

ETF 的筛选分类与轮动模型

——金融工程专题

分析师：宋旻

SAC NO: S1150517100002

2024 年 12 月 30 日

证券分析师

宋旻

022-28451131

songyang@bhzq.com

核心观点：

- 在以往有关行业轮动策略的系列研究报告里，我们详细研究了如何运用各种数据来建立行业轮动模型的策略，实现了模型收益与稳定性的不断优化。然而，在实际投资中，直接投资于行业指数往往不可行，因此我们需要借助资本市场的工具来实现这一目标。ETF 基金作为一种高效、灵活的投资工具，成为了理想的选择。
- 本篇报告中，我们使用 K-means 聚类算法和分层聚类算法，对目前市场上的 492 只行业主题类 ETF 基金进行了筛选和分类，旨在识别出最具代表性的 ETF 基金。最终，我们选出了 12 个大类 40 只 ETF 基金产品，构成了模型的备选基金池。这些 ETF 基金覆盖了 31 个申万一级行业中的 25 个。市值排名前 20 的申万一级行业中，每一个行业都能够找到至少一只较为合适的 ETF 基金追踪。而在医药、计算机和有色金属等重点行业中，不仅有追踪整个行业的 ETF 基金产品可供选择，还可以找到追踪二级子行业的各种 ETF 基金。ETF 基金的布局日益全面且精细化，为投资者提供了更加高效和精准的资产配置工具。而商贸零售、轻工制造、环保、纺织服饰、美容护理以及综合这 6 个行业因为相对较为细分或市场规模较小，目前仍缺乏流动性较好的专门的 ETF 产品覆盖。
- 我们提取量价、基本面和资金面三大类共计 14 个因子 2020 年至今的周度数据，进行单因子检测。最终，我们选出了单因子检测中表现较好，彼此相关性较低，并且本身具有代表性的 5 个因子组成 ETF 轮动模型。这 5 个因子为：指数 250 日涨跌幅因子、BIAS 乖离率因子、EPS 增速环比因子、主力净流入因子和 10 日大单净买入因子。
- 我们将最终选出的 5 个因子等权合成，构建 ETF 轮动模型。模型对于收益的提升十分显著，2020 年至今年化收益 19.70%，年化超额收益 17.30%，信息比率 1.526，周度胜率 59.18%，显著跑赢所有单因子模型。
- 我们还针对 ETF 轮动模型每期选定基金数量 N 做了遍历性敏感度测试。测试表明，当每期选定基金数量为 3 至 5 只时，模型在收益和波动的平衡型上表现更好。不过，无论选定的基金数量如何变化，模型相对于基准始终保持了明显的超额收益，证明我们的 ETF 轮动策略具有较好的稳定性和适应性。
- 未来，我们将继续致力于模型的精进优化工作，持续挖掘和整合多元化因子，构建更为精准稳定的 ETF 轮动模型。

风险提示：市场风格转换风险，第三方数据提供不准确风险，历史结果不代表未来收益，本报告不构成投资建议。

目 录

1. 前言	4
2. ETF 的聚类与筛选	4
2.1 K-means 聚类	4
2.2 分层聚类	7
2.3 申万一级行业分类	8
3. ETF 轮动模型因子的选择	10
4. ETF 轮动模型的建立	12
5. 总结与未来展望	15

图 目 录

图 1: 肘部法则确定 K-means 最佳 K 值.....	5
图 2: 间隔统计量确定 K-means 最佳 K 值.....	5
图 3: ETF 分层聚类结果.....	8
图 4: 因子相关性	11
图 5: ETF 轮动模型相对收益分层回测对比	12
图 6: 不同 N 值的 ETF 轮动模型相对收益分层回测对比	14

表 目 录

表 1: ETF 筛选结果.....	6
表 2: 申万一级行业对应 ETF	9
表 3: ETF 轮动模型使用因子介绍	10
表 4: 各类因子对应统计量	10
表 5: 各类因子对应分层回测结果	11
表 6: ETF 轮动模型因子回测表现对比.....	12
表 7: ETF 轮动模型分年度收益统计	13
表 8: 不同 N 值的 ETF 轮动模型回测表现对比.....	13
表 9: 不同 N 值的 ETF 轮动模型分年度收益统计	14
表 10: ETF 轮动模型近半年持仓情况	14

1. 前言

在以往有关行业轮动策略的系列研究报告里，我们详细研究了如何运用各种数据，如：量价数据、行业基本面数据、资金流向数据及市场一致预期数据等等来建立行业轮动模型的策略，实现了模型收益与稳定性的不断优化。然而，在实际投资中，直接投资于行业指数往往不可行，因此我们需要借助资本市场的工具来实现这一目标。ETF 基金作为一种高效、灵活的投资工具，成为了理想的选择。

在本篇报告中，我们采用了两种聚类算法对现有的行业主题类 ETF 基金进行了分类筛选，旨在识别出最具代表性的 ETF 基金。对于筛选完成后的 ETF 基金，我们仿照行业轮动模型的建立方法，提取量价、基本面和资金面三大类数据，构建专门针对 ETF 基金的轮动模型。该模型不仅能够在不同市场条件下提供稳定的超额收益，还能帮助投资者更精准地捕捉行业轮动带来的投资机会。

2. ETF 的聚类与筛选

2.1 K-means 聚类

我们提取了 A 股市场中所有行业类与主题类 ETF 基金共 492 只。首先，剔除规模不足 1 亿以及成立不足 3 个月的基金。然后，对于跟踪同一指数的 ETF 基金，综合其规模、流动性、费率以及基金管理人等因素，仅保留最具竞争力的一只。最后，提取剩余 150 只基金对应指数的日度收益率，对其进行初步的 K-means 聚类分析。

