

2021 年 12 月 27 日

## 行业泡沫膨胀与破裂的识别：以拥挤之名

金融工程研究团队

——开源量化评论（45）

魏建榕（首席分析师）

证书编号：S0790519120001

张翔（分析师）

证书编号：S0790520110001

傅开波（分析师）

证书编号：S0790520090003

高鹏（分析师）

证书编号：S0790520090002

苏俊豪（研究员）

证书编号：S0790120020012

胡亮勇（研究员）

证书编号：S0790120030040

王志豪（研究员）

证书编号：S0790120070080

盛少成（研究员）

证书编号：S0790121070009

苏良（研究员）

证书编号：S0790121070008

魏建榕（分析师）

weijianrong@kysec.cn

证书编号：S0790519120001

胡亮勇（联系人）

huliangyong@kysec.cn

证书编号：S0790120030040

### ● 单个指标捕捉行业拥挤情况存在不达预期和参数敏感的问题

一个好的行业轮动模型至少应该包含三个维度的信息，跟踪行业价格走势变化的动量，代表行业基本面强弱的景气度，和体现投资者交易热度的拥挤度。

单一指标在识别拥挤度方面存在不足，通常存在效果不达预期且参数敏感的问题。比如，换手率具有较优的识别泡沫破裂的能力；配对相关性由于本身只能识别泡沫是否存在，无法预测泡沫破裂时间节点，在表现方面稍逊一筹；下行波动率、乖离率和估值水平测试效果不达预期。

### ● 以资产集中度和相对价值构建的泡沫识别模型不适用于 A 股市场

资产集中度衡量了行业的波动性和关联度，其能够识别泡沫是否存在，但无法预测泡沫何时破裂。

相对价值的走势能够较好拟合泡沫破裂前后的形态特征，即泡沫膨胀期相对价值迅速提升，泡沫破裂期则快速下降。相对价值在识别出泡沫的基础上，可以用来判断泡沫从膨胀到破裂的拐点。

根据参考文献观点，理论上高资产集中度和低相对价值的泡沫膨胀组表现最好，高资产集中度和高相对价值的泡沫破裂组表现最差。但 A 股的测试结果表明，泡沫破裂组反而表现最好，与预期不符。

### ● 配对相关性和换手率或许是识别泡沫拐点更好的指标

配对相关性能够有效地衡量行业内个股的同涨同跌现象。且在行业股价处于相对高位时，无论股价上涨或下跌，配对相关性均能够稳定地保持在高位，适合用来判断泡沫是否存在。

换手率能够有效跟踪泡沫破裂前后的形态变化，在泡沫膨胀期，由于赚钱效应，换手率提升；在泡沫破裂后，由于处置效应，换手率快速衰减。

基于配对相关性和换手率的象限切分，能够有效复现参考文献中的观点。高配对相关性和低换手率的泡沫膨胀期跑赢其他象限，测试期内年化收益率 9.3%，最大回撤 42.3%；高配对相关性和高换手率的泡沫破裂期大幅跑输其他象限，测试期内年化收益率-3.5%，最大回撤 64.4%；同期以行业等权构建的参照组年化收益率 7.1%，最大回撤 51.5%。

以“拥挤”之名构建的泡沫识别模型对既有行业轮动信号有一定改善作用。以开源金融工程行业轮动模型为例，在每期选出的 9 个行业中，进行拥挤行业剔除，测试期内年化收益率提升 2%，剔除胜率约 66%。

