

Quantitative Investment Study of Crowding Factor from Sector Rotation Perspective

Sichen Ye

College of Economics and Management, China Jiliang University, Hangzhou, Zhejiang 310018, China

Email: 2564173140@qq.com

Abstract

This paper focuses on the crowding phenomenon existing in the A-share market. By fitting the most effective market crowding indicator through relevant validity tests of crowding proxy variables, high-quality industries and individual stocks in rotation using multi-factor composite crowding indicator then selects are selected to construct a dynamic stock pool. A return test of the crowding factor strategy is conducted using quantitative backtesting software to prove its validity and relevant ability. The risk of overheating in various industries is quantitatively analyzed through the composite congestion indicator, so as to construct industry allocation strategy. The continuous congestion indicators are constructed from the perspective of industry index and constituent stocks, and the validity of congestion indicators is verified using the logistic restriction regression method, in order to explore the indicator that can stably reveal the overheating state of market trading. After backtesting and optimization, the final composite congestion indicator is formed, according to which safe and risky industries are screened, and the final industry allocation strategy is constructed, which proves the uniqueness of congestion in the field of market trading.

Keywords: *A-share Market; Sector Rotation; Congestion Proxy Variables; Quantitative Trading*

行业轮动视角下的拥挤度因子量化投资研究

叶思辰

中国计量大学，经济与管理学院，杭州 310018

摘要: 本文主要对A股市场中存在的拥挤度现象进行了研究。通过对拥挤度代理变量进行相关有效性测试，拟合出最有效的市场拥挤度指标，随后采用多因子复合拥挤度指标轮动选择优质行业与个股，构建动态股票池，采用量化回测软件对拥挤度因子策略进行收益测试，证明其有效性与相关能力。本策略主要通过复合拥挤度指标来定量分析各行业的交易过热风险，构建行业配置策略。从行业指数和成分股两个角度进行连续型拥挤度指标的构建，采用逻辑限制回归方法对拥挤度指标的有效性进行验证，寻找能够稳定揭示市场交易过热状态的指标。经过回测验证并优化完善形成最终的复合拥挤度指标，据此筛选安全行业与风险行业，构建最终行业配置策略，证明了拥挤度因子在市场交易领域的独特性。

关键词: A股市场；行业轮动；拥挤度代理变量；量化交易

引言

相较于传统交易，量化投资存在着效率高、交易和盈利机会多、结果直观等诸多优势；较国外数十年的量化投资发展而言，国内的量化交易仍是新兴投资模式，但其优异的风险规避与盈利能力让它拥有主观交易所难以企及的高度。同时，中国股票市场的数据优势与市场优势使得量化投资与交易发挥独到优势、在中国金融市场上大有可为。经过几十年发展，专业人士在市场中发掘出了众多指标、并以此构建了许多相关策略，本文专注于拥挤度指标的研究，创新性的构建了复合拥挤度指标并以此配合择时和风控进行量化回测，证明了拥挤度指标在股市交易中的有效性，相对于市场指数获得的超额收益，可以为投资者更好的搭建策略框架

提供一定参考。通过对国内外关于因子拥挤特征的研究成果进行深入综合与对比分析，发现股票市场的核心拥挤特征主要体现在风险溢价、敞口暴露以及相关联动性等方面。因子拥挤度的研究为量化资金市场提供了新的参考维度，使投资者能够及时发现策略的超配效应，进而避免损失的进一步扩大。本文在国内外有关学者和量化分析师对拥挤度指标的研究基础上，进一步提出了拥挤度指标的实际运用方案并进行了相关测试，拥挤度的研究对于揭示市场效应及预测能力具有至关重要的意义。

1 宏观背景

2022 年末以来，受全球经济大环境和国内各项政策影响，我国股市正处于大幅波动的调整阶段，市场体量大，风格轮动速度快，股市交投过旺，保持供不应求的态势，市场规范尚待完善，市场投机行为仍占较大比重。A 股指数在 23 年中结束上升趋势，随后到达阶段峰顶 3400 点附近，之后大盘与多数个股进入下跌阶段，在年底跌破 3000 点重要关口，在 24 年初出现加速下行态势，市场情绪弱，亏钱效应明显。

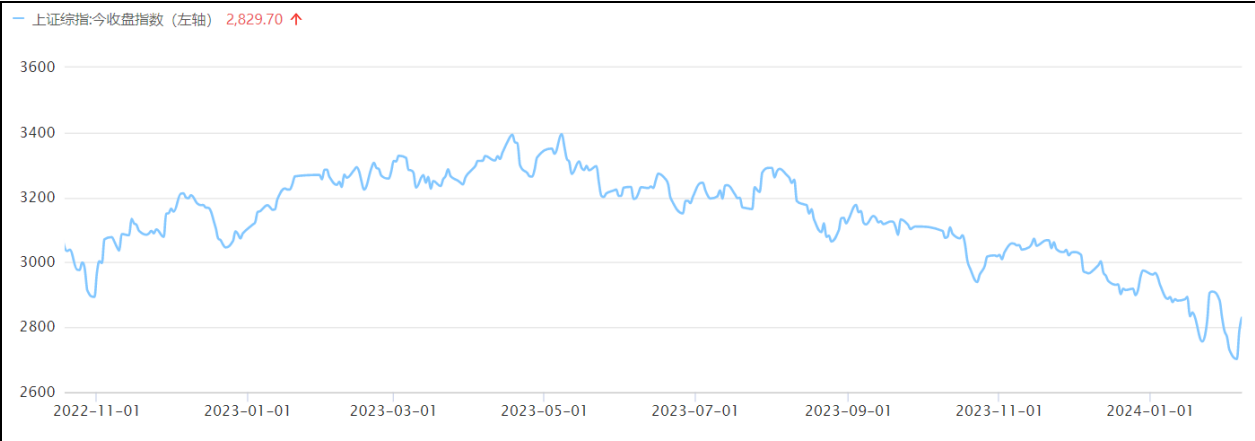


图 1 上证综指 2022.11-2024.02

宏观市场层面，国内形势趋于复杂，国内稳增长政策仍在陆续推出，降准已超预期落地，部分重点城市的商品房限购政策正陆续松绑，但地产消费等相关数据并未明显起色，微观经济对政策的反应明显钝化。23 年以来，我国 CPI 同比增长率总体呈现下降态势，反映了国内消费疲软、投资需求不足，居民的购买能力下降，国内经济有一定的通缩现象预兆。特别是近期 M1-M2 逆剪刀差的扩大，也说明了市场中“死钱”较多，货币流动性在金融体系内空转，货币政策传导至实体经济滞后，面临结构性“流动性陷阱”的挑战。