K-means 聚类是一种常用的无监督学习算法。其基本思想是通过迭代的方式找到最佳的簇中心，使得所有数据点到各自簇中心的距离之和最小。

K-means 的实现步骤如下：

1. **选择初始簇中心**：随机选择 K 个数据点作为初始的簇中心。
2. **分配数据点到最近的簇**：计算每个数据点到每个簇中心的距离，并将其分配到距离最近的簇。
3. **更新簇中心**：重新计算每个簇的新簇中心，即该簇中所有数据点的平均值。
4. **重复步骤 2 和 3**：直到簇中心不再发生变化或达到最大迭代次数。

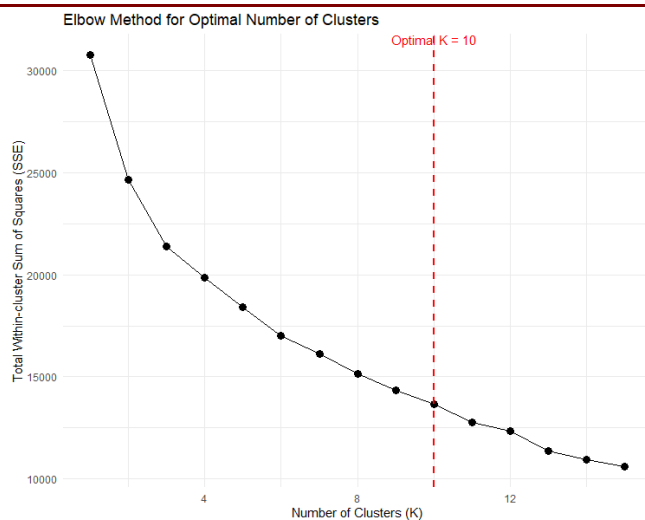
K-means 的算法的优点是简单易懂、计算速度快、结果易于解释；缺点是需要预先指定簇的数量 K，对初始簇中心的选择敏感，并且容易陷入局部最优解。

在确定 K-means 算法的最优 K 值时，常用的方法有以下几种：

1. **肘部法则 (Elbow Method)**：肘部法则是通过观察不同 K 值下的组内平方和 (Within-cluster Sum of Squares, WSS) 的变化来确定最佳 K 值的方法。
2. **轮廓系数 (Silhouette Coefficient)**：轮廓系数是一种评估聚类质量的指标，它考虑了每个数据点与其自身簇内的其他数据点的距离以及与其他簇的距离。
3. **间隔统计量 (Gap Statistic)**：间隔统计量是一种用于确定 K-means 聚类中最佳簇数 (K 值) 的方法。它通过比较实际数据集与参考数据集在不同 K 值下的组内平方和 (Within-cluster Sum of Squares, WSS) 来评估聚类的效果。

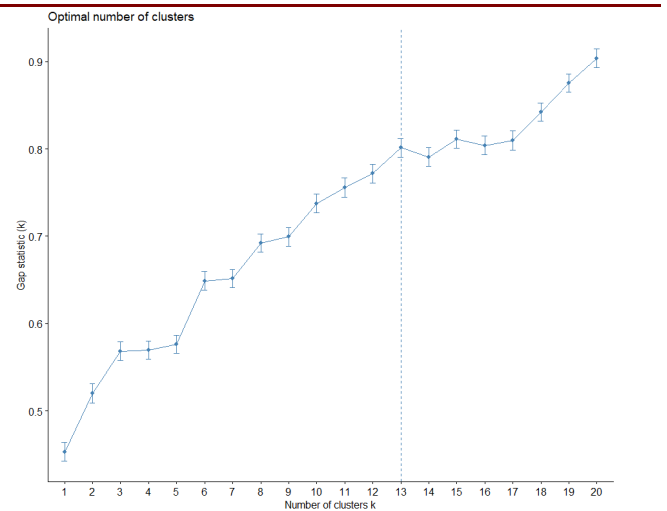
在众多判断最佳 K 值的方法中，肘部法则最为简单直观。但是在实际应用中，当数据集的簇结构不够明显或数据分布较为复杂时，肘部法则可能无法清晰地显示出一个明显的“肘部”点。在 ETF 基金的 K-means 聚类分析中我们正面临了这样的情况。此外，由于 K-means 算法对初始中心的选择非常敏感，不同的随机数种子可能会导致不同的结果。于是，我们综合考量了间隔统计量、轮廓系数和肘部算法的运行结果，来确定 ETF 基金 K-means 聚类的最佳 K 值。最终，我们判断最佳 K 值可能在 10-15 之间。