### ● 风险提示：模型测试基于历史数据，市场未来可能发生变化。

## 相关研究报告

《开源量化评论（41）-可转债估值择时策略初探》-2021.10.21

《开源量化评论（42）-知情交易者背后的择时信息》-2021.10.24

《开源量化评论（43）-量化私募行为的蛛丝马迹：龙虎榜营业部新视角》-2021.10.31

《开源量化评论（44）-北向关联持仓中的 Alpha》-2021.11.07

## 目 录

1、 拥挤度代理变量初筛 .....	4
1.1、 换手率是市场交易热度的直观体现.....	4
1.2、 下行波动率可用来表征市场波动 .....	5
1.3、 配对相关性代表了行业内个股的联动性.....	6
1.4、 乖离率刻画了当前股价偏离均值的幅度.....	6
1.5、 相对估值衡量了市场的估值水平 .....	6
2、 数据预处理 .....	7
3、 拥挤度代理变量测试 .....	8
3.1、 换手率：高换手率是市场情绪高涨的有效映射.....	8
3.2、 下行波动率：高下行波动率并不是刻画拥挤程度的良好指标.....	9
3.3、 配对相关性：配对相关性能有效衡量行业内个股的同涨同跌现象.....	9
3.4、 成交量乖离率：高成交量乖离率或许是新增资金入场的信号 .....	10
3.5、 相对估值：高估值并不是行业未来下跌的充分理由 .....	10
4、 拥挤度代理变量融合 .....	11
4.1、 资产集中度可识别泡沫存在与否 .....	11
4.2、 相对价值能够判别泡沫从膨胀到破裂的拐点.....	12
4.3、 象限切分后不同区域代表行情的不同阶段.....	13
4.4、 代理变量的选择 .....	15
5、 拥挤度的实践 .....	16
5.1、 基于拥挤度的行业剔除 .....	16
5.2、 纳入拥挤度的行业轮动 .....	17
6、 风险提示 .....	18

## 图表目录

图 1： 高换手率通常是潜在回撤的先行指标.....	5
图 2： 下行波动率与波动率走势整体接近，部分时候存在差异.....	6
图 3： 金融、周期板块包含的行业具有更低的换手率 .....	7
图 4： 指数加权后，两种标准化方法的走势接近.....	8
图 5： 高相对换手率组在 2016 年后逐渐表现出负超额收益.....	8
图 6： 高下行波动率组并不是拥挤度的良好代理变量.....	9
图 7： 高配对相关性组有一定的反向预测能力，但波动较大.....	10
图 8： 成交量乖离率并不适合用来刻画行业拥挤状态.....	10
图 9： 高相对估值组稳定跑赢基准，与预期相反.....	11
图 10： A 股不同行业资产集中度变动更加频繁.....	12
图 11： 相对价值指标能够较好地反应市场上各行业的冷热状态变化 .....	13
图 12： 资产集中度和相对价值无法有效地捕获 A 股泡沫.....	14
图 13： 高分组纳入的行业数量越少，未来超额收益越高.....	14
图 14： 配对相关性可用来识别泡沫是否存在.....	15
图 15： 换手率可用来识别泡沫何时破裂 .....	15
图 16： 配对相关性+换手率能够较好地识别行业泡沫膨胀和破裂阶段.....	16
图 17： 基于配对相关性和换手率的每月拥挤行业个数中位数为 3 个 .....	16
图 18： 不含拥挤行业的轮动模型表现年化收益率提升约 2% .....	17

图 19: 高动量+低拥挤组表现优异，超额稳定性较好 ..... 18

表 1: 泡沫破裂组 C 显著跑输等权参照组 ..... 16

表 2: 高动量+低拥挤组合能有效改善策略表现 ..... 18

一个好的行业轮动模型至少应该包含三个维度的信息，跟踪行业价格走势变化的动量，代表行业基本面强弱的景气度，和体现投资者交易热度的拥挤度。价格趋势向上的行业能够给投资者带来直观的收益；行业景气度向上是推动行业趋势形成的底层逻辑；而交易是否拥挤则预示着未来是否会产生较大的回撤。根据以上三点，一个理想的行业应该是趋势良好，基本面强劲，但市场参与热度并没有达到狂热状态。

索罗斯曾经说过，“世界经济史是一部基于假象和谎言的连续剧，要获得财富，做法就是认清其假象，投入其中，然后在假象被公众识别之前退出游戏”。泡沫膨胀与破裂的过程与此有异曲同工之妙，行业泡沫膨胀阶段是收益率增长最迅速的区间，但是泡沫破灭后亦是回撤最大的过程。如何识别泡沫膨胀并参与其中，并在泡沫破裂前及时抽身退出，是我们本期想要解决的问题，或许答案便隐藏在“拥挤”表象之下。

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。接着我们参考了学术文献《Crowded Trades: Implications for Sector Rotation and Factor Timing》中的思路，选取两类指标进行象限切分，对泡沫膨胀和泡沫破裂阶段进行定义，并测试其在A股市场的适用性。然后我们在此基础上针对A股市场特征对泡沫识别指标进行了相应修改，发现改进后的融合指标能够更好地规避泡沫破裂带来的损失。最后，我们在此基础上，测试了纳入拥挤度考核对既有行业轮动模型信号的改善效果。同时根据测试结果，在参考象限切分的思路，提出了拥挤度指标的其他应用方案。