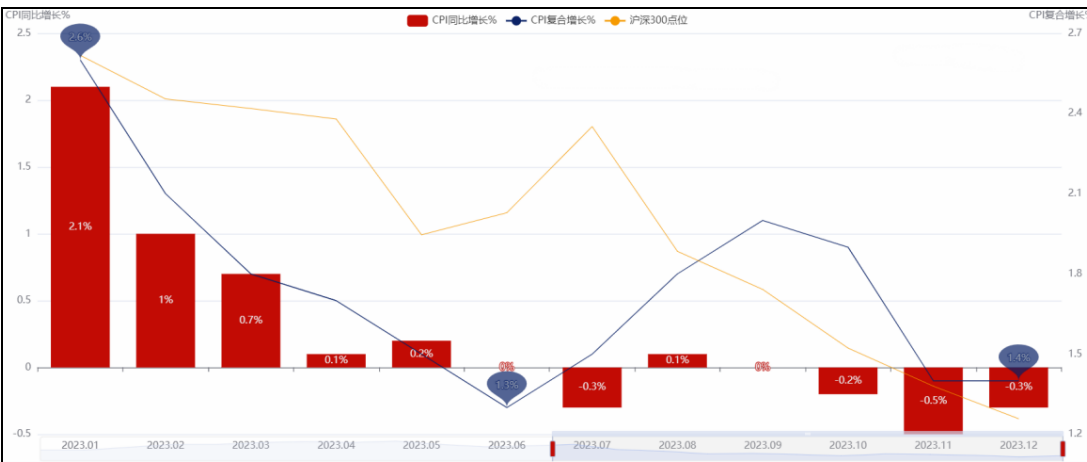


图 2 中国 CPI 同比增长数据（2023 年）

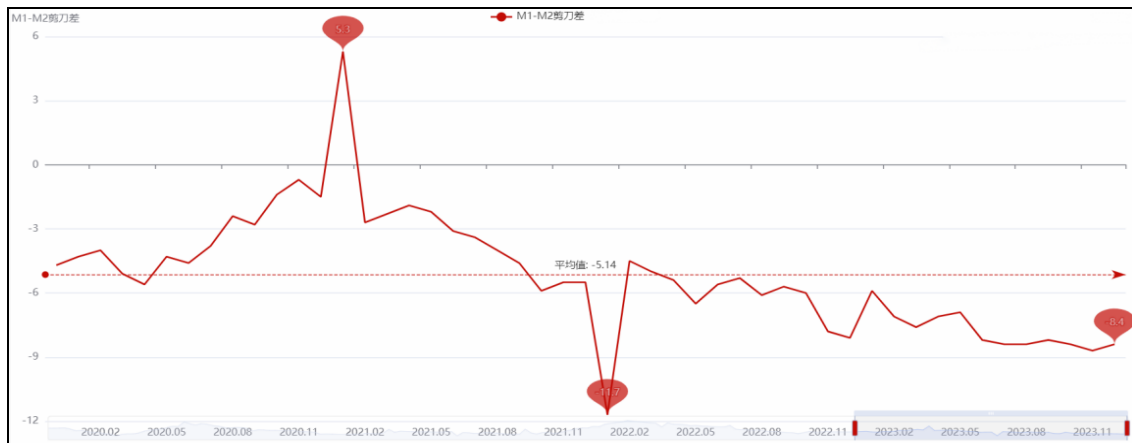


图 3 M1-M2 剪刀差

海外地缘政治复杂严峻，美国大选贯穿全年，鹰派言论对 A 股也将带来间歇性的冲击。除此以外，近期全国多省发布了 2024 年新增 IPO 的计划，一级市场对二级市场的“吸血”现象仍然存在，资本市场尚待一场全面的改革，从规范 IPO、上市公司经营管理乃至股东减持行为等方面，维护二级市场投资者的核心利益。多重因素影响下，市场情绪仍然低迷。

进入 2024 年，市场环境发生变化，板块轮动速度加快，其不确定性持续增大，而整体支撑动力却不断变弱，操作难度逐渐增大，拥有一个稳定且有效应对回撤风险的策略是当下市场的重中之重。行业轮动是 A 股市场上一个经久不息的话题。为了初步验证基于行业轮动策略的整体有效性，首先基于行业轮动假设对 A 股市场进行了 10 年回溯：以 28 个申万一级行业为例，取每年涨幅靠前的 6 个行业等权买入，每一年度终了换仓，手续费设置为双边 3%。从策略的结果来看，2010-2021 期间策略的累计净值为 33.5，年化收益率为 35%，相对申万一级行业的 28.13%，长期大幅跑赢行业等权配置，说明灵活的行业轮动对策略整体收益有正向积极的作用。