图 1：肘部法则确定 K-means 最佳 K 值



资料来源：iFind，渤海证券研究所

图 2：间隔统计量确定 K-means 最佳 K 值



资料来源：iFind，渤海证券研究所

借助 K-means 聚类算法，我们对 ETF 备选池进行了进一步筛选。最终选出 12 个大类 40 只具有更强代表性的 ETF 基金。

表 1: ETF 筛选结果

证券代码	证券名称	基金规模 (亿元)	基金管理人	重仓行业	分类结果
512400.SH	有色金属 ETF	45.70	南方基金	有色金属	1
562800.SH	稀有金属 ETF	9.93	嘉实基金	有色金属	1
159713.SZ	稀土 ETF	2.34	富国基金	有色金属	1
159562.SZ	黄金股 ETF	4.28	华夏基金	有色金属	1
510410.SH	资源 ETF	5.17	博时基金	有色金属	1
515220.SH	煤炭 ETF	34.06	国泰基金	煤炭	1
159995.SZ	芯片 ETF	269.36	华夏基金	电子	2
516510.SH	云计算 ETF	12.61	易方达基金	计算机	2
562500.SH	机器人 ETF	12.26	华夏基金	机械设备	2
515880.SH	通信 ETF	27.93	国泰基金	通信	2
159819.SZ	人工智能 ETF	65.74	易方达基金	计算机	2
515000.SH	科技 ETF	32.04	华宝基金	电子	2
512720.SH	计算机 ETF	12.65	国泰基金	计算机	2
512710.SH	军工龙头 ETF	52.16	富国基金	国防军工	3
512800.SH	银行 ETF	49.21	华宝基金	银行	4
515030.SH	新能源车 ETF	58.32	华夏基金	电力设备	5
159790.SZ	碳中和 ETF	24.57	华夏基金	电力设备	5
515790.SH	光伏 ETF	92.11	华泰柏瑞基金	电力设备	5
159611.SZ	电力 ETF	18.60	广发基金	公用事业	5
516970.SH	基建 50ETF	32.18	广发基金	建筑装饰	6
512880.SH	证券 ETF	351.69	国泰基金	非银金融	7
159928.SZ	消费 ETF	111.12	汇添富基金	食品饮料	8
512690.SH	酒 ETF	112.26	鹏华基金	食品饮料	8
159766.SZ	旅游 ETF	30.12	富国基金	社会服务	8
515170.SH	食品饮料 ETF	31.15	华夏基金	食品饮料	8
159936.SZ	可选消费 ETF	2.40	广发基金	汽车	8
516110.SH	汽车 ETF	8.24	国泰基金	汽车	8
159996.SZ	家电 ETF	24.86	国泰基金	家用电器	8
159870.SZ	化工 ETF	16.47	鹏华基金	基础化工	9
159745.SZ	建材 ETF	7.41	国泰基金	建筑材料	9
159825.SZ	农业 ETF	15.98	富国基金	农林牧渔	9
159865.SZ	养殖 ETF	48.19	国泰基金	农林牧渔	9
515210.SH	钢铁 ETF	12.17	国泰基金	钢铁	9
159883.SZ	医疗器械 ETF	19.05	永赢基金	医药生物	10
512010.SH	医药 ETF	215.92	易方达基金	医药生物	10
159992.SZ	创新药 ETF	83.57	银华基金	医药生物	10
560080.SH	中药 ETF	23.43	汇添富基金	医药生物	10
512200.SH	房地产 ETF	73.09	南方基金	房地产	11
159869.SZ	游戏 ETF	81.20	华夏基金	传媒	12
512980.SH	传媒 ETF	40.45	广发基金	传媒	12

资料来源: iFind, 渤海证券研究所

2.2 分层聚类

在将 ETF 备选池缩小至 40 只之后，我们进一步采用了分层聚类（Hierarchical Clustering）的算法对剩余的 ETF 进行了更为细致的分类。

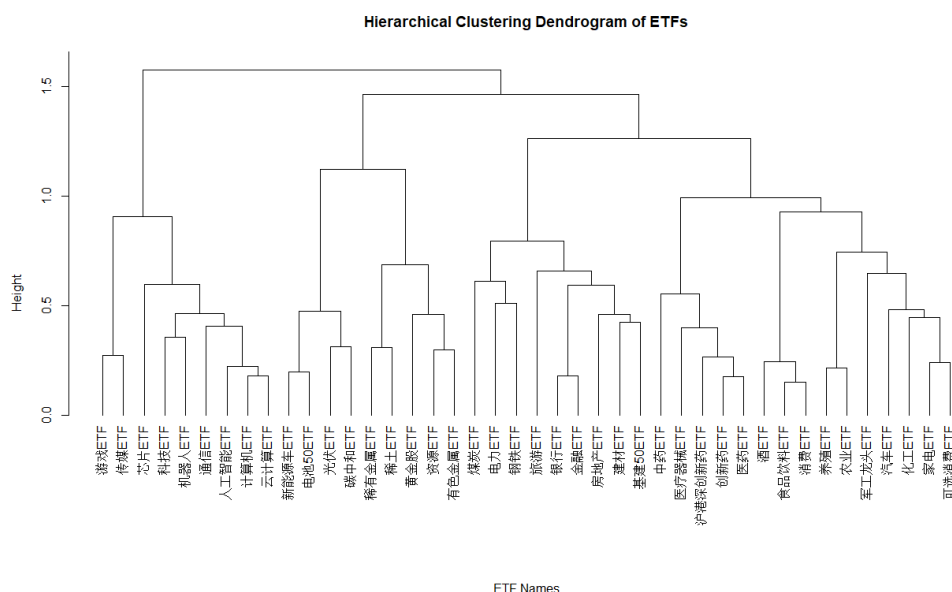
分层聚类是一种聚类分析技术，它通过构建一个树状结构来系统地组织和展示观察数据之间的关系。与 K-means 聚类不同，分层聚类无需预先确定聚类的数量，而是依据数据本身的相似性程度，自下而上或自上而下地逐步形成聚类。

分层聚类的具体步骤如下：

- a. **计算距离矩阵：**首先，计算每一对 ETF 之间的欧几里得距离，组成距离矩阵，作为后续聚类分析的基础。
- b. **执行分层聚类：**基于上述距离矩阵，算法开始按照预设的相似性度量标准，逐步合并最相似的数据点或集群，直至所有数据点最终合并为一个单一的簇。
- c. **绘制树状图：**随着聚类过程的推进，每次合并都会在树状图中表现为一个新的分支点，从而形成一个完整的层次结构
- d. **剪切树状图以获取聚类：**最后，根据分析需求或特定阈值，“剪切”树枝，以此获得所需的聚类数量和成员组成。

