公司数量较多时，我们无法对不同公司的业务进行批量化对比。如下表所示，这些上市公司均生产伺服微电机，但是公司披露的产品名称却大相径庭。

表 1：相同业务不同名称示例

股票代码	公司简称	公司按产品披露名称	标准产品代码	标准产品名称
002979.SZ	雷赛智能	电机	3903050204	伺服微电机
002008.SZ	大族激光	激光及自动化配套设备	3903050204	伺服微电机
002025.SZ	航天电器	电机	3903050204	伺服微电机
603416.SH	信捷电气	驱动系统	3903050204	伺服微电机
300048.SZ	合康新能	节能设备高端制造	3903050204	伺服微电机
002139.SZ	拓邦股份	电机及控制系统类	3903050204	伺服微电机

资料来源：招商证券、秩鼎

另一方面，不同公司披露的经营业务的详细程度也存在区别。有的公司披露的业务较为详细具体，有的公司则较为粗略，也同样会造成不同公司业务复杂度等指标不可比的现象。比如，Wind 显示根据产品分类中航西飞（000768.SZ）2021 年经营业务仅有“航空产品和其他”，并没有对航空产品进行进一步的分类。而广联航空（300900.SZ）2021 年经营业务则包含“航天航空零部件及无人机”、“航空工装”和“航空辅助工具及其他”，相对而言更为具体。

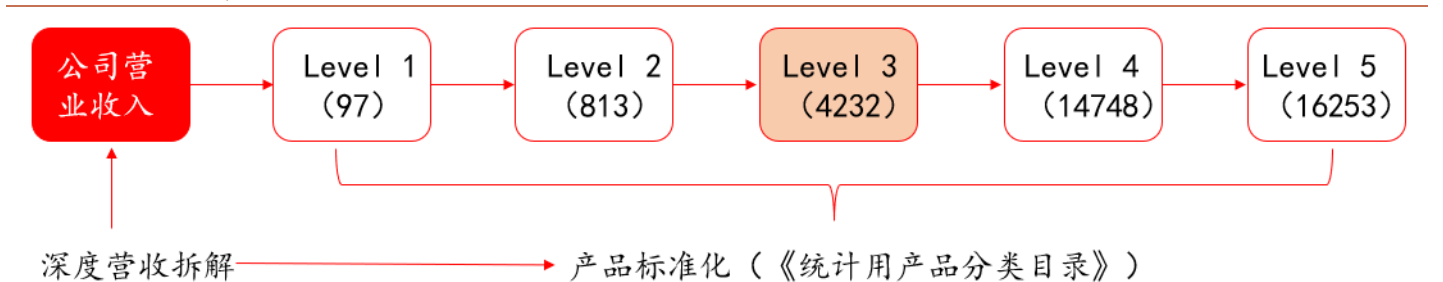
1.2 标准化营收拆解数据

在上一节中提到上市公司在年报中披露的营收数据没有统一的标准，在量化中直接使用该数据容易出现不同公司之间营收数据不可比的问题。**因此，我们需要遵循一定的规则将所有上市公司的营收数据标准化。**

秩鼎基于公司披露的分行业、分业务领域或分产品营业收入占比，结合公司年报、公告或官网对自身的经营情况描述、供应链交易信息、研报等信息，对公司营业收入进行经营领域拆解，并将拆解结果对照至《统计用产品分类目录》的标准产品代码。

《统计用产品分类目录》是由国家统计局发布，是国际上广泛采用的分类标准，涉及国民经济行业活动的全部产出，且建立了与联合国《产品总分类》的转换关系，可用于国际间的数据对比。其产品包括实物类产品和服务类产品，共计97个一级产品、813个二级产品、4232个三级产品、14748个四级产品和16253个五级产品。**目前，秩鼎已实现对2018年以来中证1800成分股每一年度的营收拆解，并将产品细分到《统计用产品分类目录》Level 3层级，对公司营收进行了标准化，如图2所示。**

图 1：营收数据标准化



资料来源：招商证券、秩鼎

下面我们以南玻A（000012.SZ）2021年年报数据为例，展示营收拆解及标准化的过程。如下图所示，2021年度南玻披露的营收主要包含玻璃产品、电子玻璃及显示器件产业、太阳能及其他产业三块。参照公司年报，我们发现南玻的营收中玻璃产品又包含浮法玻璃、光伏玻璃和工程玻璃，电子玻璃及显示器件产业包含电子玻璃与显示器件，而太阳能及其他产业包含电池片、高纯多晶硅、硅片和光伏电站四个部分。在对公司业务进行深度拆解后，我们把拆分后的产品匹配到《统计用产品分类目录》的对应层级，得到标准化后的产品列表。

图 2: 营收数据标准化 (以南玻 A (000012.SZ) 2021 年年报数据为例)



资料来源: 招商证券、秩鼎

表 2 中我们列出了南玻 2021 年营收拆解标准化后的结果。可以看到在原始披露中玻璃产品、电子玻璃及显示器产业和太阳能及其产业各占南玻 2021 年度营收的 78.81%、13.52% 和 7.68%。在营收拆解后, 由于没有信息披露南玻在细分产品的营收比例, 这里我们做一个均分的处理, 如玻璃产品分为了浮法玻璃、工程玻璃和光伏玻璃三块, 因此这三个细分产品各占玻璃产品营收占比 (78.81%) 的 1/3。

表 2: 营收数据标准化结果 (以南玻 A (000012.SZ) 2021 年年报数据为例)

原始披露	营收占比	营收拆解	标准产品名称	产品占比
玻璃产品	78.81%	浮法玻璃	平板玻璃	26.27%
玻璃产品	78.81%	工程玻璃	玻璃	26.27%
玻璃产品	78.81%	光伏玻璃	特种玻璃	26.27%
电子玻璃及显示器产业	13.52%	电子玻璃	特种玻璃	6.76%
电子玻璃及显示器产业	13.52%	显示器件	电真空光电子器件	6.76%
太阳能及其他产业	7.68%	电池片	物理电池	1.92%
太阳能及其他产业	7.68%	高纯多晶硅	电子半导体材料	1.92%
太阳能及其他产业	7.68%	硅片	电子半导体材料	1.92%
太阳能及其他产业	7.68%	光伏电站	发电量	1.92%

资料来源: 招商证券、秩鼎

二、基于营收相似性的动量效应

2.1 基于营收相似性的绝对动量效应

在第一章中我们具体介绍了我们如何对公司的营收进行详细的拆解并根据《统计用产品分类目录》进行标准化。在对公司的营收进行标准化后，不同公司之间的业务变得可比，我们可以利用该数据计算公司之间的营收相关度，从营收的角度探索上市公司之间的相似性。

具体来说，首先我们基于每个年度公司在不同标准产品的营收比例（若无某项标准产品业务，则取 0），构造 t 时刻公司的营收向量 P_t 。其次，我们计算两两公司营收向量之间的余弦相似度。具体计算公式如下：

$$r_{ij,t} = \frac{P_{i,t}P'_{j,t}}{(P_{i,t}P'_{i,t})^{\frac{1}{2}}(P_{j,t}P'_{j,t})^{\frac{1}{2}}} \quad (1)$$

其中： $P_{i,t}$ 是 i 公司 t 时刻的营收向量，是一个 $1 \times m$ 的向量，m 为全市场标准产品数量， $P_{j,t}$ 则是 j 公司 t 时刻的营收向量。 $r_{ij,t}$ 是公司 i 和 j 在 t 时刻的营收相似度。

通过上述公式我们可以计算出两两公司之间的营收相似度，**营收相似度越高，代表两个公司经营的业务和营收结构越相似。当出现针对公司经营业务的信息冲击时，两个公司股价的相关性越高。**由于上市公司年报的公告期是下一年的 1 月 1 日到 4 月 30 日，因此在每年 5 月之后我们使用的是上一年度公司之间的营收相似度，而在每年 5 月之前我们使用的是上上年度公司之间的营收相似度。

与我们在《基于供应链动量效应的选股策略》报告中构建的供应链动量因子类似，我们首先考虑基于公司之间的营收相似性构建营收相似绝对动量因子（Rev_MOM）。该因子刻画的是对于公司 i，其 t 时刻的因子值是与之具有营收相似性的其他公司过去一段时间的关联度加权收益。该因子的计算公式如下：

$$\text{Rev_MOM}_{i,t} = \frac{\sum_{j \neq i} r_{ij,t} * \text{Ret}_{j,t}}{\sum_{j \neq i} r_{ij,t}} \quad (2)$$