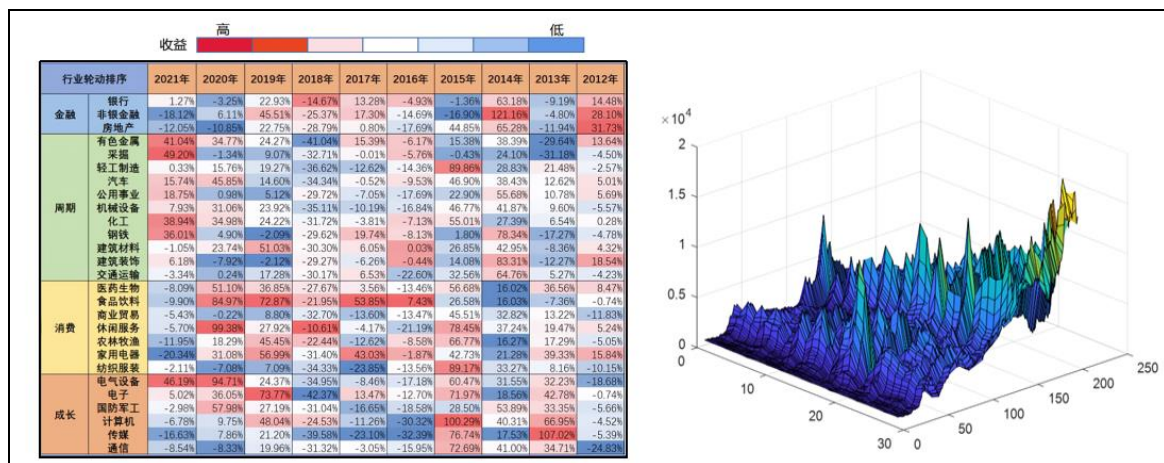


图 4 A 股行业轮动

2 文献综述

“拥挤度”这一概念，首先被广泛运用于交通运输和机械、医学等领域。金融市场的拥挤特征的客观存在表现，早在 2004 年便已经被 Hogan 等人发现，后续的相关深度研究也就从此展开。

2008 年金融危机之后外国对拥挤度的研究便已开始，学者 Pojarliev 与 Lenich^[14]对于货币基金日评数据进行了基金层面拥挤策略的构建，得出了拥挤度与收益率之间的负向关系；Chue 发现在基金市场中过往的高

收益可能会导致过度的回撤；Zhong 等^[15]利用换手率指标对单只个股拥挤度高低对该股回报率产生的影响进行了分析；拥挤指标也被 Bonne、Volpati 等学者在不同层面和不同指数中进行了研究。

国内，王昭栋、张利平针对 A 股财务和交易数据进行了拥挤度研究；贾丽娜等对国内基金拥挤交易进行了研究；刘玉夕根据九类风格因子，通过非线性加权，对市场组合进行了研究；华泰证券的林晓明和陈烨^[7]分析师在单因子测试的系列金工研报中测试了估值类因子、成长类因子、换手率类因子、波动率类因子、动量类因子在中国市场的表现，在每个大类中挖掘出了在中国市场上一些表现较好的子类因子，他们在《确定研究主体：行业解析与分类》一文中确立了目标行业，通过对它们进行划分和分类，随后在《行业全貌描绘：从宏观角度分析》及其后的四篇文献中，从宏观战略风险、中观投资风格、微观管理方式及产业链结构四个不同层面对行业特征进行了深入的剖析与归类，旨在多角度解析关键动力因素，后续研究中，林晓明团队的焦点转向对行业的投资策略布局，分别以宏观环境因素、趋势跟踪以及行业景气水平为核心进行探讨；海通证券的冯佳睿和袁林青分析师也进行了拥挤度指标的研究实证，结果显示拥挤状态的短期抬升和下降与净值的局部高点和低点相关联。国盛证券的林志朋和刘富兵^[9]分析师基于 A 股市场的特征，从交易数据的角度构建了“国盛金工因子拥挤模型”，使用多空换手率、波动率和 beta 比率进行因子拥挤度的衡量指标，研究发现所构建的拥挤度指标可以有效提升组合胜率，提高组合收益。

我国学者目前对拥挤度的研究属于起步阶段，是较新领域，相关研究不多，所以基于 A 股行业轮动视角对拥挤度及其相关指标的研究是有必要的。

3 创新点与研究方法

3.1 研究思路与方法

本文首先测试了常用的刻画拥挤度的代理变量的表现，发现单一指标在识别拥挤度方面存在不足，单一指标通常存在效果不达预期且参数敏感的问题，故在验证阶段自行拟合、构建复合拥挤度指标以更加灵敏和科学地反映市场或个股拥挤情况。

拥挤度可作为判断指数或单一股票是否出现逆转信号的衡量指标。拥挤交易是指一笔资金在同一资产或一系列类似性质的资产中进行大规模的收购和销售，从而造成了一种资产的价格变动。股价往往围绕其内在价值上下波动。在基本面改善、宏观政策持续加大的情况下，优质股的价值中枢不断上升，给投资者提供长远收益。但如果基本面超出了预期或是政策突然变好，股价和成交量就会猛增。过热交易行为使股价在一开始就涨到很高的位置，到达拥挤的状态，容易引起市场的踩踏，有较大下跌风险。

基于理论研究的严谨性使用定量指标来捕获行业的拥挤程度。衡量市场拥挤度的主要指标有：换手率、波动率、乖离率、相对估值、成交金额比率、行业动量等，后续，我们将对其进行进一步的量化与统计分析、回归检验。本策略倾向于选取多个有逻辑基础的原始指标，从多个维度构建应用场景，从而更全面地讨论行业拥挤度相关问题。本报告主要关注行业拥挤度的相关问题，因此后续所有指标的定义都应用于申万一级行业。

在特定行业或领域内，资本注入确实能够推动利润增长，然而，一旦涌入资金超出合理范围，反而会提升利润回撤和波动的风险。鉴于此，为了降低投资组合的风险水平，投资者宜考虑避开那些资金过度密集的行业或市场板块。

2023 年到 24 年初，A 股市场的主要趋势为市场情绪的降温导致了密集持股的解构，彼时共同基金持有的重仓股普遍下挫，发生了一定程度的恐慌性抛售，而分红率较高、股票集中度较低的板块则成了资金的安全港，即便在整体市场走低之际也实现了较为稳健的盈利。当前市场情绪以寻求风险防范为主，正处在一个板块更迭的关键时期，市场动荡有所上升，近期采取守势策略为佳。本策略通过挖掘低拥挤度高收益行业，构建市场优选模型，先于市场热点与投机风波提前布局相应板块中的优质股票。