分层聚类的优点是可以使用树状图直观地展示不同水平的聚类关系，帮助我们理解 ETF 分类的具体层级结构。ETF 备选池分层聚类结果如下：

图 3: ETF 分层聚类结果



资料来源: iFind, 渤海证券研究所

2.3 申万一级行业分类

对于最终选出的 40 只 ETF 基金, 我们通过分析其持仓数据, 将其与申万一级行业进行了匹配。我们发现, 在市值排名前 20 的申万一级行业中, 每一个行业都能够找到至少一只较为合适的追踪 ETF 基金。而在医药、计算机和有色金属等重点行业中, 不仅有追踪整个行业的 ETF 基金产品可供选择, 还可以找到追踪二级子行业的各种 ETF 基金。ETF 基金的布局日益全面且精细化, 为投资者提供了更加高效和精准的资产配置工具。

另一方面, 在 31 个申万一级行业中, 有 6 个行业当前尚未找到对应的 ETF 基金, 这些行业是: 商贸零售、轻工制造、环保、纺织服饰、美容护理以及综合。这 6 个行业相对较为细分或市场规模较小, 目前仍缺乏专门的 ETF 产品覆盖。

表 2：申万一级行业对应 ETF

证券名称	市值占比	流通市值 (亿元)	对应 ETF
银行	11.87%	93,012.35	金融 ETF、银行 ETF
电子	8.97%	70,235.80	科技 ETF、芯片 ETF
非银金融	7.52%	58,906.61	证券 ETF
医药生物	6.51%	51,022.79	医药 ETF、医疗器械 ETF、中药 ETF、创新药 ETF
电力设备	6.23%	48,806.30	新能源车 ETF、光伏 ETF、碳中和 ETF
食品饮料	6.13%	47,994.55	消费 ETF、食品饮料 ETF、酒 ETF
计算机	4.51%	35,370.43	计算机 ETF、人工智能 ETF、云计算 ETF
机械设备	4.22%	33,028.71	机器人 ETF
汽车	4.14%	32,459.16	汽车 ETF、可选消费 ETF
公用事业	3.93%	30,822.04	电力 ETF
基础化工	3.75%	29,379.80	化工 ETF
交通运输	3.67%	28,777.14	旅游 ETF (33.7%)
石油石化	3.64%	28,516.16	资源 ETF (20.4%)
有色金属	2.99%	23,402.73	资源 ETF、黄金股 ETF、有色金属 ETF、稀土 ETF、稀有金属 ETF
国防军工	2.60%	20,359.62	军工龙头 ETF
通信	2.24%	17,514.58	通信 ETF
家用电器	2.19%	17,129.31	家电 ETF
煤炭	2.12%	16,586.23	煤炭 ETF
建筑装饰	2.09%	16,350.66	基建 50ETF
传媒	1.70%	13,309.25	传媒 ETF、游戏 ETF
房地产	1.46%	11,404.00	房地产 ETF
商贸零售	1.16%	9,118.61	
农林牧渔	1.15%	9,041.51	农业 ETF、养殖 ETF
钢铁	1.00%	7,806.64	钢铁 ETF
轻工制造	0.97%	7,624.61	
环保	0.82%	6,404.06	
建筑材料	0.75%	5,839.07	建材 ETF
纺织服饰	0.72%	5,624.36	
社会服务	0.51%	3,989.31	旅游 ETF
美容护理	0.32%	2,484.95	
综合	0.14%	1,081.84	

资料来源：iFind，渤海证券研究所

3. ETF 轮动模型因子的选择

我们参考行业轮动模型的研究成果，选定量价、基本面和资金面三大类共计 14 个因子。提取 ETF 基金对应指数 2020 年至今的周度因子数据，用于建立 ETF 轮动模型。

综合 T 检验、IC 检验、分层回测和因子相关性测试结果，我们最终选出了单因子检测中表现较好，彼此相关性较低，并且本身具有代表性的 5 个因子，构成 ETF 轮动模型，这些因子是：指数 250 日涨跌幅因子、BIAS 乖离率因子、EPS 增速环比因子、主力净流入因子和 10 日大单净买入因子。

表 3：ETF 轮动模型使用因子介绍

因子大类	因子名称	因子定义
量价因子	mom20	指数 20 日涨跌幅
	mom250	指数 250 日涨跌幅
	RSI	相对强弱指标
	BIAS	乖离率
基本面因子	ROE_TTM	滚动 12 个月的净资产收益率
	ROE_Q	最新一期的净资产收益率
	NP_YOY	最新一期的净利润增长率
	净利润增速环比	净利润同比增长率的环比增速
	ROE 增速环比	ROE_TTM 同比增长率的环比增速
	EPS 增速环比	最新一期 EPS 同比增长率的环比增速
资金面因子	净买入额	指数最新一期净买入额
	净主动买入额	指数最新一期净主动买入额
	主力净流入	指数最新一期主力净流入
	DDE10	指数 10 日大单净买入量

资料来源：iFind，渤海证券研究所

表 4：各类因子对应统计量

因子大类	因子名称	t 值绝对值 平均值	t 值绝对 值>2 概率	IC 平均值	IC 标准差	IR	IC>0 概率
量价因子	mom5	1.797	0.196	0.009	0.313	0.029	52.24%
	mom20	1.995	0.224	0.014	0.327	0.042	53.47%
	mom250	2.260	0.265	0.023	0.354	0.065	52.65%
	RSI	1.931	0.200	0.005	0.322	0.017	49.80%
	BIAS	1.948	0.237	0.017	0.316	0.055	52.65%
基本面因子	ROE_TTM	1.421	0.110	-0.009	0.267	-0.035	49.39%
	ROE_Q	1.403	0.122	-0.007	0.279	-0.025	48.16%
	NP_YOY	1.140	0.078	-0.009	0.274	-0.034	49.80%
	净利润增速	0.815	0.029	0.006	0.206	0.030	50.61%
	EPS 增速	0.894	0.057	0.019	0.207	0.094	55.51%