其中， $\text{Rev_MOM}_{i,t}$ 为公司 i 在 t 时刻的因子值， $\text{Ret}_{j,t}$ 为公司 j 在 t 时刻过去一段时间的收益率。我们认为当公司之间的营收相似度小于一定阈值时，公司之间的关联其实是较弱的，股价的关联也相对较弱，对我们构建的因子来说是一个噪音。为了剔除这部分的影响，我们仅保留营收相似度大于 0.1 的营收相似公司（即 $r_{ij,t} > 0.1$ ）。

表 3 中我们列出了因子回测的框架。回测区间为 2019 年 5 月 1 日至 2022 年 11 月 30 日，每个月最后一个交易日进行调仓，股票权重为等权方式，股票样本池为中证 1800 成分股。本文的因子测试均采用此因子回测框架。

表 3: 因子回测框架

项目	内容
回测区间	2019.5.1-2022.11.30
调仓频率	月度
调仓日	每个月最后一个交易日
样本空间	中证 1800
动量回溯期	20 个交易日
IC 测试	IC 指标为因子值与下一期股票收益率的秩相关系数
分组测试	在每个月最后一个交易日后，根据因子值大小将样本空间内的股票分成 10 组，每组组内进行等权配置计算各组历史表现。多头组为因子值最大的组，空头组为因子值最小的组
基准	样本空间内股票的等权组合
交易费率	暂不考虑交易费率

资料来源：招商证券

从下表中我们可以看到，构建的 Rev_MOM 因子在横截面上对股票收益的预测能力和稳定性较为一般。 Rev_MOM 因子在回测区间内的 IC 均值为 1.46%，t 值为 0.73。

表 4: 营收相似绝对动量因子 (Rev_MOM) IC 测试结果

因子	平均值	胜率	最大值	最小值	IC_IR	t 统计量
Rev_MOM	1.46%	52.50%	40.66%	-23.48%	0.12	0.73

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

2.2 基于营收相似性的相对动量效应

上一节中，我们验证了营收相似绝对动量因子 (Rev_MOM) 的回测效果较为一般。从因子构建的逻辑出发，我们认为具有相似经营业务的公司，受限于投资者有限关注，可能出现股票价格在时间维度上的联动效应。因此，除了从绝对动量的角度，我们也可以考虑从公司之间相对收益的角度来构建因子，即具有相似经营业务公司“补涨”的逻辑。基于这个逻辑，我们构建了营收相似相对动量因子 (RevRelative_MOM)，该因子的计算公式如下：

$$\text{RevRelative_MOM}_{i,t} = \frac{\sum_{j \neq i} r_{ij,t} * (\text{Ret}_{j,t} - \text{Ret}_{i,t})}{\sum_{j \neq i} r_{ij,t}} \quad (3)$$

其中， $\text{RevRelative_MOM}_{i,t}$ 为公司 i 在 t 时刻的因子值， $\text{Ret}_{i,t}$ 为公司 i 在 t 时刻过去一段时间的收益率。该因子刻画的是对于公司 i，其 t 时刻的因子值是与之具有营收相似性的其他公司过去一段时间相对于公司 i 的关联度加权超额收益。

从下表中我们可以看到，构建的 RevRelative_MOM 因子在横截面上对股票收益的预测能力和稳定性均较强。 RevRelative_MOM 因子在回测区间内的 IC 均值为 3.63%，t 值为 2.93，均较为显著。

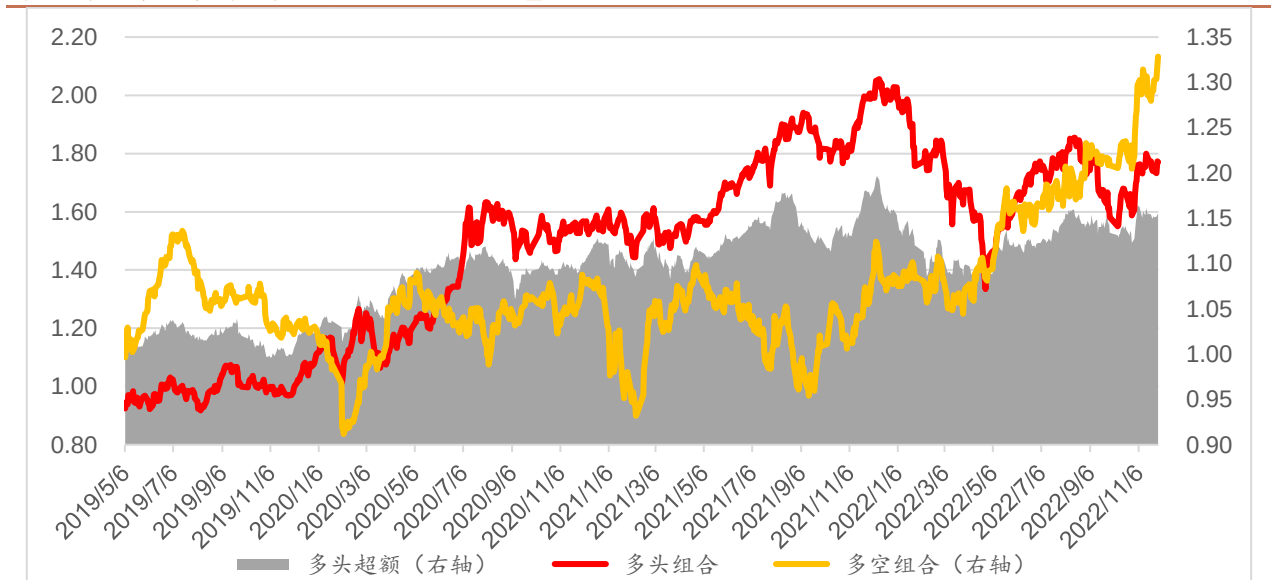
表 5: 营收相似相对动量因子 (RevRelative_MOM) IC 测试结果

因子	平均值	胜率	最大值	最小值	IC_IR	t 统计量
RevRelative_MOM	3.63%	58.14%	23.56%	-8.31%	0.45	2.93

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

下表我们列出 RevRelative_MOM 因子的分组测试结果, 可以看到多头组合表现较好, 近三年总收益 77.12%, 年化收益 17.74%, 夏普比和卡玛比分别为 0.64 和 0.51。多空组合与多头超额年化收益分别为 8.45% 和 4.20%。

图 3: 营收相似相对动量因子 (RevRelative_MOM) 因子分组测试



回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

表 6: 营收相似相对动量因子 (RevRelative_MOM) 因子分组测试结果展示

组别	总收益	年化收益	年化波动率	最大回撤	夏普比率	卡玛比率	年化双边换手率
多头组合	77.12%	17.74%	24.75%	-35.09%	0.64	0.51	330.52%
多空组合	32.84%	8.45%	13.68%	-19.74%	0.47	0.43	\
多头超额	15.50%	4.20%	6.55%	-9.98%	0.34	0.42	\

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

三、基于营收相似性的相对动量效应增强

第二章中我们验证了在 A 股市场中确实存在基于营收相似性的相对动量效应，通俗的说就是营收相似的公司可能存在“补涨”的现象，并基于此逻辑构建营收相似相对动量因子（RevRelative_MOM）。这一章我们尝试分别从信息来源和经营业务复杂度两个维度尝试对该因子进行增强。

3.1 信息来源

投资者将全市场上市公司根据经营业务的不同划分成不同的行业，如中信证券将 A 股划分成 30 个一级行业。一般来说，处于相同行业的公司其经营业务相似性较高。投资者通过行业分类的方法能够综合概括市场行业信息、提升信息处理效率。

近年来，随着公司业务不断发展以及企业分工逐渐精细化，处于相同行业的公司其经营业务可能也有较大区别，处于不同行业的公司其经营业务可能也存在相似性。然而，在投资者的潜意识里，投资者可能认为同行业公司的信息与目标公司更相关，而其他行业公司的信息与目标公司的相关性较低。这样就导致投资者对于来自不同行业公司的关注力不足，来自不同行业的信息更不容易反映到目标公司的股价中，最终导致来自不同行业但是营收相似的公司与目标公司股价在时间维度的联动效应更强。

基于这个逻辑，我们将第二章中构建的营收相似相对动量因子的计算方法进行细微的调整，即仅考虑来自不同行业但营收相似公司的相对动量，构建不同行业的营收相似相对动量因子（Ind_RevRelative_MOM）。该因子的计算公式如下：

$$\text{Ind_RevRelative_MOM}_{i,t} = \frac{\sum_{j \neq i} r_{ij,t} * (\text{Ret}_{j,t} - \text{Ret}_{i,t})}{\sum_{j \neq i} r_{ij,t}} (i \text{ 与 } j \text{ 属于不同行业}) \quad (4)$$