基于市场的真实情况，构建挖掘低拥挤度高收益行业的策略思路，进行量化投资研究。

通过对量化的产业周期性模型进行改良，专注于市场密集交易现象的影响，并对这种密集度要素进行深入研究，同时借鉴动能效应的原理，认定各产业持续增长的趋势是有延续性的，可以在短时间内通过比较不同产业在截面上的表现来针对性地投资于涨势较强的产业。另一方面，将密集度要素视作评估风险的一个标准。一旦某产业出现密集交易的信号，便表示该产业可能面临短期的过度交易和风险提升，此时我们应迅速卖出股票退出市场。

3.2 创新点及意义

本策略通过重构量化行业轮动框架，关注市场的拥挤度效应，深度挖掘拥挤度因子，参考动量效应，认为行业的上涨趋势具有惯性，短期内可以通过行业截面比较寻找涨幅靠前的行业进行配置。其次，我们把拥挤度因子作为风险效应的指标。当行业触发拥挤度信号后，提示行业短期内交易过热，风险上升，这时应立即清仓离场。

策略主要利用拥挤度效应对市场情绪的精确把握，从流动性、乖离率等角度来筛选拥挤度因子，不仅把握了市场“速度”，还抓住了市场“加速度”，更好地反映了市场的拥挤度效应。基于动量理念构建的产业投资组合平台，进而利用拥挤度指标法来筛查并排除那些并未经历拥挤现象的行业，进而淘汰那些风险较高的行业。观察发现，复合策略可以较完美地避开即将大幅回撤的行业，保留未来涨幅排名靠前的行业。

该策略的创新点主要集中在动态的行业与个股选择上，利用拥挤度综合指标判断、选择优质行业，运用动态股票池构建，最大程度地抵御市场风险；通过严谨的信号有效性检验，修正叠加信号的误判，及时止盈止损，达到较低回撤率基础上的效益最大化。同时，基于整个量化策略设计投资组合优化，采用最佳投资比例进行组合，从而有效提高收益、减少风险。

总体而言，结合行业轮动的拥挤度选股策略的亮点更多在于防御市场大幅下行和提前预判反转或新趋势的到来；遇到极端风险时，处于估值洼地的价值型股票拥有更强的支撑力，能更有效地控制回撤。在定位上，此复合策略对盈利水平的提升并不突出，但其抗跌能力有所增强，尤其是对将到来的重大风险的感知能力。在近年市场波动较大，不确定性仍在延续的环境下，这类防御性策略更能发挥其配置价值。

4 理论分析

本策略旨在通过研究短期内某一产业是否遭受了过多交易行为，从而躲避那些已经人满为患的产业，以此达到降低交易风险的目的。文中为了方便多个指标之间进行比较，将会直接使用历史分位数来量化产业的密集程度，密集程度的指标读数愈高，表明该产业面临的拥挤问题愈加重。

4.1 拥挤度代理变量

(1) 集中度

行业 i 在某一时刻的“集中度”定义如下：

$$C_i = \frac{\sum_{j=1}^n (AR^j * |w_i^j|)}{\sum_{j=1}^n AR^j} \quad (1)$$

$$AR^j = \frac{\sigma_{Ej}^2}{\sum_{i=1}^N \sigma_{Ai}^2}$$

其中， σ_{Ej}^2 是第 j 个特征向量的方差，也就是第 j 个特征值； σ_{Ai}^2 是第 i 个资产的方差； AR^j 是第 j 个特征向量的“吸收度”； w_i^j 内 j 是资产 i 在第 j 个特征向量中的权重。直观的来说， AR^j 测量了特征向量对样本总方差的解释力度，而 C_i 测量了资产 j 对于总样本方差的影响力。

(2) 乖离率指标

乖离率指标（BIAS）：显示了在特定时间段内市场价格相对其移动平均线（MA）的脱节水平，用以判断价格因远离其平均走势而引起的剧烈波动，进而预测其调整或跳跃的可能性与信度。乖离率通常是指最新市场价格与近期平均值之间的偏离度，以 t 时刻 K 行业的收盘价乖离率（窗口期 N）为例：

$$\text{close_bias}_{t,k} = \frac{\text{close}_{t,k} - \sum_{i=0}^{N-1} \text{close}_{t-i,k} / N}{\sum_{i=0}^{N-1} \text{close}_{t-i,k} / N} \quad (2)$$

本策略改进了乖离率，通过计算成交量、成交额和换手率的乖离率，通过与 N 日平均值分析比较，及时反馈市场最新的流动性指标出现了明显的偏离。正向偏差加剧时，短线盈利及利润回撤的可能性增加；反向偏差加大时，做空者回补的机遇亦随之提高。

$$\begin{aligned} \text{volume_bias}_{t,k} &= \frac{\text{volume}_{t,k} - \sum_{i=0}^{N-1} \text{volume}_{t-i,k} / N}{\sum_{i=0}^{N-1} \text{volume}_{t-i,k} / N} \\ \text{amt_bias}_{t,k} &= \frac{\text{amt}_{t,k} - \sum_{i=0}^{N-1} \text{amt}_{t-i,k} / N}{\sum_{i=0}^{N-1} \text{amt}_{t-i,k} / N} \\ \text{turn_bias}_{t,k} &= \frac{\text{turn}_{t,k} - \sum_{i=0}^{N-1} \text{turn}_{t-i,k} / N}{\sum_{i=0}^{N-1} \text{turn}_{t-i,k} / N} \end{aligned} \quad (3)$$