请务必阅读正文之后的声明

10 of 18

	ROE 增速	0.815	0.041	0.019	0.207	0.090	50.61%
资金面因子	净买入额	1.648	0.188	0.042	0.292	0.144	55.10%
	净主动买入额	1.754	0.204	0.046	0.305	0.151	55.92%
	主力净流入	1.638	0.200	0.039	0.297	0.132	55.51%
	DDE	1.413	0.159	0.030	0.267	0.113	55.10%

资料来源：iFind，渤海证券研究所

表 5：各类因子对应分层回测结果

因子名称	累计收益	年化收益	波动率	最大回撤	夏普比率	信息比率	胜率
mom5	18.39%	3.65%	22.75%	43.63%	0.160	0.103	51.43%
mom20	56.12%	9.92%	24.34%	40.44%	0.407	0.601	56.33%
mom250	48.55%	8.76%	24.81%	41.04%	0.353	0.464	54.69%
RSI	67.91%	11.63%	23.40%	37.30%	0.497	0.709	51.43%
BIAS	89.41%	14.52%	23.86%	28.79%	0.609	0.974	55.10%
ROE_TTM	-18.75%	-4.31%	22.04%	56.49%	-0.196	-0.674	45.31%
ROE_Q	-15.22%	-3.44%	22.30%	54.39%	-0.154	-0.565	50.20%
NP_YOY	-0.96%	-0.20%	22.92%	50.31%	-0.009	-0.254	51.43%
净利润增速	9.23%	1.89%	20.55%	42.22%	0.092	-0.065	50.20%
EPS 增速	56.39%	9.96%	22.13%	36.56%	0.450	0.915	60.00%
ROE 增速	34.03%	6.41%	22.57%	43.91%	0.284	0.519	54.29%
净买入额	56.17%	9.92%	23.41%	30.70%	0.424	0.671	53.88%
净主动买入额	41.20%	7.60%	23.47%	29.19%	0.324	0.442	53.88%
主力净流入	55.01%	9.75%	22.94%	26.20%	0.425	0.648	53.47%
DDE	80.91%	13.41%	21.18%	34.65%	0.633	1.108	52.65%
平均	11.81%	2.40%	20.11%	39.07%	0.119	--	--

资料来源：iFind，渤海证券研究所

图 4：因子相关性

	mom5	mom20	mom250	RSI	DDE	BIAS	净利润增速	EPS 增速	ROE 增速	净买入额	净主动买入额	主力净流入
mom5	1.00	0.49	0.31	0.62	0.45	0.76	0.13	0.11	0.12	0.34	0.37	0.37
mom20	0.49	1.00	0.34	0.86	0.32	0.83	0.14	0.12	0.12	0.26	0.28	0.25
mom250	0.31	0.34	1.00	0.41	0.25	0.32	0.19	0.11	0.12	0.28	0.28	0.27
RSI	0.62	0.86	0.41	1.00	0.33	0.86	0.12	0.12	0.10	0.24	0.29	0.25
DDE	0.45	0.32	0.25	0.33	1.00	0.48	0.13	0.14	0.11	0.39	0.35	0.42
BIAS	0.76	0.83	0.32	0.86	0.48	1.00	0.13	0.12	0.11	0.27	0.30	0.28
净利润增速	0.13	0.14	0.19	0.12	0.13	0.13	1.00	0.53	0.46	0.13	0.15	0.15
EPS 增速	0.11	0.12	0.11	0.12	0.14	0.12	0.53	1.00	0.90	0.12	0.12	0.12
ROE 增速	0.12	0.12	0.12	0.10	0.11	0.11	0.46	0.90	1.00	0.10	0.11	0.10
净买入额	0.34	0.26	0.28	0.24	0.39	0.27	0.13	0.12	0.10	1.00	0.91	0.97
净主动买入额	0.37	0.28	0.28	0.29	0.35	0.30	0.15	0.12	0.11	0.91	1.00	0.92
主力净流入	0.37	0.25	0.27	0.25	0.42	0.28	0.15	0.12	0.10	0.97	0.92	1.00