## 1、拥挤度代理变量初筛

拥挤交易是指资金大量购买或出售同一项资产或一组具有相似特征的资产，导致资产价格发生重大变化。当行业产生拥挤交易时，资金争相涌入该行业并不断推高行业内股票的价格，行业内个股的关联度亦逐渐提升，形成同涨同跌现象。

除了这种定性的描述，我们希望通过定量指标来捕获行业的拥挤程度。本篇报告根据经验和惯例，选取了换手率、下行波动率、配对相关性、乖离率和估值水平等量价指标，试图定量刻画当下的行业拥挤情况。

### 1.1、换手率是市场交易热度的直观体现

换手率作为衡量市场交易活跃度的直观指标，能够比较有效的判断市场情绪的拐点。若某个行业换手处于持续上升的通道，通常意味着行业的交易热度逐渐加强，市场的关注度在升温。在资产价格达到阶段高点时，换手率新高通常对应着市场情绪亢奋，但分歧也开始加大，未来发生回撤的风险增加。在后视镜视角下，以2021年市场关注度较高的周期板块采掘（现煤炭）行业为例，在2021年9月16日，采掘行业指数达到阶段高点，同期换手率亦达到阶段高点，后续股价开始了一轮大幅回撤。但换手率达到阶段高点时，并不必然会立即产生较大的回撤。比如2021年8月20日，采掘行业换手率便处于上行趋势且在阶段高点，但采掘指数依然在此之后继续上涨。

图1：高换手率通常是潜在回撤的先行指标



资料来源：Wind、开源证券研究所

计算行业换手率时，我们通过对行业内成分股换手率进行市值加权计算，其中权重为个股的流通市值。

$$ind_k = \frac{\sum_i^n w_i tr_i}{\sum_i^n w_i}$$

其中， $w_i$ 表示行业内个股 $i$ 的流通市值权重占比， $tr_i$ 表示个股 $i$ 的换手率， $n$ 表示行业内个股的数量。

## 1.2、下行波动率可用来表征市场波动

波动率是市场参与者分歧程度的映射。当市场对行业未来表现的分歧逐渐增加时，股价的波动会相应增加。当行业处于波动高位时，是否对应着行业表现开始出现反转呢？

相比于波动率，下行波动率能够更好地反应极端下跌风险。根据波动率的定义，下行波动率可以通过对小于目标收益率水平的数值进行计算得到，具体计算公式如下：

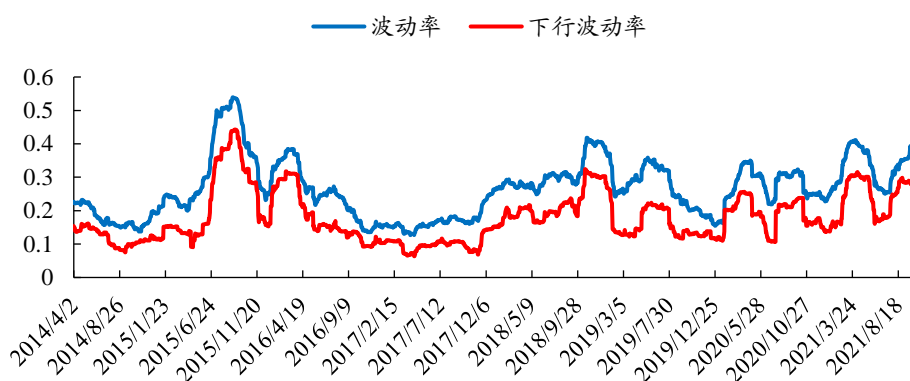
$$\sigma_k^d = \sqrt{\frac{1}{N} \sum_{r_i < \bar{r}} (r_i - \bar{r})^2}$$

其中， $\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i$ ，表示过去一段时间的平均收益率， $N$ 表示低于均值的样本数量。

在实际处理过程中，我们将收益率低于 0 值的交易日即判定为下行风险期。以食品饮料行业为例，其波动率和下行波动率的走势如图 2 所示。可以看到波动率和下行波动率走势在大部分时间基本一致，仅在特定时刻有所差异。