其中，公司 j 是与公司 i 处于不同中信一级行业分类但是营收相似（ $r_{ij,t} > 0.1$ ）的公司。

从下表中我们可以看到，构建的 Ind_RevRelative_MOM 因子在横截面上对股票收益的预测能力和稳定性均较强。同时，与 RevRelative_MOM 因子对比，Ind_RevRelative_MOM 因子有效性有显著提升，IC 均值从 3.63% 提升到 4.65%，t 统计量从 2.93 提升到 3.54。

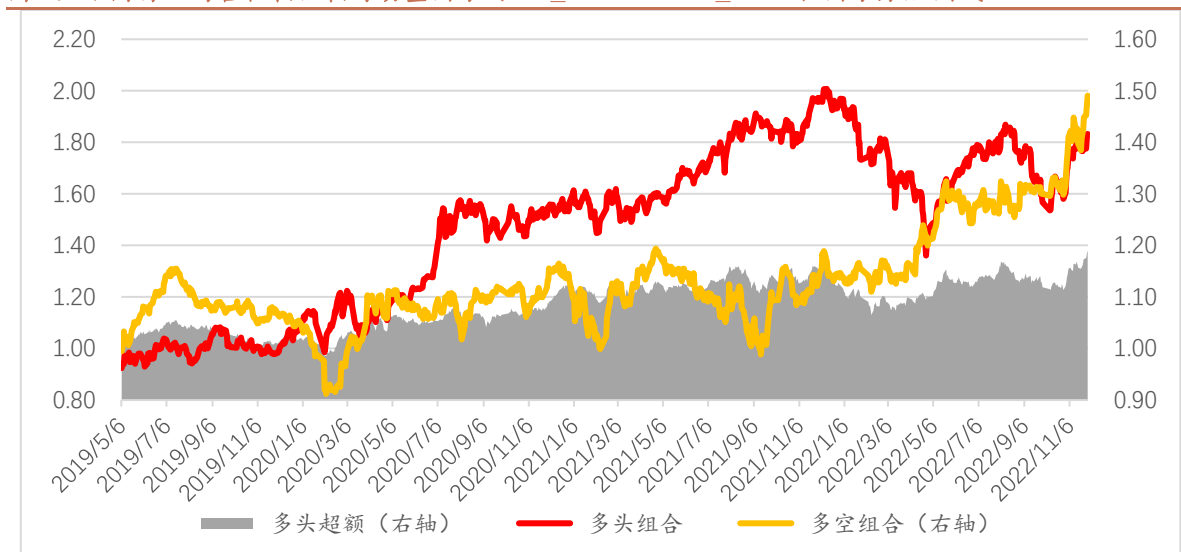
表 7: IC 测试结果比较

因子	平均值	胜率	最大值	最小值	IC_IR	t 统计量
RevRelative_MOM	3.63%	58.14%	23.56%	-8.31%	0.45	2.93
Ind_RevRelative_MOM	4.65%	62.79%	28.77%	-8.35%	0.54	3.54

回溯区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

下表我们列出 Ind_RevRelative_MOM 因子的分组测试结果，可以看到多头组合表现较好，近三年总收益 83.25%，年化收益 18.89%，夏普比和卡玛比分别为 0.69 和 0.59，收益表现优于原本的 RevRelative_MOM 因子。多空组合与多头超额年化收益同样有所提升，两者分别为 12.08% 和 5.10%。

图 4: 不同行业的营收相似相对动量因子 (Ind_RevRelative_MOM) 因子分组测试



回溯区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

表 8: 不同行业的营收相似相对动量因子 (Ind_RevRelative_MOM) 因子分组测试结果展示

组别	总收益	年化收益	年化波动率	最大回撤	夏普比率	卡玛比率	年化双边换手率
多头组合	83.25%	18.89%	24.41%	-32.22%	0.69	0.59	317.62%
多空组合	49.06%	12.08%	16.20%	-21.02%	0.62	0.57	\
多头超额	19.01%	5.10%	7.04%	-8.72%	0.44	0.58	\

回溯区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

3.2 经营业务复杂度

上一小节我们从信息来源的角度验证了投资者更容易倾向于忽视来自不同行业的信息, 从而影响其接收处理信息的效率。那么如果目标公司的经营业务复杂度较高, 当投资者接收到来自外界的信息之后, 其可能需要花费更多的精力和时间处理分析该信息对目标公司的影响, 最终也会导致目标公司与其营收相似公司的股价在时间维度上的联动效应更强。基于这个逻辑, 我们同样对第二章中构建的营收相似相对动量因子进行改进, 构建考虑业务复杂度的营收相似相对动量因子 (Comp_RevRelative_MOM)。

那么首先, 我们需要考虑如何量化上市公司的复杂度 (Rev_Comp)。第一种方法, 我们可以借用产业组织学中用来衡量行业市场集中度的赫芬达尔-赫兹曼指数 (Herfindahl-Hirschman Index, HHI)。其原理是计算行业内各市场竞争主体所占行业总收入或总资产的百分比的平方和。当 HHI 较大时, 表明该市场行业集中度较高, 存在一定的寡头格局, 反之则行业集中度较低, 竞争较大。与之类似, 我们利用 HHI 的概念来衡量公司的主营业务的复杂度, 即计算公司各项业务占其总营收比例的平方和。当公司主营业务较为分散时, 其 HHI 较小, 我们认为其主营业务复杂程度较高。而主营业务较为集中的公司, 其 HHI 较大, 我们认为其主营业务复杂程度较低。为了与第二种方法计算的业务复杂度单调性方向一致, $Rev_Comp_{i,t} = 1 - HHI_{i,t}$, $Rev_Comp_{i,t}$ 为 t 时刻公司 i 的业务复杂度

第二种方法, 我们可以用下式来衡量公司业务复杂度:

$$Rev_Comp_{i,t} = \frac{(\sum_{d=1}^d r_{id,t})/d}{(\sum_{s=1}^s r_{is,t})/s} \quad (5)$$

其中, $Rev_Comp_{i,t}$ 为 t 时刻公司 i 的业务复杂度, $r_{id,t}$ 为 t 时刻与公司 i 处于不同中信一级行业的公司与公司 i 的营收相似度, $r_{is,t}$ 则是 t 时刻与公司 i 处于相同中信一级行业的公司与公司 i 的营收相似度。该公式表示的是 t 时刻公司 i 与不同行业的公司的营收相似度的均值除以公司 i 与相同行业的公司的营收相似度的均值。通过该公式计算出来的值越大, 代表该公司与不同行业公司经营业务相似性越高, 该公司跨行业多元经营的程度越高, 我们也就认为其经营业务越复杂。

接下来，我们就将目标公司经营业务的复杂度纳入到因子计算中，构建考虑业务复杂度的营收相似相对动量因子（Comp_RevRelative_MOM）。该因子的计算公式如下：

$$\text{Comp_RevRelative_MOM}_{i,t} = \text{Rev_Comp_Rank}_{i,t} * \frac{\sum_{j \neq i} r_{ij,t} * (\text{Ret}_{j,t} - \text{Ret}_{i,t})}{\sum_{j \neq i} r_{ij,t}} \quad (6)$$

其中，乘号右边与RevRelative_MOM因子的计算公式一致，而Rev_Comp_Rank_{i,t}为t时刻公司i的业务复杂度Rev_Comp_{i,t}在其所在的中信一级行业的百分比排名。从式子中我们可以看到，若乘号右侧的值相同，当目标公司的业务复杂度越低，Rev_Comp_Rank_{i,t}的值越小，它将缩小信息从其他业务相似公司传导到目标公司的时滞，本期目标公司的因子值越小。我们暂且将以 HHI 方法计算的公司复杂度构建的考虑业务复杂度的营收相似相对动量因子命名为 Comp_RevRelative_MOM_1，以第二种方法计算的公司业务复杂度构建的因子命名为 Comp_RevRelative_MOM_2，将两个因子进行对比。

从下表中，我们可以看到通过第二种方法计算公司业务复杂度构建的因子 IC 回测结果明显较好：Comp_RevRelative_MOM_1 因子与 RevRelative_MOM 因子 IC 测试结果较为接近；而 Comp_RevRelative_MOM_2 因子相比 RevRelative_MOM 因子有较为明显的提升，IC 均值由之前的 3.63% 提升到 3.92%，t 统计量也从之前的 2.93 提升到 3.30。因此，我们选择使用第二种方法来衡量公司业务的复杂度。