资料来源：iFind，渤海证券研究所

4. ETF 轮动模型的建立

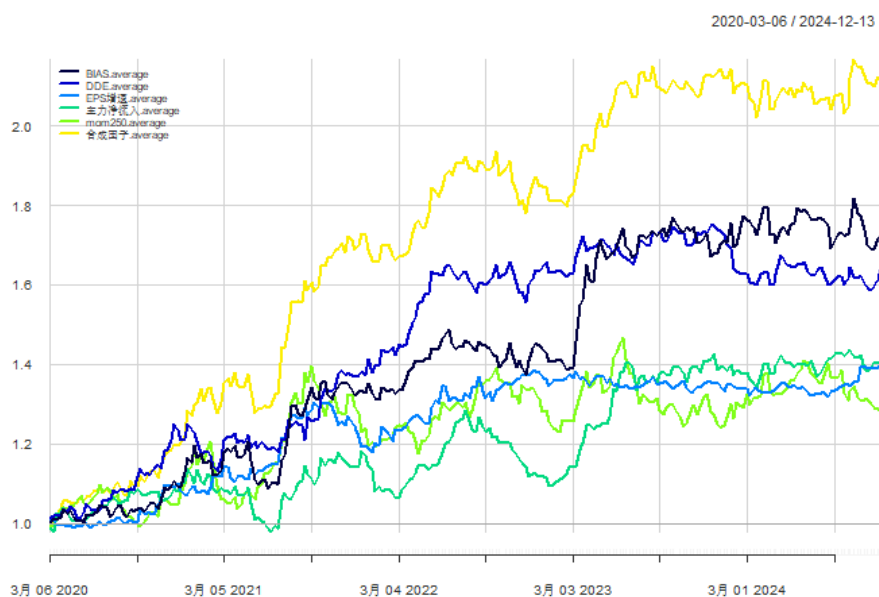
我们将最终选出的 5 个因子等权合成，构建最终的 ETF 轮动模型。合成因子对于模型收益的提升十分显著。2020 年至今年化收益 19.70%，年化超额收益 17.30%，信息比率 1.526，周度胜率 59.18%。表现显著跑赢所有单因子模型。

表 6：ETF 轮动模型因子回测表现对比

	累计收益	年化收益	波动率	最大回撤	夏普比率	信息比率	胜率
合成因子	133.29%	19.70%	23.13%	29.98%	0.852	1.526	59.18%
mom250	48.55%	8.76%	24.81%	41.04%	0.353	0.464	54.69%
主力净流入	55.01%	9.75%	22.94%	26.20%	0.425	0.648	53.47%
EPS 增速	56.39%	9.96%	22.13%	36.56%	0.450	0.915	60.00%
DDE	80.91%	13.41%	21.18%	34.65%	0.633	1.108	52.65%
BIAS	89.41%	14.52%	23.86%	28.79%	0.609	0.974	55.10%
平均	11.81%	2.40%	20.11%	39.07%	0.119	--	--

资料来源：iFind，渤海证券研究所

图 5：ETF 轮动模型相对收益分层回测对比



资料来源：iFind，渤海证券研究所

分年度统计来看，复合因子在市场强势的 2020 和 2021 年大幅跑赢基准，在市场弱势的 2022 和 2023 年也表现出了较好的防守性。过去 5 年中，仅仅在 2024 年跑输基准 2.68%，但年度收益仍达到了 8.82%。

表 7：ETF 轮动模型分年度收益统计

年份	合成因子	mom250	主力净流入	EPS 增速	DDE	BIAS	平均
2020	67.20%	52.63%	41.83%	40.95%	58.10%	54.13%	28.67%
2021	37.62%	6.55%	9.88%	14.84%	21.87%	19.69%	5.87%
2022	-6.90%	-8.54%	-16.56%	-1.96%	-1.03%	-9.51%	-14.89%
2023	0.08%	-15.62%	6.96%	-14.28%	-8.57%	1.04%	-13.50%
2024	8.82%	18.35%	11.45%	14.96%	3.77%	12.29%	11.50%

资料来源：iFind，渤海证券研究所

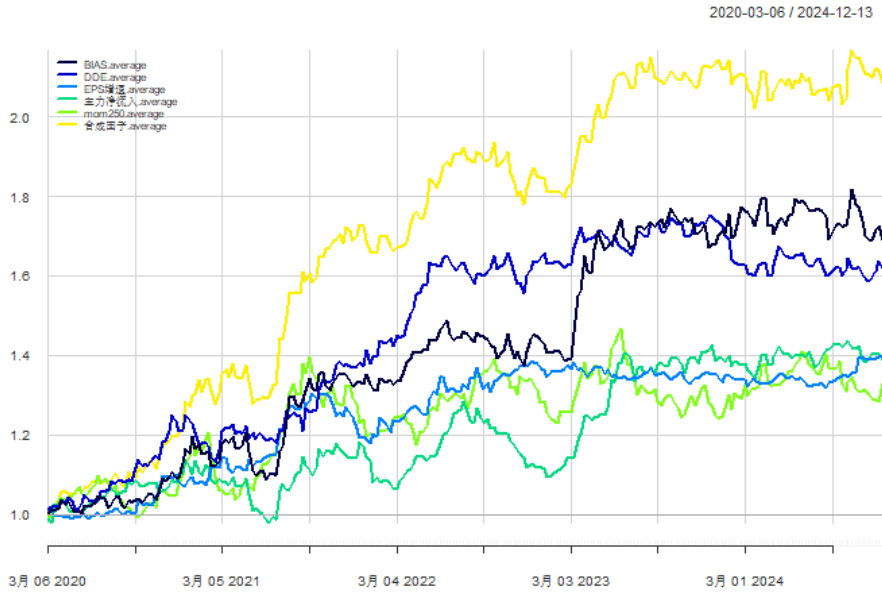
我们还针对 ETF 轮动模型每期选定基金数量 N 做了遍历性敏感度测试。测试表明，当每期选定基金数量为 3 至 5 只时，模型在收益和波动的平衡型上表现更好。不过，无论选定的基金数量如何变化，模型相对于基准始终保持了明显的超额收益，证明我们的 ETF 轮动策略具有较好的稳定性和适应性。

表 8：不同 N 值的 ETF 轮动模型回测表现对比

N 值	累计收益	年化收益	波动率	最大回撤	夏普比率	信息比率	胜率
2	217.45%	27.78%	31.10%	41.79%	0.893	1.150	51.02%
3	282.04%	32.91%	28.36%	30.45%	1.160	1.639	59.18%
4	176.87%	24.13%	26.43%	31.62%	0.913	1.345	55.92%
5	165.99%	23.08%	25.38%	32.91%	0.909	1.385	54.69%
6	129.13%	19.24%	24.55%	34.58%	0.784	1.239	55.51%
7	129.26%	19.26%	23.58%	31.39%	0.817	1.379	57.55%
8	133.29%	19.70%	23.13%	29.98%	0.852	1.526	59.18%
平均	11.81%	2.40%	20.11%	39.07%	0.119	--	--