图2：下行波动率与波动率走势整体接近，部分时候存在差异



数据来源：Wind、开源证券研究所

### 1.3、配对相关性代表了行业内个股的联动性

配对相关性是指行业内个股之间两两相关性的均值或中位数。一般而言，当某个行业成为市场热门时，此时市场中活跃的资金会较行情平淡时更关注行业的 beta 属性，而相应淡化行业内个股的 alpha 特征，体现到资金操作上便是行业内个股同涨同跌的现象凸显。资金的快速大幅流入，推升行业内个股配对相关性逐步提升时，行业也开始变得拥挤。

行业配对相关性的定义为：

$$corr_k = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=1, i \neq j}^n corr(r_i, r_j)$$

其中， $r_i$ 和 $r_j$ 分别是股票 $i$ 和 $j$ 的超额收益序列，基准为个股所在行业收益率。

### 1.4、乖离率刻画了当前股价偏离均值的幅度

价格乖离率代表标的资产当日收盘价与其移动均线的差距，其刻画了股价偏离市场平均成本线的幅度。乖离度越高，表明市场获利盘越大，未来投资者兑现收益的需求升高，股价承压增加，股价发生下跌的风险上升。以价格乖离率为例，其定义如下：

$$Bias_k^n = \frac{Close - MA_n}{MA_n}$$

其中， $Bias_k^n$ 表示行业 $k$ 的 $n$ 日乖离率， $Close$ 表示当日收盘价， $MA_n$ 表示 $n$ 日移动平均值。

### 1.5、相对估值衡量了市场的估值水平

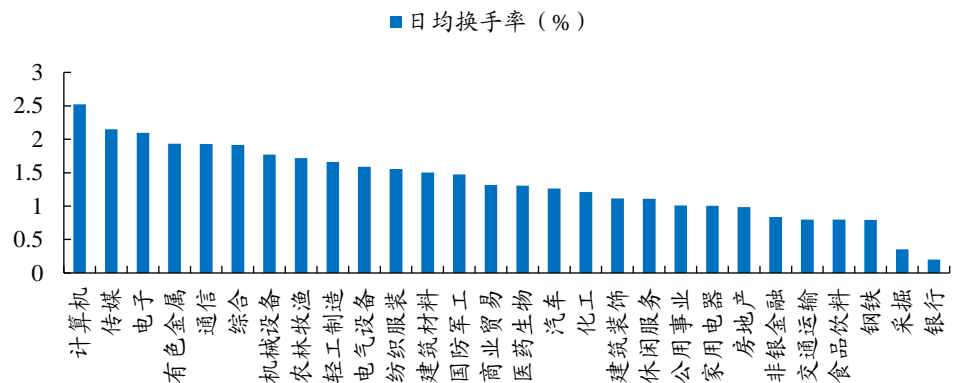
谚语有云：“树不会长到天上去”。任何行业都有估值的天花板，当估值达到极端状态时，若盈利增长不足以消化估值压力，后续高估值行业将面临回调压力。不同行业之间估值也存在着量级差异，高成长板块市场通常愿意给予更高的估值水平，而对于成长性相对较低的板块，市场通常给出的估值较低。但无论是高成长还是低成长行业，行业估值均有一定的上下限，当超过历史上限时，通常认为行业过于拥挤。

## 2、数据预处理

行业拥挤度提升通常对应着未来资产标的发生回撤的风险，为了及时规避下跌风险，我们选取多个代理变量尝试对行业拥挤情况进行定量判断。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，如大牛市期间，不同行业的估值水平均会被抬升，以估值绝对水平刻画行业拥挤状态会导致拥挤行业数量大幅提升。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。

以换手率为例，不同行业由于特性差异，行业换手率不在同一个范围内，比如大金融、大周期板块中的行业换手率相对较低，而大科技板块中的行业由于市值整体偏小，业绩不确定性更高，换手率通常更高。