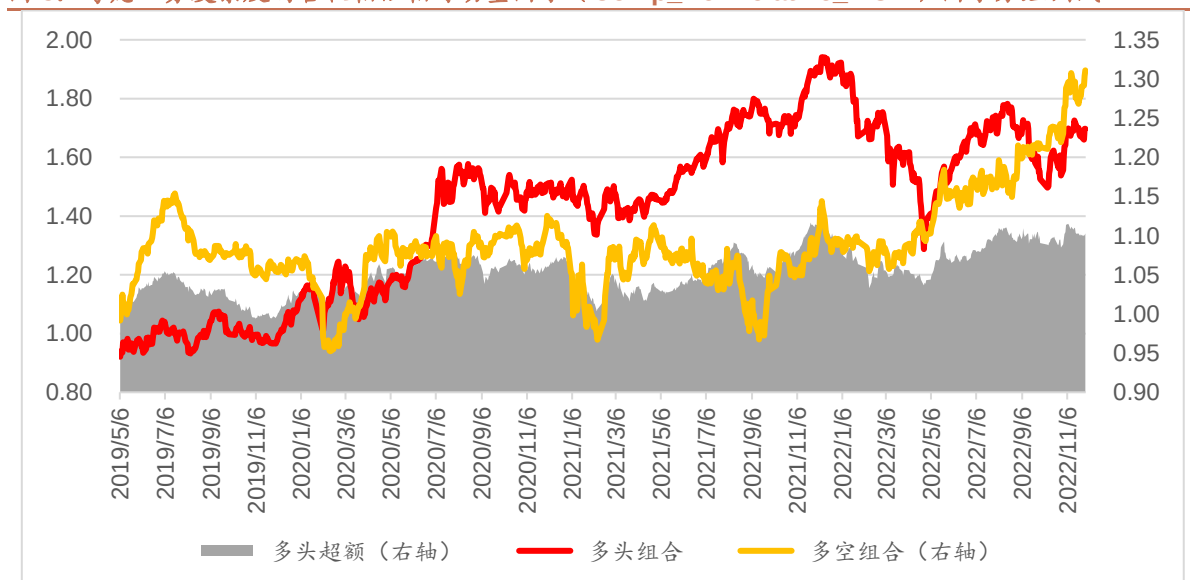
表 9: IC 测试结果比较

因子	平均值	胜率	最大值	最小值	IC_IR	t 统计量
RevRelative_MOM	3.63%	58.14%	23.56%	-8.31%	0.45	2.93
Comp_RevRelative_MOM_1	3.56%	58.14%	23.95%	-6.57%	0.46	2.99
Comp_RevRelative_MOM_2	3.92%	60.47%	23.21%	-7.01%	0.50	3.30

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

下表我们列出 Comp_RevRelative_MOM 因子的分组测试结果，可以看到多头组合表现较好，近三年总收益 69.71%，年化收益 16.31%，夏普比和卡玛比分别为 0.59 和 0.48，多头端收益表现略弱于原本的 RevRelative_MOM 因子。

图 5: 考虑业务复杂度的营收相似相对动量因子（Comp_RevRelative_MOM）因子分组测试



回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

表 10: 考虑业务复杂度的营收相似相对动量因子 (Comp_RevRelative_MOM) 因子分组测试结果展示

组别	总收益	年化收益	年化波动率	最大回撤	夏普比率	卡玛比率	年化双边换手率
多头组合	69.71%	16.31%	24.27%	-33.65%	0.59	0.48	497.26%
多空组合	31.10%	8.04%	13.53%	-17.45%	0.45	0.46	\
多头超额	10.29%	2.84%	6.24%	-8.53%	0.13	0.33	\

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

3.3 同时考虑信息来源与经营业务复杂度

在这一节中, 我们将前面两节的内容进行结合, 构建同时考虑信息来源与经营业务复杂度的因子 (IndComp_RevRelative_MOM)。该因子的计算公式为:

$$\text{IndComp_RevRelative_MOM}_{i,t} = \text{Rev_Comp_Rank}_{i,t} * \frac{\sum_{j \neq i} r_{ij,t} * (\text{Ret}_{j,t} - \text{Ret}_{i,t})}{\sum_{j \neq i} r_{ij,t}} \quad (i \text{ 与 } j \text{ 属于不同行业}) \quad (7)$$

该公式与 Comp_RevRelative_MOM 因子的计算公式相似, 唯一的区别在于上式仅考虑与目标公司 i 处于不同中信一级行业分类但是营收相似 ($r_{ij,t} > 0.1$) 的公司 j。

下表我们同时展示了四个因子的 IC 测试结果。可以看到 IndComp_RevRelative_MOM 因子对股票收益的预测能力和稳定性最好: 该因子 IC 均值为 4.73%, t 统计量为 3.76。即当我们同时考虑信息来源与目标公司的经营业务复杂度时, 其他业务相似公司过去 20 个交易日相对于目标公司的平均超额收益与目标公司股价未来涨跌的关联性更强, 符合我们构建因子的逻辑与预期。

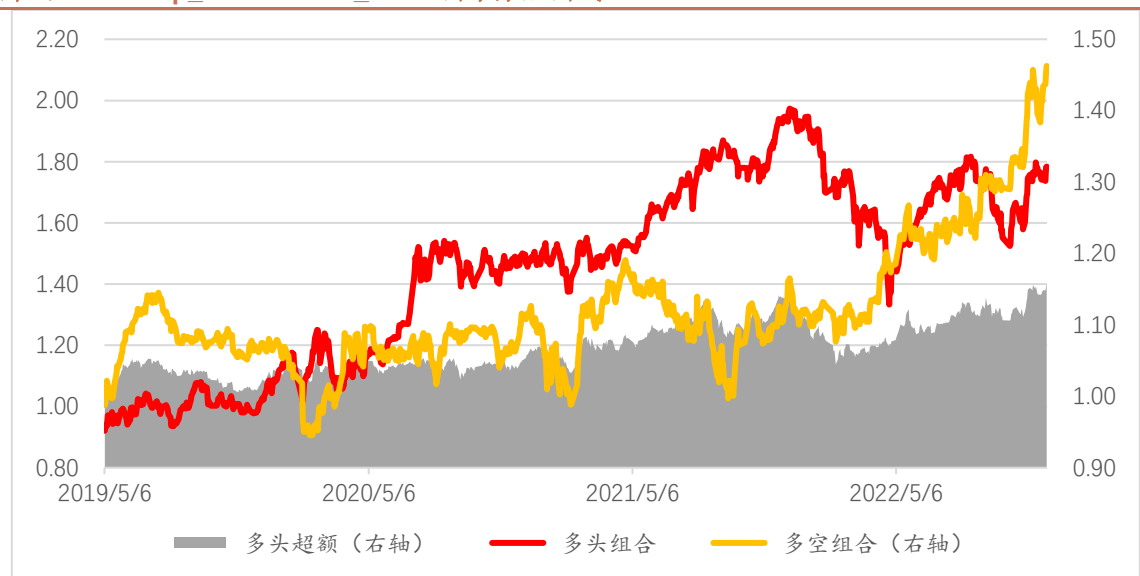
表 11: IC 测试结果比较

因子	平均值	胜率	最大值	最小值	IC_IR	t 统计量
RevRelative_MOM	3.63%	58.14%	23.56%	-8.31%	0.45	2.93
Ind_RevRelative_MOM	4.65%	62.79%	28.77%	-8.35%	0.54	3.54
Comp_RevRelative_MOM	3.92%	60.47%	23.21%	-7.01%	0.50	3.30
IndComp_RevRelative_MOM	4.73%	65.12%	25.69%	-6.91%	0.57	3.76

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

下表我们列出 IndComp_RevRelative_MOM 因子的分组测试结果, 同样可以看到多头组合表现较好, 近三年总收益 78.45%, 年化收益 18.00%, 夏普比和卡玛比分别为 0.66 和 0.55。

图 6: IndComp_RevRelative_MOM 因子分组测试



回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

表 12: IndComp_RevRelative_MOM 因子分组测试结果展示

组别	总收益	年化收益	年化波动率	最大回撤	夏普比率	卡玛比率	年化双边换手率
多头组合	78.45%	18.00%	24.11%	-32.43%	0.66	0.55	415.02%
多空组合	46.27%	11.48%	15.27%	-17.46%	0.62	0.66	/
多头超额	15.74%	4.27%	6.53%	-8.86%	0.35	0.48	/

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

四、考虑业务关联性的营收相似相对动量因子

4.1 考虑业务关联性的营收相似相对动量因子

在第二章中我们介绍了我们基于每个年度公司在不同标准产品的营收比例（若无某项标准产品业务，则取 0），构造 t 时刻公司的营收向量 P_t ，再计算两两公司营收向量之间的余弦相似度，以此来代表两个公司经营业务的相似度。但是，应用该方法计算的公司间经营业务相似度隐含的假设是不同的标准产品之间是相互独立的。