（3）量价相关系数

通常情况下，股票的流动性和价格走势变化趋势相同，量价相关系数可以用于判断是否存在量价背离，价格趋势是否反转。我们采用行业指数收盘价和流动性指标的相关系数来定量分析量价背离现象。量价相关系数越大，说明量价指标的走势越接近，反之则表示两者出现背离。行业 K 在时刻 t 的成交量与收盘价相关系数可以表示为：

$$\begin{aligned} \text{corr_volume_close}_{t,k} &= \frac{\sum_{i=0}^{N-1} (\text{close}_{t-i,k} - \overline{\text{close}_{t,k}})(\text{volume}_{t-i,k} - \overline{\text{volume}_{t,k}})}{\sum_{i=0}^{N-1} (\text{close}_{t-i,k} - \overline{\text{close}_{t,k}})^2 \sum_{i=0}^{N-1} (\text{volume}_{t-i,k} - \overline{\text{volume}_{t,k}})^2} \\ \overline{\text{close}_{t,k}} &= \sum_{i=0}^{N-1} \text{close}_{t-i,k} / N \\ \overline{\text{volume}_{t,k}} &= \sum_{i=0}^{N-1} \text{volume}_{t-i,k} / N \end{aligned} \quad (4)$$

当股价上涨时，如果某个行业指数的量价相关系数下降，意味着出现了量价背离的现象。当量价相关系数处于历史低位时，需要警惕可能出现的下跌风险。可通过收盘价和成交额以及换手率的相关系数对量价背离现象进行分析，成交额（turnover）与之同理，公式如下：

$$\begin{aligned} \text{corr_amt_close}_{t,k} &= \frac{\sum_{i=0}^{N-1} (\text{close}_{t-i,k} - \overline{\text{close}_{t,k}})(\text{amt}_{t-i,k} - \overline{\text{amt}_{t,k}})}{\sum_{i=0}^{N-1} (\text{close}_{t-i,k} - \overline{\text{close}_{t,k}})^2 \sum_{i=0}^{N-1} (\text{amt}_{t-i,k} - \overline{\text{amt}_{t,k}})^2} \\ \text{corr_turn_close}_{t,k} &= \frac{\sum_{i=0}^{N-1} (\text{close}_{t-i,k} - \overline{\text{close}_{t,k}})(\text{turn}_{t-i,k} - \overline{\text{turn}_{t,k}})}{\sum_{i=0}^{N-1} (\text{close}_{t-i,k} - \overline{\text{close}_{t,k}})^2 \sum_{i=0}^{N-1} (\text{turn}_{t-i,k} - \overline{\text{turn}_{t,k}})^2} \end{aligned} \quad (5)$$

（4）收益率峰度

通常状况下行业内成分股收益率往往存在分化，如果个别时段成分股无差别上涨、分化度下降，可能意味着资金入场时没有进行个股筛选、市场热度较高。我们采用标准差和峰度两个指标衡量个股收益率分化程度。峰度公式为：

$$K = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} \quad (6)$$

式中, K 表示峰度 (无量纲); i 表示第 i 个数值; \bar{x} 表示平均值; n 表示采样数量。

5 实证设计

5.1 拥挤度指标构建

行业拥挤是一种交易现象, 当投资者情绪过于高涨, 资金集中进入某个行业时, 就形成了行业拥挤。行业拥挤度过高后, 会带来绝对收益或者相对收益的下降。

行业的长期走势由基本面决定, 但行业拥挤是一种交易现象, 因此可以使用量价交易数据来衡量拥挤度。衍生技术指标种类繁多, 通过对指标的不断尝试, 确实可以取得较好的拥挤度识别效果, 但是计算量也会非常巨大, 这将增加工作量, 而且技术指标都会涉及到参数, 指标越多参数也会越多, 这会大大增加过拟合可能。

我们对拥挤度单因子进行初步测试。本策略统一采用历史分位数方法判断每项量价指标在历史上所处位置, 基于行业轮动进行有效性回测, 在时序的框架下, 计算一个原始的量价指标, 判断该指标在历史上是否处于临界位置。

首先, 确定各指标的窗口期长度 N , 计算各个量价指标的原始数值, 各指标窗口期取值范围设定如下:

对于各指标测试窗口长度设定, 经过数据推演, 为达最佳效果, 我们统一采用 60 日作为标准回测时长, 基于各指标确定触发拥挤状态临界阈值。

根据不同指标的方向性, 历史分位数的阈值设定如下:

(a) 正向指标: 当其量价指标历史分位数低于规定阈值时, 判定当前属于拥挤状态, 设定为 1%、5%、10% 和 20%。

(b) 负向指标: 指标历史分位数需高出规定阈值时才能判定处于拥挤状态, 设定指标分位数阈值为 80%、90%、95% 和 99%。

数据预处理: 历史分位数的计算统一起始于 2012.1.1-2022.1.1, 评估在过去 20 个交易日行业指数收益率大于零时拥挤度指标的有效性, 所有拥挤度指标对应一个时间窗口参数, 在测试中同时关注拥挤度指标的方向与正负相关系数。

在代理变量构建过程中, 经常面临的问题是行业间同一维度数据的量级存在天然差异, 无法直接比较, 同时指标时序上的绝对值高低亦受市场大环境影响, 为了减缓以上的影响, 我们首先需要对指标进行去量纲处理。

5.2 拥挤度指标有效性检验

拥挤度指标的有效性体现在能否及时对市场交易过热风险做出提示, 避免投资组合出现大幅亏损, 指标有效性主要体现在两个方面:

(1) 拥挤度指标需要与行业指数未来一段时间收益率存在负相关关系, 也就是拥挤度指标越高时指数下跌概率也应该随之增大;

(2) 拥挤度指标需要存在合理的门槛阈值设定, 当拥挤度指标处于阈值之上时, 行业指数未来一段时间收益率的期望为负值, 如果找不到合适的阈值, 我们就难以通过拥挤度指标生成止损信号。

我们以显著性检验、信号胜率检验、回测收益检验、过拟合检验为基础构建门限回归检验模型。

指标检验的回测区间设定为 2012.01.01 日至 2022.01.01, 实验分为样本检验与样本外检验。门限回归模型对于这种数据变化现象可以进行有效识别和预测, 其基本思想是通过门限值的控制作用构建回归方程: 当给出回归方程的变量信息后, 可以首先根据门限值对数据进行筛选, 在不同门限值设定下使用合适的回归方程。门限回归方法本质上是将样本数据按状态空间的取值进行分类, 采用分段的线性回归模型表示数据的相关性。