资料来源：iFind，渤海证券研究所

图 6: 不同 N 值的 ETF 轮动模型相对收益分层回测对比



资料来源: iFind, 渤海证券研究所

表 9: 不同 N 值的 ETF 轮动模型分年度收益统计

年份	N=2	N=3	N=4	N=5	N=6	N=7	N=8	平均
2020	90.22%	108.15%	95.21%	86.07%	76.34%	69.03%	67.20%	28.67%
2021	62.03%	72.82%	49.12%	47.83%	43.05%	39.92%	37.62%	5.87%
2022	-2.28%	-12.64%	-12.34%	-6.24%	-9.79%	-8.56%	-6.90%	-14.89%
2023	-6.66%	0.02%	-2.79%	-2.12%	-4.83%	-0.04%	0.08%	-13.50%
2024	12.91%	21.54%	11.62%	5.37%	5.79%	6.05%	8.82%	11.50%

资料来源: iFind, 渤海证券研究所

ETF 轮动模型最近半年的持仓情况如下:

表 10: ETF 轮动模型近半年持仓情况

时间	第一名	第二名	第三名	第四名	第五名
2024/7/5	黄金股 ETF	汽车 ETF	养殖 ETF	化工 ETF	钢铁 ETF
2024/7/12	汽车 ETF	黄金股 ETF	芯片 ETF	房地产 ETF	化工 ETF
2024/7/19	芯片 ETF	食品饮料 ETF	酒 ETF	消费 ETF	汽车 ETF
2024/7/26	家电 ETF	军工龙头 ETF	医疗器械 ETF	汽车 ETF	证券 ETF
2024/8/2	养殖 ETF	医疗器械 ETF	家电 ETF	创新药 ETF	基建 50ETF
2024/8/9	房地产 ETF	医疗器械 ETF	黄金股 ETF	家电 ETF	养殖 ETF
2024/8/16	医疗器械 ETF	煤炭 ETF	银行 ETF	家电 ETF	酒 ETF
2024/8/23	家电 ETF	黄金股 ETF	煤炭 ETF	资源 ETF	可选消费 ETF
2024/8/30	汽车 ETF	家电 ETF	游戏 ETF	化工 ETF	可选消费 ETF
2024/9/6	汽车 ETF	证券 ETF	钢铁 ETF	医疗器械 ETF	游戏 ETF
2024/9/13	证券 ETF	汽车 ETF	钢铁 ETF	医疗器械 ETF	黄金股 ETF
2024/9/20	黄金股 ETF	有色金属 ETF	云计算 ETF	家电 ETF	汽车 ETF
2024/9/27	证券 ETF	云计算 ETF	科技 ETF	有色金属 ETF	计算机 ETF

请务必阅读正文之后的声明

14 of 18

2024/9/30	证券 ETF	云计算 ETF	芯片 ETF	计算机 ETF	新能源车 ETF
2024/10/11	钢铁 ETF	建材 ETF	煤炭 ETF	稀土 ETF	家电 ETF
2024/10/18	芯片 ETF	证券 ETF	云计算 ETF	军工龙头 ETF	科技 ETF
2024/10/25	稀土 ETF	稀有金属 ETF	新能源车 ETF	碳中和 ETF	军工龙头 ETF
2024/11/1	稀土 ETF	钢铁 ETF	有色金属 ETF	稀有金属 ETF	农业 ETF
2024/11/8	云计算 ETF	钢铁 ETF	人工智能 ETF	军工龙头 ETF	计算机 ETF
2024/11/15	传媒 ETF	游戏 ETF	汽车 ETF	银行 ETF	钢铁 ETF
2024/11/22	传媒 ETF	游戏 ETF	稀土 ETF	黄金股 ETF	电力 ETF
2024/11/29	传媒 ETF	人工智能 ETF	云计算 ETF	银行 ETF	机器人 ETF
2024/12/6	钢铁 ETF	银行 ETF	游戏 ETF	资源 ETF	建材 ETF
2024/12/13	钢铁 ETF	游戏 ETF	养殖 ETF	电力 ETF	汽车 ETF

资料来源: iFind, 渤海证券研究所

5. 总结与未来展望

本篇报告中, 我们使用 K-means 聚类算法和分层聚类算法, 对目前市场上的 492 只行业主题类 ETF 基金进行了筛选和分类。最终, 我们选出了 12 个大类 40 只 ETF 基金产品。这些 ETF 基金覆盖了 31 个申万一级行业中的 25 个。其中不仅有追踪整个行业的 ETF 基金, 还有追踪二级子行业的 ETF 基金。为投资者提供了更加高效和精准的资产配置工具。

我们提取量价、基本面和资金面三大类共计 14 个因子 2020 年至今的周度数据, 进行单因子检测。最终, 我们选出了单因子检测中表现较好, 彼此相关性较低, 并且本身具有代表性的 5 个因子组成 ETF 轮动模型。这 5 个因子为: 指数 250 日涨跌幅因子、BIAS 乖离率因子、EPS 增速环比因子、主力净流入因子和 10 日大单净买入因子。

我们将最终选出的 5 个因子等权合成, 构建 ETF 轮动模型。模型对于收益的提升十分显著, 2020 年至今年化收益 19.70%, 年化超额收益 17.30%, 信息比率 1.526, 周度胜率 59.18%, 显著跑赢所有单因子模型。