但是，我们知道在实际的生产生活中，许多产品或者服务之间往往是存在一定关联的，产品之间可能存在供应关系，而上市公司之间也因为这些经济关系而联系起来。正如我们在《“蓝海启航”系列研究之五：基于供应链动量溢出效应的选股策略》中所探讨的，上市公司之间由于存在产品的供应关系，其基本面和股价具有相关性。以光伏产业链为例，硅片对应到国家统计局《统计用产品分类目录》中的三级标准产品名称为电子半导体材料，而产业链下游电池片对应电池零部件，光伏系统设备则对应物理电池。假设市场中出现针对硅片的消息时，也将同样会影响到产业链下游生产电池片、经营光伏系统设备的公司。

基于上述逻辑，我们在这一章中，将不再认为不同的标准产品之间是相互独立的，而是创新性地引入标准产品之间的关联性，构建考虑业务关联性的营收相似相对动量因子。首先，我们需要先刻画业务关联度，与我们计算公司经营业务相似度类似，我们采用该式计算不同标准产品之间的关联度：

$$Cor_P_{mn,t} = \frac{C_{m,t}C'_{n,t}}{(C_{m,t}C'_{m,t})^{1/2}*(C_{n,t}C'_{n,t})^{1/2}} \quad (8)$$

其中： $C_{m,t}$ 是标准产品 m 在 t 时刻的公司分布向量，是一个维数 1×1800 的向量。 $C_{n,t}$ 则是另一标准产品 n 在 t 时刻的公司分布向量。 $Cor_P_{mn,t}$ 就是计算出来的是标准产品 m 和标准产品 n 在 t 时刻的业务关联度。

计算得到标准产品之间的关联度之后，我们就对第二章中初始的营收相似度 $r_{ij,t}$ 进行修正，计算考虑业务关联性的营收相似度 $r_adj_{ij,t}$ ：

$$r_adj_{ij,t} = \frac{P_{i,t}Cor_P_tP'_{j,t}}{(P_{i,t}P'_{i,t})^{1/2}*(P_{j,t}P'_{j,t})^{1/2}} \quad (9)$$

其中，式（1）与式（9）的区别在于在分子端多乘了 Cor_P_t ， Cor_P_t 是一个维数为 $m \times m$ 的矩阵， m 为全市场标准产品数量，其代表标准产品之间的关联度矩阵。 $r_adj_{ij,t}$ 为修正的公司 i 与公司 j 的营收相似度，该营收相似度考虑了公司 i 与公司 j 经营业务之间的业务关联性。值得一提的是，当 Cor_P_t 是一个单位矩阵时，即不同标准产品之间相互独立，则 $r_adj_{ij,t}$ 等于 $r_{ij,t}$ 。

下面，我们就举一个简单的例子来说明 $r_{ij,t}$ 与 $r_adj_{ij,t}$ 之间的区别。假设有两家公司 i 和 j ，公司 i 只经营 a 业务，公司 j 只经营 b 业务， a 业务和 b 业务相关性为 0.5，此时 $P_{i,t} = [1 \ 0]$ ， $P_{j,t} = [0 \ 1]$ ， $Cor_P_t = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ ，通过公式（1）我们可以计算得到两家公司的营收关联度为 0，因为两家公司分别经营两项业务；而通过公式（9）我们可以计算得到两家公司的营收关联度为 0.5，两家公司虽然经营两项不同的业务，但是两项业务存在相关性，因此我们认为两个公司的营收也存在相似度。

在对公司之间的营收相似度进行修正后，我们接下来便对第二章和第三章构建的因子进行改进，下面我们给出改进后因子的计算公式：

$$RevRelative_MOM_adj_{i,t} = \frac{\sum_{j \neq i} r_adj_{ij,t} * (Ret_{j,t} - Ret_{i,t})}{\sum_{j \neq i} r_adj_{ij,t}} \quad (10)$$

$$Ind_RevRelative_MOM_adj_{i,t} = \frac{\sum_{j \neq i} r_adj_{ij,t} * (Ret_{j,t} - Ret_{i,t})}{\sum_{j \neq i} r_adj_{ij,t}} \quad (i \text{ 与 } j \text{ 属于不同行业}) \quad (11)$$

$$Comp_RevRelative_MOM_adj_{i,t} = Rev_Comp_Rank_{i,t} * \frac{\sum_{j \neq i} r_adj_{ij,t} * (Ret_{j,t} - Ret_{i,t})}{\sum_{j \neq i} r_adj_{ij,t}} \quad (12)$$

$$\text{IndComp_RevRelative_MOM_adj}_{i,t} = \text{Rev_Comp_Rank}_{i,t} * \frac{\sum_{j \neq i} r_{ij,t} * (\text{Ret}_{j,t} - \text{Ret}_{i,t})}{\sum_{j \neq i} r_{-adj_{ij,t}}} \quad (i \text{ 与 } j \text{ 属于不同行业}) \quad (13)$$

其中，RevRelative_MOM_adj_{i,t} 为改进的营收相似相对动量因子（RevRelative_MOM），Ind_RevRelative_MOM_adj_{i,t} 为改进的不同行业的营收相似相对动量因子（Ind_RevRelative_MOM），Comp_RevRelative_MOM_adj_{i,t} 为改进的考虑业务复杂度的营收相似相对动量因子（Comp_RevRelative_MOM），IndComp_RevRelative_MOM_adj_{i,t} 为改进的考虑信息来源与经营业务复杂度的因子（IndComp_RevRelative_MOM）。

下表我们给出了考虑业务关联度后改进的因子的 IC 测试结果，并与原因子的测试结果进行对比。我们可以总结出两点：（1）从因子 IC 测试结果来看，IndComp RevRelative MOM adj 因子的整体效果最好，该因子既考虑了信息来源和经营业务复杂度，也考虑了业务之间的关联性。该因子 IC 均值达到 5.21%，ICIR 和 t 统计量分别达到 0.65 和 4.26，横截面上对股票未来收益的预测能力和稳定性较强。（2）对比原因子与考虑业务关联性后改进的因子，可以看到改进后因子效果有显著提升。从 IC 均值来看，仅 Comp_RevRelative_MOM_adj 因子略低于原因子，其他因子均高于原因子；而从 ICIR 和 t 统计量来看，改进后的因子均高于原因子。

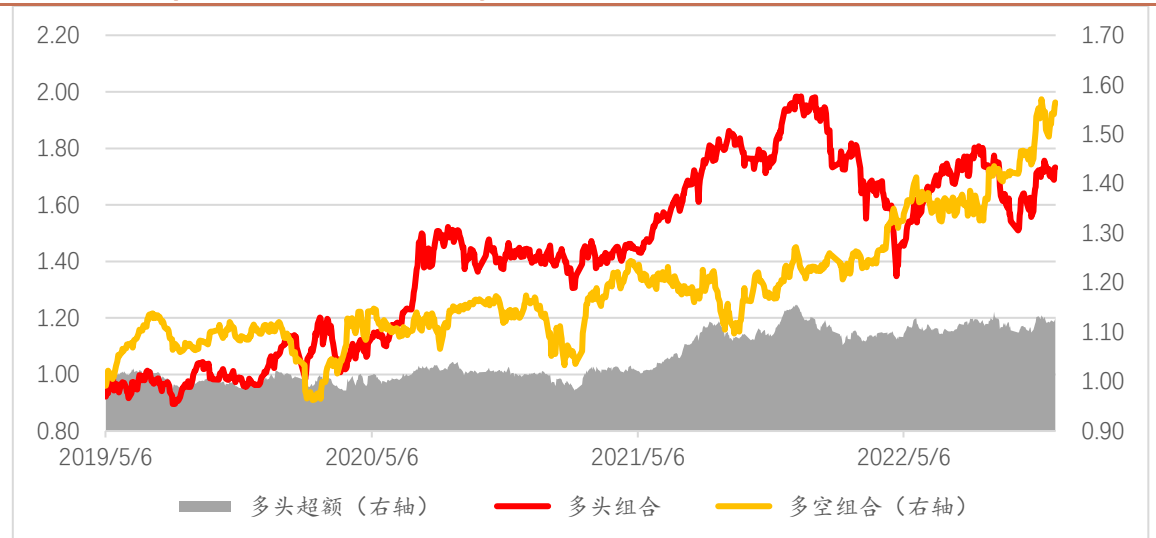
表 13: IC 测试结果比较

因子	平均值	胜率	最大值	最小值	IC_IR	t 统计量
RevRelative_MOM	3.63%	58.14%	23.56%	-8.31%	0.45	2.93
RevRelative_MOM_adj	3.84%	55.81%	21.74%	-6.54%	0.50	3.30
Ind_RevRelative_MOM	4.65%	62.79%	28.77%	-8.35%	0.54	3.54
Ind_RevRelative_MOM_adj	5.14%	65.12%	29.12%	-5.54%	0.62	4.05
Comp_RevRelative_MOM	3.92%	60.47%	23.21%	-7.01%	0.50	3.30
Comp_RevRelative_MOM_adj	3.42%	65.12%	18.02%	-4.66%	0.54	3.57
IndComp_RevRelative_MOM	4.73%	65.12%	25.69%	-6.91%	0.57	3.76
IndComp_RevRelative_MOM_adj	5.21%	62.79%	26.19%	-4.82%	0.65	4.26