由于门限回归采用了分段线性化的方法，因此可以充分利用线性回归模型的处理方法。一般门限回归的模型形式为：

$$y = \begin{cases} \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1m}x_m & x_{i0} < x_i < x_{i1} \\ \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \dots + \beta_{2m}x_m & x_{i1} < x_i < x_{i2} \\ \dots \dots \dots \\ \beta_{k0} + \beta_{k1}x_1 + \beta_{k2}x_2 + \dots + \beta_{km}x_m & x_{i(k-1)} < x_i < x_{ik} \end{cases} \quad (7)$$

其中 x_i 为自变量， x_{ij} 为不同的门限值， β_{ij} 为不同门限值下的回归系数。在不同门限值设定下的回归方程中，回归系数可以存在很大差异。

拥挤度指标的有效性检验：

(1) 回归系数随门限值增大呈变小趋势。随着门限值增大，门限回归方程的系数也将不断发生变化。门限回归计算因变量选取的是换手率、峰度以及成交额与收盘价的相关系数，一个有效的拥挤度指标，随着门限值不断变大系数应不断变小。

(2) 为了保证门限回归结果的有效性，在设定的门限值取值范围内（50%-70%），回归系数需要显著为负。

(3) 随门槛值提高，超过该门槛的指数回报率中正值的所占比重预期将逐步减少，两者之间显示出明显的相反相关性。高于此门槛的指数回报率的中位数理应逐步向负值靠拢，并显示出负向的相关性；若中位数为负 0.5 或更小，则可视作判断拥挤度指标是否有效的标准。

在完成了涵盖显著性评估、相关系数的回归检验、成功率与盈利能力的相关性测定等四重验证后，精选了三项堵塞度的参照指数：10 日收益峰度、近 20 日均值转手率、以及 40 日的交易额与收盘价的相关系数。这三项指标通过了 50%至 70%这一显著性门槛的检验，对它们分别进行了显著性分析。基于这三种单独指标之和，构建了一个综合的堵塞度指标后，又对其进行了显著性的检验。

Sequential F-statistic determined thresholds: 0 Significant F-statistic largest thresholds: 0				Sequential F-statistic determined thresholds: 0 Significant F-statistic largest thresholds: 2			
Threshold Test	F-statistic	Scaled F-statistic	Critical Value**	Threshold Test	F-statistic	Scaled F-statistic	Critical Value**
0 vs. 1	1.743510	8.717552	18.23	0 vs. 1	2.672438	13.36219	18.23
1 vs. 2	1.462260	7.311298	19.91	1 vs. 2 *	4.654062	23.27031	19.91
2 vs. 3	1.866763	9.333817	20.99	2 vs. 3	2.068341	10.34170	20.99
3 vs. 4	0.604497	3.022487	21.71	3 vs. 4	1.306089	6.530443	21.71
4 vs. 5	0.096406	0.482028	22.37	4 vs. 5	0.000000	0.000000	22.37
Sequential F-statistic determined thresholds: 0 Significant F-statistic largest thresholds: 0				Sequential F-statistic determined thresholds: 1 Significant F-statistic largest thresholds: 1			
Threshold Test	F-statistic	Scaled F-statistic	Critical Value**	Threshold Test	F-statistic	Scaled F-statistic	Critical Value**
0 vs. 1	2.889101	14.44551	18.23	0 vs. 1 *	4.631850	23.15925	18.23
1 vs. 2	1.020636	5.103179	19.91	1 vs. 2	3.659793	18.29897	19.91
2 vs. 3	1.571237	7.856184	20.99	2 vs. 3	1.436755	7.183775	20.99
3 vs. 4	0.567091	2.835457	21.71	3 vs. 4	0.774382	3.871908	21.71
4 vs. 5	0.567307	2.836535	22.37	4 vs. 5	0.784001	3.920004	22.37

图 5 检验结果

可以看出当门槛在 50%-70%之间时，F 值都较为显著，通过检验，模型科学。

5.3 最终指标构成

基于上节门限回归与遍历计算，最终经过拟合优化的 3 个有效衡量指标。

筛选指标 1：成分股 10 日收益率峰度。

comp_turn_kurtosis_10 指标表现非常稳定，在不同门限值设定下回归系数均显著为负。考虑到指标本身的单调性非常明显，我们采用 95% 阈值进行市场拥挤与否的判别。

筛选指标 2：过去 20 日平均换手率。

turn_20 指标在门限值较大时显著性会下降，门限值取 95% 时显著性在 0.1 以上，说明 turn_20 指标不宜设置过高的阈值，我们采用 90% 作为 turn_20 指标的判定阈值。

筛选指标 3：40 日成交额收盘价相关系数。

corr_amount_close_40 也非常稳定，直接取最高的 95% 作为判定阈值。

筛选出的 comp_turn_kurtosis_10、turn_20、corr_amount_close_40 三个拥挤度指标都具有提示市场下行风险的作用，当每个指标处于阈值之上时，行业指数下跌概率较高。最终，经过拟合优化处理，对行业拥挤度判定的三个门限条件分别以 95%、90%、95% 的阈值。

6 实证检验与回测

基于以上指标构建的拥挤度指标作为行业拥挤度与个股拥挤度判断标准，在进行了指标构建后瞄准不同综合指数进行检验。

针对沪深 300、中证 1000、小市值三类标的进行实证检验，得出结论：在拥挤度指标在区域极点时，往往指数也接近阶段高点（低点），当拥挤度指标反转后（斜率正负发生变化），指数也开始反转。一般来说，低拥挤度往往按存在买入机会或反转的多头信号；高拥挤度往往是存在泡沫、空方强势的信号，这一现象在市值较小的标的组成的指数中更为明显。当然，该指标在部分时刻存在一定滞后性，在未来需要进一步研究与优化。