**图3：金融、周期板块包含的行业具有更低的换手率**



数据来源：Wind、开源证券研究所

为了通过比较不同行业换手率的变动来跟踪各行业当下的拥挤情况，我们首先需要对行业换手率进行去量纲调整，常用的去量纲方法有历史分位点法和两步标准化法。

### ● 历史分位点法

回溯过去一段时间，计算本行业当前换手率在回溯期内的历史分位情况。分位数能够有效剔除指标之间量纲差异的影响。

### ● 两步标准化法

两步标准化整体思路如下：首先在时序层面对各个行业分别进行标准化，再在横截面上在行业间进行标准化。具体如下：

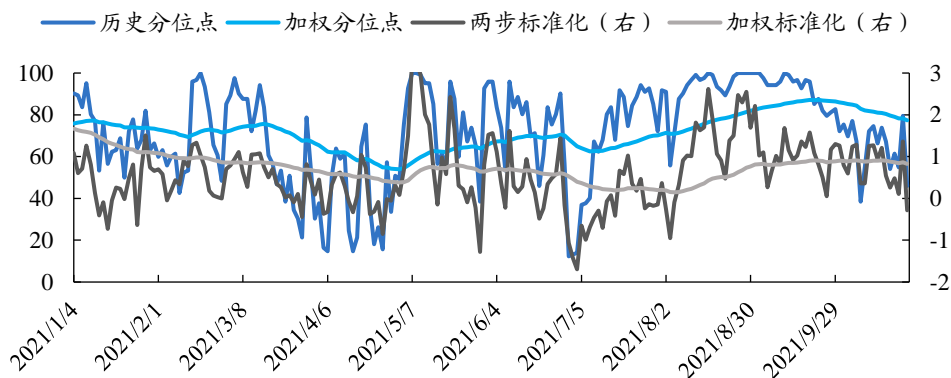
- (1). 对某个行业，将当前换手率减去过去一段时间换手率的均值，并除以区间内该行业换手率的标准差，得到当前时刻时序标准化后的换手率值；
- (2). 将时序标准化后的换手率值减去截面上各行业换手率的均值，并除以截面上换手率的标准差，得到当前行业调整后的换手率。

在对相关指标进行去量纲化后，我们发现去量纲后得到的数据波动相对较大。为了平抑波动，且对更新的信息赋予更高的权重，我们采用指数加权的方式对去量纲后的数值进行二次调整。

从调整后的结果来看，加权分位点和加权标准化两种方法对应的走势近似，表

明两种去量纲方法殊途同归。由于加权分位点法对应的数值更加直观，且相对更加灵敏，以下测试我们主要基于加权后历史分位点来进行相关指标测算。

图4：指数加权后，两种标准化方法的走势接近



数据来源：Wind、开源证券研究所

### 3、拥挤度代理变量测试

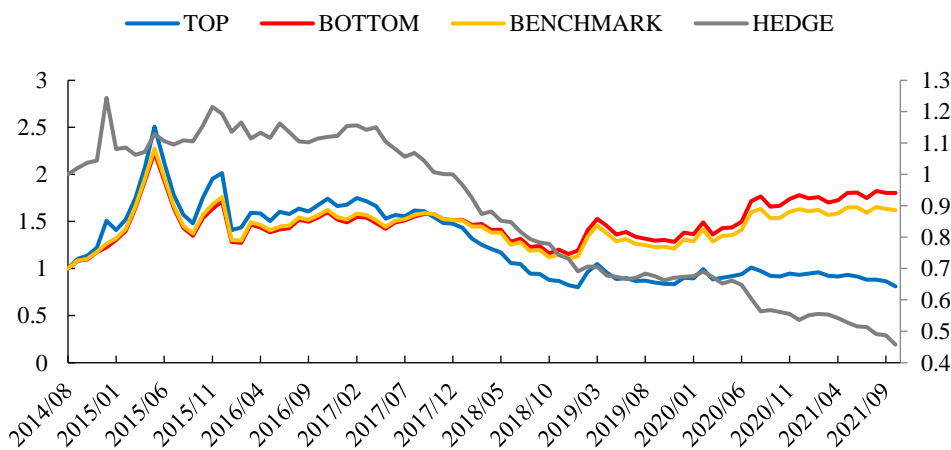
在对构建的拥挤度代理变量进行去量纲处理后，我们开始评估所构建的拥挤度代理变量是否有效。有效性的判断标准为：基于拥挤度代理变量筛选出的高拥挤行业，未来是否显著跑输非拥挤行业。在进行回测时，我们按照月度频率进行调仓，取指数加权后的最新值作为月末的信号进行下一期拥挤行业的判断。由于行业拥挤与否是一个极端状态，所以在高低分组时，我们采取不同的配置比例。高分组选择排序值最高的四个行业，剩余的行业作为低分组，比较高低两组的表現情况。