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

受篇幅限制，这里我们仅展示 IC 测试中表现最好的 IndComp_RevRelative_MOM_adj 因子的分组回测效果，如果投资者对其他因子的分组效果感兴趣，欢迎与我们交流。下表我们列出 IndComp_RevRelative_MOM_adj 因子的分组测试结果，可以看到多头组合表现较好，近三年总收益 73.32%，年化收益 17.02%，夏普比和卡玛比分别为 0.61 和 0.53，多头组合整体表现与 IndComp RevRelative MOM 因子相差不大。对比多空组合的表现，可以看到 IndComp RevRelative MOM adj 因子的多空组合表现明显优于 IndComp RevRelative MOM 因子，两个因子多空组合年化收益分别为 13.63% 与 11.48%，卡玛比分别为 0.88 与 0.66。

图 7: IndComp_RevRelative_MOM_adj 因子分组测试



回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

表 14: IndComp_RevRelative_MOM_adj 与 IndComp_RevRelative_MOM 因子分组测试对比

因子	组别	总收益	年化收益	年化波动率	最大回撤	夏普比率	卡玛比率	年化双边换手率
IndComp_RevRelative_MOM	多头组合	78.45%	18.00%	24.11%	-32.43%	0.66	0.55	415.02%
	多头超额	15.74%	4.27%	6.53%	-8.86%	0.35	0.48	/
	多空组合	46.27%	11.48%	15.27%	-17.46%	0.62	0.66	/
IndComp_RevRelative_MOM_adj	多头组合	73.32%	17.02%	24.44%	-32.09%	0.61	0.53	406.42%
	多头超额	12.76%	3.49%	6.39%	-7.05%	0.23	0.50	/
	多空组合	56.40%	13.63%	14.71%	-15.42%	0.79	0.88	/

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

4.2 大类因子相关性分析

为检验上文构建的 IndComp_RevRelative_MOM_adj 因子与常见风格因子之间的相关性，我们计算了各因子之间的截面 Spearman 秩相关系数均值，具体结果如下表所示。

从下表中我们可以发现以下两点：（1）IndComp_RevRelative_MOM_adj 因子与成长因子、价量因子和价值因子的相关性较低。其中，该因子与价量因子相关性最高，相关系数均值为 11.59%。（2）IndComp_RevRelative_MOM_adj 因子与反转因子的相关性较高，呈现出较为显著的正相关关系。

表 15: IndComp_RevRelative_MOM_adj 因子与常见风格因子相关性 (%)

	IndComp_RevRelative_MOM_adj	反转因子	成长因子	价量因子	价值因子
IndComp_RevRelative_MOM_adj	100.00%				
反转因子	71.22%	100.00%			
成长因子	-6.52%	-8.54%	100.00%		
价量因子	11.59%	14.85%	1.23%	100.00%	
价值因子	6.14%	8.41%	-10.50%	3.88%	100.00%

资料来源: 招商证券、Wind

注: 反转因子计算公式为股票 20 日收益*(-1)

追本溯源，其实 IndComp_RevRelative_MOM_adj 因子与动量因子较高的相关性是因为在因子构造过程中，公式中包含了营收相似公司相对于目标公司的超额收益所导致的。换句话说，该因子的有效性既包含了目标公司相对于营收相似公司“补涨”的逻辑，又包含了目标公司短期反转的逻辑。那么，该因子的有效性中有多少“补涨”逻辑所带来的，有多少是短期反转效应所带来的呢？

这里我们尝试将 IndComp_RevRelative_MOM_adj 因子与反转因子进行截面回归取残差的中性化处理。下表我们列出了中性化处理后的 IndComp_RevRelative_MOM_adj 因子与反转因子的 IC 测试。通过下表我们可以看到在对 IndComp_RevRelative_MOM_adj 因子做中性化处理之后，其 IC 均值为 2.48%，t 统计量为 2.10，胜率 72%，因子的有效性仍然较高。同时，我们看到反转因子虽然有较高的 IC 均值，但是稳定性较差，IC 胜率 51%，t 统计量为 1.65。可以说，IndComp_RevRelative_MOM_adj 因子有效性是目标公司相对于营收相似公司“补涨”效应与目标公司短期反转效应的结合，具有 1+1>2 的效果。

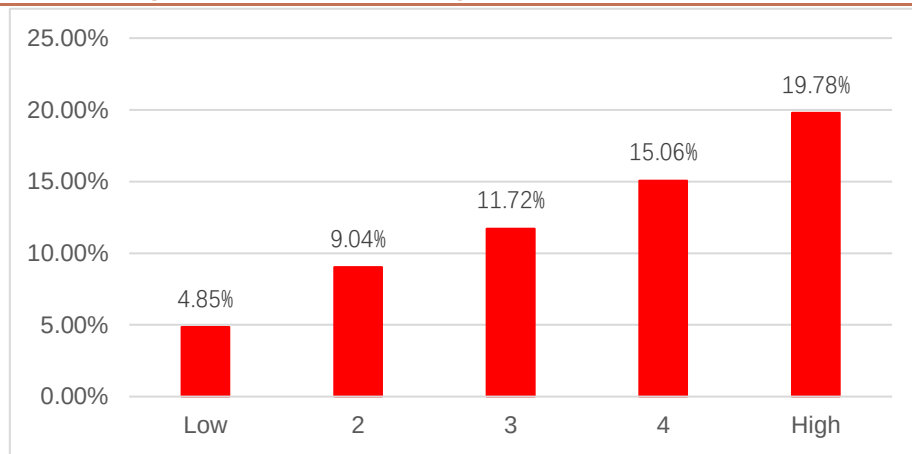
表 16: IC 测试结果比较

因子	平均值	胜率	最大值	最小值	IC_IR	t 统计量
中性化处理后的 IndComp_RevRelative_MOM_adj	2.48%	72.09%	22.49%	-15.83%	0.32	2.10
反转因子	3.56%	51.16%	39.01%	-30.33%	0.25	1.65
IndComp_RevRelative_MOM_adj	5.21%	62.79%	26.19%	-4.82%	0.65	4.26

回溯区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

此外，我们还对 IndComp_RevRelative_MOM_adj 因子与反转因子进行双变量投资组合分析，来控制反转因子对投资组合的影响。具体操作为在每个横截面首先根据反转因子的值将市场股票分成 5 等份，在每个分组里再根据 IndComp_RevRelative_MOM_adj 因子值分成 5 等份，再平均 5 个反转等分组中相同 IndComp_RevRelative_MOM_adj 因子排名的等权投资组合收益。如下图所示，在控制了反转因子的影响后，IndComp_RevRelative_MOM_adj 因子仍具有较强的分组能力，且单调性较好。

图 8: IndComp_RevRelative_MOM_adj 因子与反转因子双变量投资组合分析



回溯区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

注: High 为因子值最大的组, Low 为因子值最小的组

4.3 不同宽基指数中的因子表现

接下来，我们把 IndComp_RevRelative_MOM_adj 因子应用到不同的宽基指数中，包括沪深 300、中证 500 和中证 1000，测试该因子在这些不同股票池中的有效性。

从 IndComp_RevRelative_MOM_adj 因子在不同股票池中的 IC 测试结果表可以看到，该因子在不同的股票池中的有效性均较为显著，t 统计量均在 2 以上；此外，相比沪深 300 与中证 500，该因子在中证 1000 股票池中的有效性更强，其 IC 均值达到 6.15%，ICIR 为 0.77，t 统计量 5.06，IC 胜率达到了 79%。

表 17: 不同股票池的 IndComp_RevRelative_MOM_adj 因子 IC 测试结果

股票池	平均值	胜率	最大值	最小值	IC_IR	t 统计量
沪深 300	5.63%	65.12%	27.74%	-23.00%	0.45	2.93
中证 500	3.51%	62.79%	19.26%	-18.63%	0.37	2.42
中证 1000	6.15%	79.07%	27.57%	-4.94%	0.77	5.06