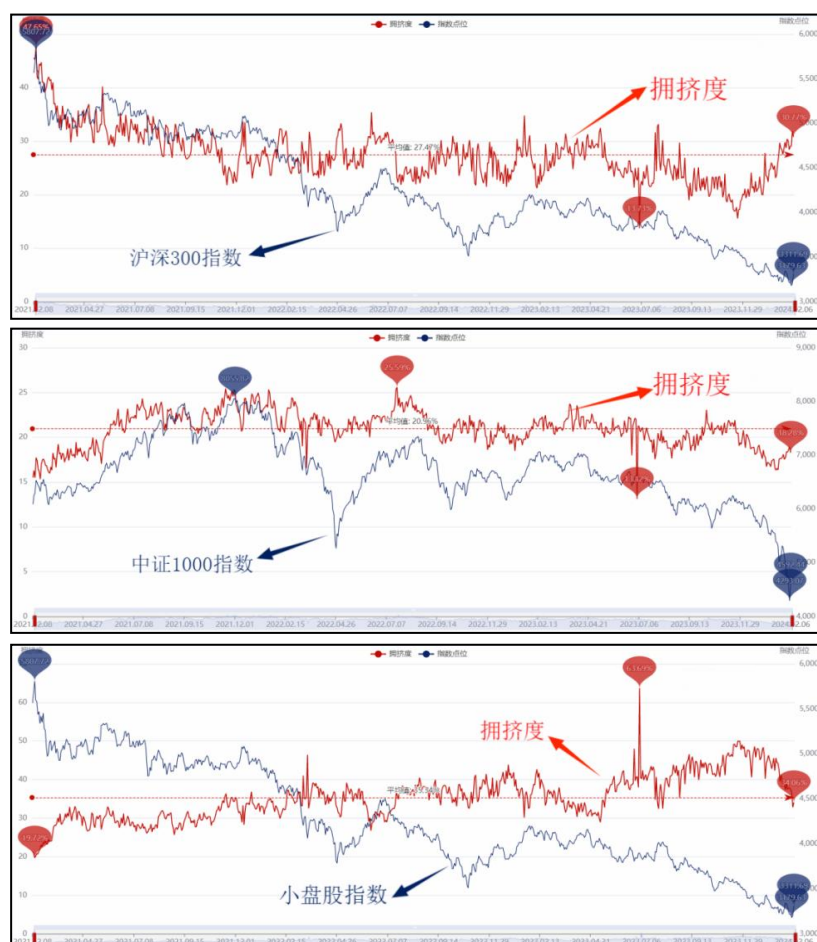


图 6 拥挤度指标作用效果（沪深 300、中证 1000 与小盘股）

由于本文以论述拥挤度指标为主，而相关量化策略的实证需要“选股-择时-风控”这一系列模块的综合运用，故在此简要叙述选股、择时、风控逻辑。

经过因子筛选，根据月频拥挤度行业轮动策略，设置两组选股条件进行大盘拥挤信号的处理，最终在大盘拥挤时选用 A，非拥挤时选用 B。

A 组：毛利率大于 15%；每股净资产相对增长率为正；市盈率在 10-45 之间；总市值介于 30-2000 亿；

B 组：净资产收益率连续三年为正；商誉净资产比低于 100%；年净利润同比增长 10%；总市值介于 30-2000 亿。

经过最终复合因子筛选的股票才能作为最终交易标的，下图是择时买卖点逻辑。

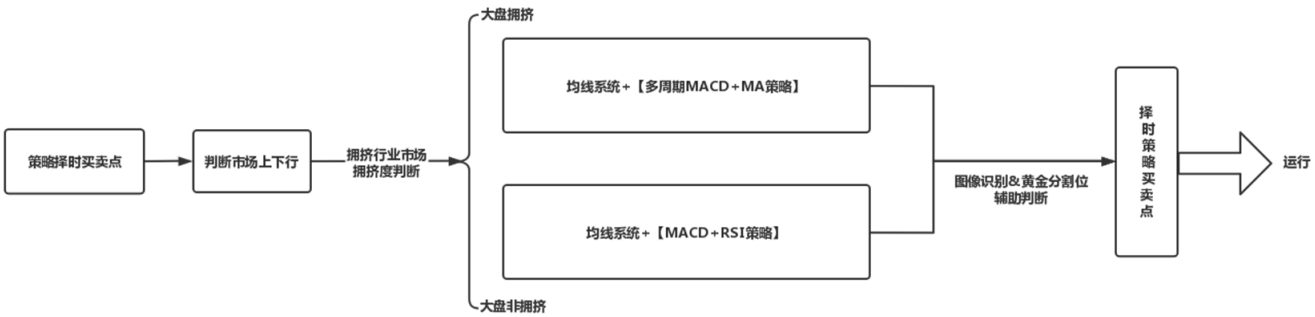
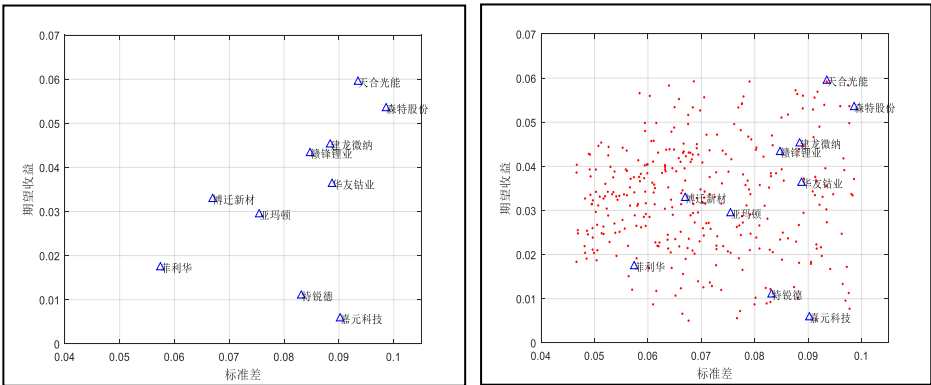