我们还针对 ETF 轮动模型每期选定基金数量 N 做了遍历性敏感度测试。测试表明, 当每期选定基金数量为 3 至 5 只时, 模型在收益和波动的平衡型上表现更好。不过, 无论选定的基金数量如何变化, 模型相对于基准始终保持了明显的超额收益, 证明我们的 ETF 轮动策略具有较好的稳定性和适应性。

未来, 我们将继续致力于模型的精进优化工作, 持续挖掘和整合多元化因子, 构建更为精准稳定的 ETF 轮动模型。

风险提示：

- 1、**市场风格转换风险：**未来如果发生市场波动和市场风格转换，模型可能失效。
- 2、**第三方数据提供不准确风险：**模型数据来源为公开的历史数据，可能存在第三方数据提供不准确风险。
- 3、**历史结果不代表未来收益：**模型基于历史规律总结，历史结果不代表未来收益。

本报告不构成投资建议。

分析师声明:

本报告署名分析师在此声明:我们具有中国证券业协会授予的证券投资咨询执业资格并登记为证券分析师,以勤勉尽责的职业态度、专业审慎的研究方法,使用合法合规的数据和信息,独立、客观地出具本报告;本报告所表述的任何观点均精准地、如实地反映研究人员的个人观点,结论不受任何第三方的授意或影响。我们所获取报酬的任何部分无论是在过去、现在及将来均不会与本报告中的具体投资分析意见或观点有直接或间接的联系。

风险提示及免责声明:

投资者应自主作出投资决策并自行承担投资风险,任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本报告由渤海证券股份有限公司(以下简称“本公司”)制作,仅供本公司的客户使用。发送本报告不构成任何合同或承诺的基础,不因接收者收到本报告而视其为本公司客户。本报告仅在相关法律许可的情况下发送,并仅为提供信息而发送,不构成任何广告。

本报告的信息来源于已公开的资料,本公司对该等信息的准确性、完整性或可靠性不作任何保证,不保证该信息未经任何更新。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断,本报告所指的证券或投资标的的价格、价值及投资收入可升可跌。过往表现不应作为日后表现的依据。在不同时期,本公司可发出与本报告所载资料、意见及推测不一致的报告。本公司不保证本报告所含信息保持在最新状态。同时,本公司对本报告所含信息可在不发出通知的情形下做出修改,投资者应当自行关注相应的更新或修改。

在任何情况下,本报告内容的全部或部分均不构成对任何人的投资建议。本报告中所指的投资及服务可能不适合个别客户,不构成客户私人咨询建议。在任何情况下,本公司、本公司员工或者关联机构不承诺投资者一定获利,不与投资者分享投资收益,也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意,其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。在任何情况下,本公司、本公司员工或者关联机构不就本报告中的任何内容对任何投资做出任何形式的担保。

在所知情的范围内,本公司、本报告撰写人以及财产上的利害关系人与本报告所评价或作出明确估值和投资评级的证券无利害关系。本公司利用信息隔离墙控制内部一个或多个领域、部门或关联机构之间的信息流动。因此,投资者应注意,在法律许可的情况下,本公司及其所属关联机构可能持有并交易本报告中所提公司的股份或其他财产权益,也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。

市场有风险,投资需谨慎。投资者不应将本报告作为作出投资决策的唯一参考因素,亦不应认为本报告可以取代自己的判断。在决定投资前,如有需要,投资者务必向专业人士咨询并谨慎决策。

本报告版权为本公司所有。未经本公司事先书面许可,任何机构和/或个人不得以任何形式刊载、转发、翻版、复制、发布或引用本报告全部或部分内容,亦不得从未经本公司书面授权的任何机构、个人或其运营的媒体平台接收、翻版、复制或引用本报告全部或部分内容。如征得本公司同意进行引用、刊载或转发的,需在允许的范围内使用,并注明出处为“渤海证券股份有限公司”且不得对本报告进行任何有悖原意的引用、删节和修改。若本公司以外的其他机构(以下简称“该机构”)发送本报告,则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息或进而交易本报告中提及的证券。本报告不构成本公司向该机构之客户提供的投资建议,本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

投资评级说明：

项目名称	投资评级	评级说明
公司评级标准	买入	未来 6 个月内相对沪深 300 指数涨幅超过 20%
	增持	未来 6 个月内相对沪深 300 指数涨幅介于 10%~20%之间
	中性	未来 6 个月内相对沪深 300 指数涨幅介于-10%~10%之间
	减持	未来 6 个月内相对沪深 300 指数跌幅超过 10%
行业评级标准	看好	未来 12 个月内相对于沪深 300 指数涨幅超过 10%
	中性	未来 12 个月内相对于沪深 300 指数涨幅介于-10%-10%之间
	看淡	未来 12 个月内相对于沪深 300 指数跌幅超过 10%

渤海证券研究所机构销售团队：
高级销售经理：朱艳君

座机：+86 22 2845 1995

手机：135 0204 0941

邮箱：zhuyanjun@bhqz.com

天津：

天津市南开区水上公园东路宁汇大厦 A 座写字楼

邮政编码：300381

电话：+86 22 2845 1888

传真：+86 22 2845 1615

高级销售经理：王文君

座机：+86 10 6810 4637

手机：186 1170 5783

邮箱：wangwj@bhqz.com

北京：

北京市西城区西直门外大街甲 143 号 凯旋大厦 A 座 2 层

邮政编码：100086

电话：+86 10 6810 4192

传真：+86 10 6810 4192

渤海证券股份有限公司网址： www.ewww.com.cn