回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

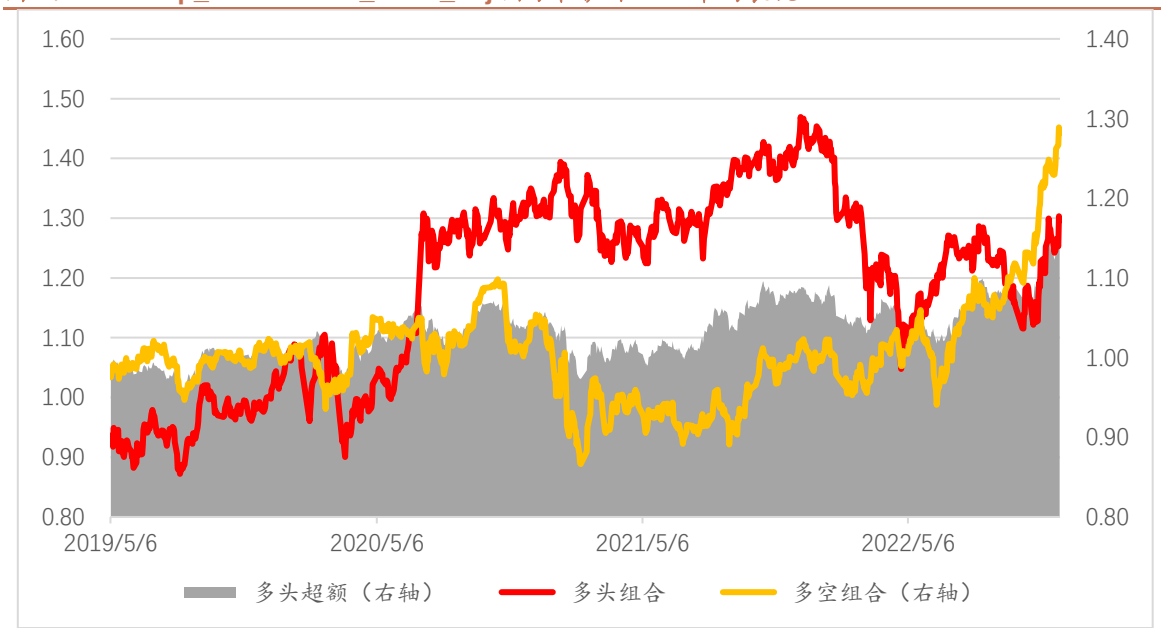
从下表我们可以看到, IndComp_RevRelative_MOM_adj 因子在沪深 300、中证 500 和中证 1000 的多头组合年化收益分别为 7.68%、13.15%和 19.51%, 多头超额年化收益分别为 3.97%、5.71%和 11.10%。同时, 该因子在中证 1000 股票池中的表现尤为亮眼: 多头超额夏普比率 1.6, 卡玛比率 1.66; 多空组合夏普比率 1.22, 卡玛比率 1.45。

表 18: 不同股票池下的 IndComp_RevRelative_MOM_adj 因子分组测试

股票池	组别	总收益	年化收益	年化波动率	最大回撤	夏普比率	卡玛比率	年化双边换手率
沪深 300	多头组合	29.55%	7.68%	22.04%	-28.68%	0.26	0.27	630.87%
	多头超额	14.59%	3.97%	7.57%	-9.20%	0.26	0.43	/
	多空组合	28.12%	7.34%	14.16%	-21.10%	0.38	0.35	/
中证 500	多头组合	54.12%	13.15%	22.43%	-28.48%	0.50	0.46	601.34%
	多头超额	21.43%	5.71%	5.92%	-5.33%	0.63	1.07	/
	多空组合	18.71%	5.02%	12.87%	-13.94%	0.23	0.36	/
中证 1000	多头组合	86.62%	19.51%	25.22%	-32.87%	0.69	0.59	599.27%
	多头超额	44.56%	11.10%	5.69%	-6.67%	1.60	1.66	/
	多空组合	72.35%	16.83%	12.16%	-11.58%	1.22	1.45	/

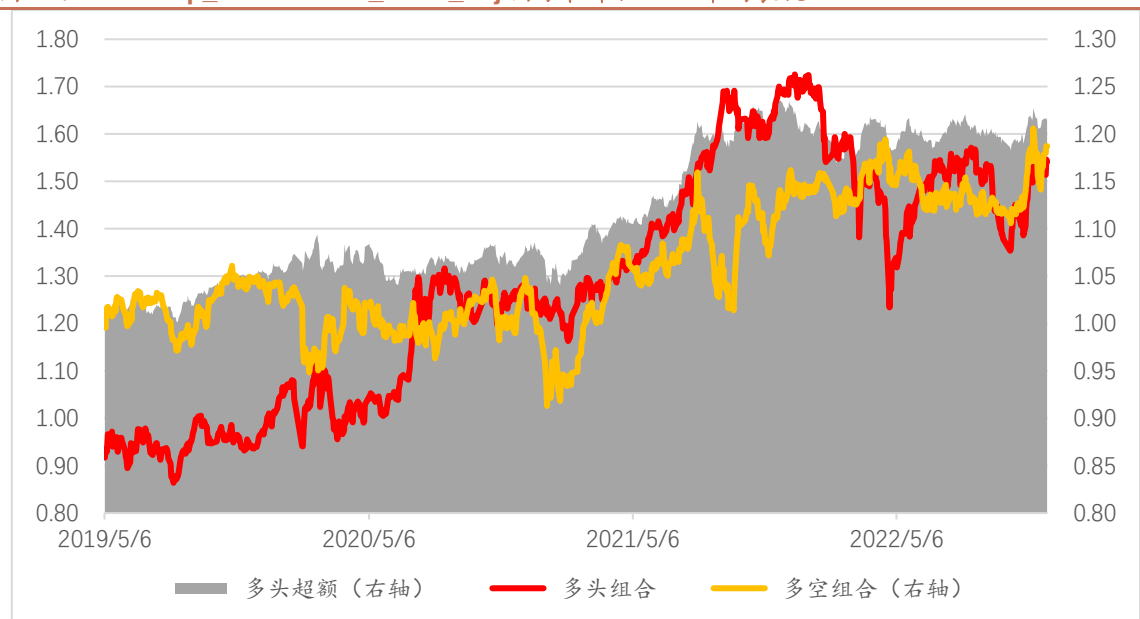
回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

图 9: IndComp_RevRelative_MOM_adj 因子在沪深 300 中的表现



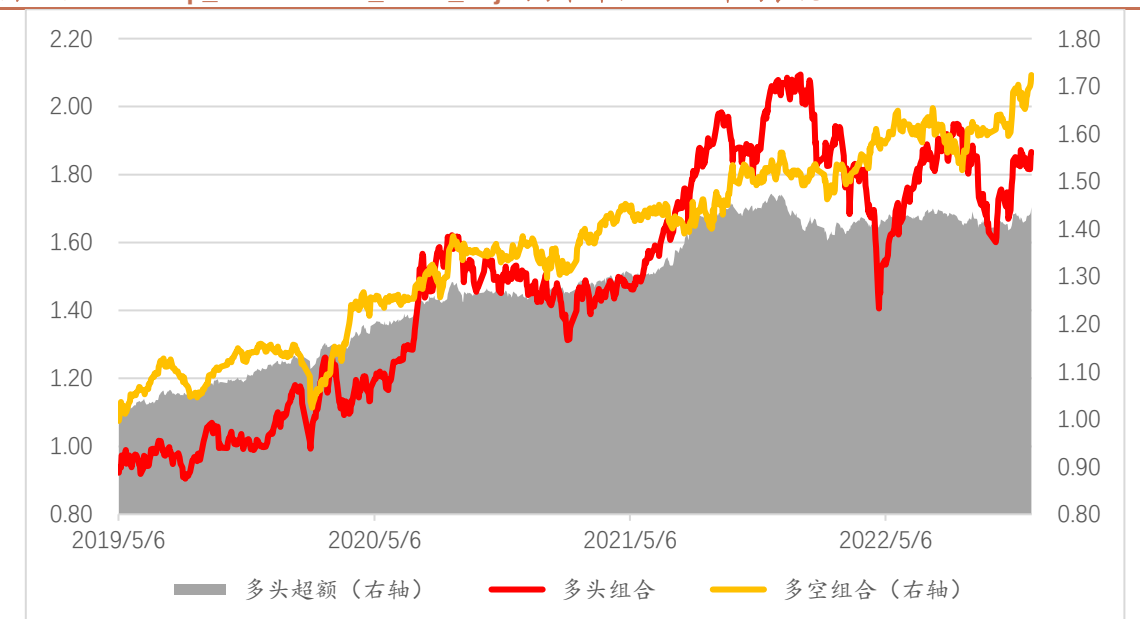
回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