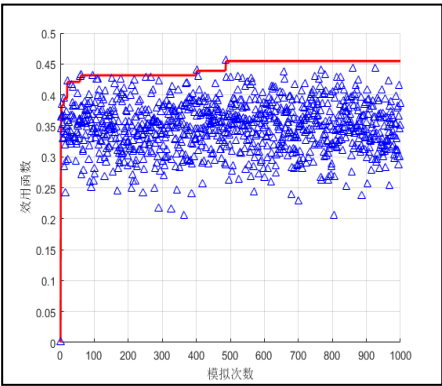
图 7 买卖点选择逻辑

在风控方面，策略构建了最优投资组合，以下是部分测试数据，策略整体满足效用最大化。基于拥挤度指标的量化投资策略风险可控，模拟得出的最佳投资组合和优化计算的结果基本一致。



图(a) 组合前的收益与风险

图(b) 组合后的收益与风险



图(c) 效用函数

图 8 通过模拟寻找最优投资组合

综合上述研究，将本策略通过同花顺 SuperMind 平台进行回测，回测时间设置为：2022 年 2 月 1 日至 2024 年 2 月 1 日；初始资金设置为：10000000 元，选股周期为每日，其中红线代表的是本策略的收益率，灰线代表了同时期沪深 300 指数的收益率，可以看出在市场行情不佳的情况下，该策略具有较高的风险规避能力、稳健的综合收益率与较低的回撤区间，因而本策略具有一定的投资价值，战胜市场平均收益、规避了大部分投资风险。

策略回测资金曲线如下：



7 结论与建议

本文主要研究了基于行业轮动的集中度复合因子量化策略，其中，资产集中度类指标定位经常导致泡沫形成的拥挤交易，而拥挤度复合因子与相对估值用来确定泡沫何时破裂。就拥挤度中任意单指标本身不足以确定泡沫何时破裂，但同时使用这些指标可以识别行业和因子泡沫的上升和破裂。通过采用基于这些指标的行业轮动策略，投资者的表现可能会显著的优于中国股市甚至全球股市的平均表现。

投资者可以通过使用这些指标来进行因子择时，并构建一系列量化投资策略，其结果显著优于 A 股和静态因子投资组合。

建立拥挤指数是为了捕捉市场中的过度交易状况。直观上看，股市最显著的拥挤特征就是交易量和价格的大幅提升。此外，市场波动率增大、量价变化趋势不一致等异常现象，也显示目前市场风险正在累积，可能存在拥挤交易。本文从以下几个角度建立拥挤指数，对市场中的交易风险提出警示：

- (1) 衡量行业指数的价格或流动性是否大幅上涨，相应指标：动量、流动性和乖离率指标；
- (2) 衡量行业指数的交易量和价格走势是否一致，相应指标：量价相关性指标；
- (3) 衡量当前市场是否存在大幅波动，相应指标：波动率指标；
- (4) 衡量近期行业指数的收益率分布与统计规律是否偏离。

关于如何去刻画行业以及因子拥挤程度，学术界和业界都有很多研究和尝试。常见的拥挤度的刻画方式有估值差、融券利用率、配对相关性、因子波动率和因子长期收益反转等。可上述指标在实际应用过程中却或多或少存在一些问题。这一策略的常见问题就是，行业或者因子的拥挤度较高的时候往往并不是行业或者因子表现见顶的时候，即高拥挤度往往并不一定对应着未来的低收益。

本文所提出的行业轮动与拥挤度复合因子结合的方法让人耳目一新。从实证结果来看，能够较好把握泡沫初期上涨带来的收益，并较好地规避泡沫卖出期的下跌损失。后续我们将继续在 A 股市上针对行业拥挤度和复合因子进行相应的实证研究，希望能收获更多实用结论！

本文提供了一种可供投资者参考的规避市场风险、获取超额收益的方法。在纷繁复杂的当代金融市场，拥有稳健收益能力和风险规避能力的多元策略投资者显然更容易应对市场的波动与风险，从而在市场交易中立足。

致谢

感谢刘家鹏教授在本文研究方面提供的指导与帮助。

参考文献

- [1] 吴文超. 中国 A 股市场行业轮动现象分析--基于动量和反转交易策略的检验[J]. 金融理论与实践, 2014,000 (009) :111-114.
- [2] 吴宇阳. 股票市场因子拥挤度研究与实证分析[D]. 东北财经大学, 2020,015-020.
- [3] 刘睿. 基于因子拥挤度的组合配置策略研究[D]. 上海财经大学, 2021,010-021.
- [4] 石川等. 投资因子方法与实践[M]. 北京:电子工业出版社, 2020,001-360.
- [5] 陈强. 机器学习与 python 应用[M]. 北京:高等教育出版社, 2020,001-360:高等教育出版社, 2020,377-456.
- [6] 梁勇. 收敛率 (BIAS) 指标在报表分析中的应用研究[J]. 中国总会计师.
- [7] 林晓明, 陈晔. 华泰多因子系列之二: 华泰单因子的估值因子检验[R]. 华泰证券, 2016.
- [8] 林晓明, 李聪, 刘志成, 王家兴. 拥挤指数在行业配置中的应用[R]. 华泰证券, 2020.
- [9] 林志鹏, 刘福兵. 因子时点、因子动量、因子离散度、因子拥挤度的三个尺度[R]. 国盛证券, 2019.
- [10] 尹明, 刘福兵. 多因子系列之一:对多因子选股体系的思考[R]. 国盛证券, 2019.
- [11] 于明明, 徐寅. 系统化资产配置系列之八. 基于因子的资产配置研究[R]. 兴业证券, 2020.
- [12] Chue,T., "Omitted Risks or Crowded Strategies: Why Mutual Fund Comovement Predicts Future Performance.", FMA working paper, 2015.
- [13] Daniel,K., Moskowitz,T., "Momentum crashes."Journal of Financial Economics,2016,122,221-247.
- [14] Pojarliev, M., Levich,R.M., "Detecting crowded trades in currency funds.",Financial Analysts Journal,2011,67(1),26-39.
- [15] Zhong,L., Ding.X.,Tay.N., "The impact on stock returns of crowding by mutual funds." The Journal of Portfolio Management,2017,43,87-99.