图 10: IndComp_RevRelative_MOM_adj 因子在中证 500 中的表现



回溯区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

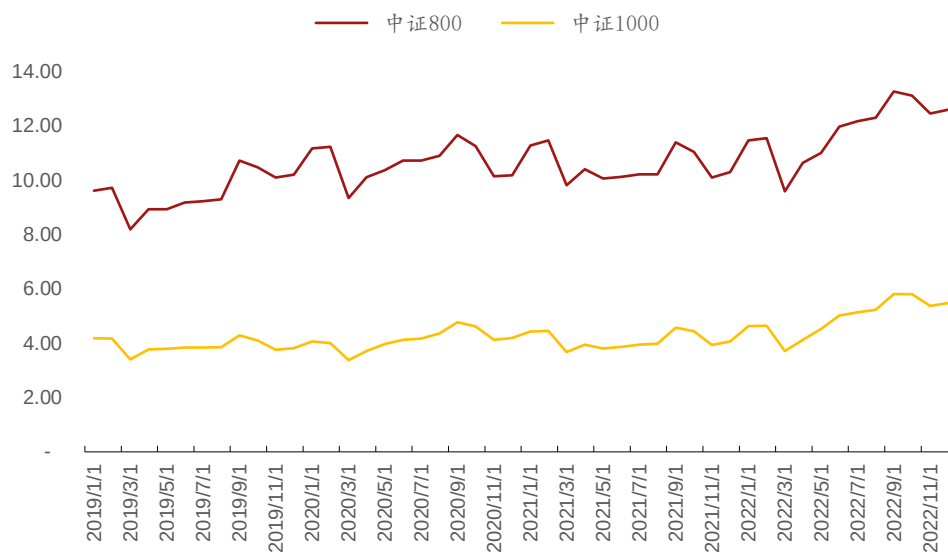
图 11: IndComp_RevRelative_MOM_adj 因子在中证 1000 中的表现



回溯区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

从上文 IndComp_RevRelative_MOM_adj 因子在沪深 300、中证 500 和中证 1000 中的回测结果可以看到，**IndComp_RevRelative_MOM_adj 因子在中证 1000 中的有效性更强。我们认为这可能与跟踪中证 1000 指数成分股的卖方机构数量较少有关。**如下图所示，我们计算了中证 800 与中证 1000 指数成分股 180 天内的平均覆盖机构数。可以看到中证 800 指数成分股平均覆盖的机构数量中枢在 10 家左右，而中证 1000 指数成分股平均覆盖的机构数量中枢在 4 家左右。卖方分析师对上市公司信息的分析处理能力较强，同时其作为沟通机构投资者与上市公司之间的桥梁，能够通过研究报告、路演等途径与投资者进行信息的交流。**因此当覆盖公司的卖方机构数量较多时，卖方分析师加速了信息在市场中的传播效率，提高了股票的市场定价效率，从而导致我们发现的营收相似的公司之间股价在时间维度上的联动效应减弱。**

图 12: 指数成分股 180 天平均覆盖机构数



回测区间: 2019.5-2022.11; 资料来源: 招商证券、Wind

五、总结

作为“蓝海启航”系列报告的第八篇，本报告采用秩鼎公司提供的营收拆解标准化数据，计算 A 股市场上市公司经营业务的相似性，构建了营收相似相对动量因子（RevRelative MOM），并从信息来源和经营业务复杂度两个维度对该因子进行增强。进一步，我们创新性地提出不同经营业务之间可能也存在相关性，并构建了考虑业务相关性的营收相似相对动量因子。最后，基于因子测试的有效性，我们探索营收相似相对动量因子在不同股票样本池内的选股表现。本报告的主要结论如下：

- 秩鼎基于公司披露的分行业、分业务领域或分产品营业收入占比，结合公司年报、公告或官网对自身的经营情况描述、供应链交易信息、研报等信息，对公司营业收入进行经营领域拆解，并将拆解结果对照至国家统计局《统计用产品分类目录》的标准产品代码。
- 在一个非完全有效的市场，对于某个特定企业（即目标公司），其他具有相同经营业务或经营业务相似的企业信息可能不会被投资者及时地关注并处理，该目标公司的股票价格可能短时间的偏离其实际价值。其后，投资者逐渐充分接收、处理相关信息，股票价格逐渐反应该信息，这样就导致了该公司股票价格与其经营业务相似公司的股票价格出现了在时间维度上的联动效应。
- 基于上述逻辑，我们构建了营收相似相对动量因子（RevRelative_MOM）。RevRelative MOM 因子在横截面上对股票收益的预测能力和稳定性均较强，该因子在回测区间内的 IC 均值为 3.63%，t 值为 2.93。从逻辑出发，我们从信息来源和经营业务复杂度两个维度对 RevRelative MOM 因子进行增强。增强后的 IndComp_RevRelative_MOM 因子对股票收益的预测能力和稳定性有所提升：该因子 IC 均值为 4.73%，t 统计量为 3.76。
- 进一步，我们创新性地引入标准产品之间的关联性，对企业之间的营收关联度进行修正，构建考虑业务关联性的营收相似相对动量因子。改进后因子效果有显著提升，IndComp_RevRelative_MOM_adj 因子 IC 均值达到 5.21%，ICIR 和 t 统计量分别达到 0.65 和 4.26。
- IndComp_RevRelative_MOM_adj 因子与反转因子的相关性较高。但是将 IndComp_RevRelative_MOM_adj 因子与反转因子进行中性化处理后，因子仍具有较强的预测能力和稳定性；此外，通过双变量投资组合分析，我们发现控制了反转因子的影响后，IndComp_RevRelative_MOM_adj 因子仍具有较强的分组能力，且单调性较好。
- IndComp_RevRelative_MOM_adj 因子在不同的宽基指数中的有效性均较为显著，t 统计量均在 2 以上；此外，相比沪深 300 与中证 500，该因子在中证 1000 股票池中的有效性更强，其 IC 均值达到 6.15%，ICIR 为 0.77，t 统计量 5.06，IC 胜率达到了 79%。

风险提示

本报告结果通过历史数据统计、建模和测算完成，在政策、市场环境发生变化时模型存在失效的风险；本报告所提及个股或基金仅表示与相关主题有一定关联性，不构成任何投资建议。

分析师承诺

负责本研究报告的每一位证券分析师，在此申明，本报告清晰、准确地反映了分析师本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告中的具体推荐或观点直接或间接相关。

任瞳： 研究发展中心执行董事，量化与基金评价团队负责人，管理学硕士，18 年证券研究经验，2010 年、2015 年、2016 年、2017 年、2018 年、2020 年新财富最佳分析师（金融工程方向）。在量化选股择时、基金研究以及衍生品投资方面有深入独到的见解。

周靖明： 武汉大学金融工程硕士，6 年量化策略研究开发经验。研究方向为多因子量化选股、另类 Alpha 研究、行业轮动、因子择时。此外，对 SmartBeta 基金和量化基金产品也有深入研究。

杨航： 美国马里兰大学公园分校数量金融硕士，厦门大学金融工程学士，2021 年加入招商证券研究发展中心。

投资评级定义

报告中所涉及的投资评级采用相对评级体系，基于报告发布日后 6-12 个月内公司股价（或行业指数）相对同期当地市场基准指数的市场表现预期。其中，A 股市场以沪深 300 指数为基准；香港市场以恒生指数为基准；美国市场以标普 500 指数为基准。具体标准如下：

股票评级

强烈推荐： 预期公司股价涨幅超越基准指数 20%以上

增持： 预期公司股价涨幅超越基准指数 5-20%之间

中性： 预期公司股价变动幅度相对基准指数介于±5%之间

减持： 预期公司股价表现弱于基准指数 5%以上

行业评级

推荐： 行业基本面向好，预期行业指数超越基准指数

中性： 行业基本面稳定，预期行业指数跟随基准指数

回避： 行业基本面转弱，预期行业指数弱于基准指数

重要声明

本报告由招商证券股份有限公司（以下简称“本公司”）编制。本公司具有中国证监会许可的证券投资咨询业务资格。本报告基于合法取得的信息，但本公司对这些信息的准确性和完整性不作任何保证。本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价，在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。除法律或规则规定必须承担的责任外，本公司及其雇员不对使用本报告及其内容所引发的任何直接或间接损失负任何责任。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。客户应当考虑到本公司可能存在可能影响本报告客观性的利益冲突。

本报告版权归本公司所有。本公司保留所有权利。未经本公司事先书面许可，任何机构和个人均不得以任何形式翻版、复制、引用或转载，否则，本公司将保留随时追究其法律责任的权利。