#### 3.1、换手率：高换手率是市场情绪高涨的有效映射

在对换手率进行去量纲处理时，会涉及到对时间窗口长度的选择，我们将回溯期设置为122个交易日（半年）。时间窗口选择相对较长期限，主要考虑到A股的市场从兴起到过热是一个逐步累积的过程，跨越的周期相对较长。

从测试结果来看，2016年以来，高换手率组表现持续跑输低换手率组，表明高换手率组未来潜在超额收益稳定为负，规避当下高换手率的行业可能是不错的选择。

图5：高相对换手率组在2016年后逐渐表现出负超额收益



数据来源：Wind、开源证券研究所

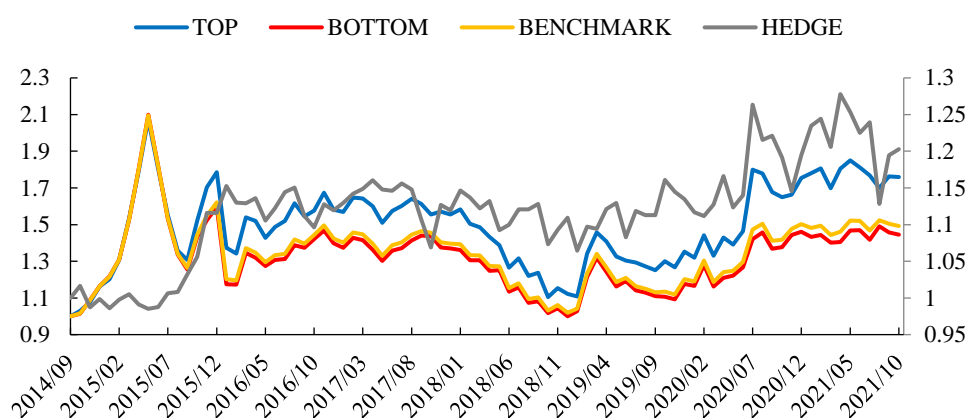


### 3.2、下行波动率：高下行波动率并不是刻画拥挤程度的良好指标

不同行业波动率之间也存在量级的差异，新兴板块包含的行业波动率会相应更大一些，传统板块则相对较低，为此我们对不同行业的波动率也做去量纲处理。

在计算下行波动率时，需要定义两个参数，第一个参数是下行波动率本身的计算周期，这里设置为 22 个交易日；第二个参数是在计算分位点时需要回溯的长度，我们秉承前文的思路，选择回溯 122 个交易日来计算下行波动率去量纲后的值。可以看到，个股下行波动率用来作为拥挤度的衡量指标存在较高失效的概率。通常股价波动率不仅体现在下跌过程中放大，更可能在股价拉涨期间放大。

图6：高下行波动率组并不是拥挤度的良好代理变量



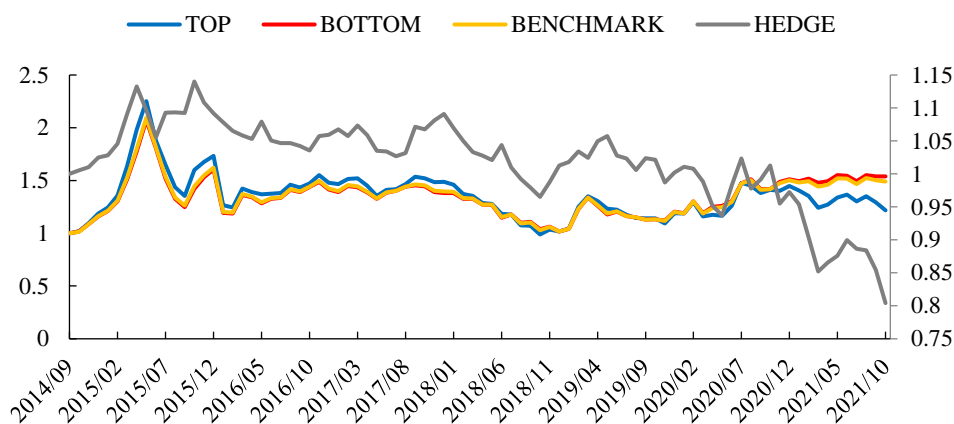
数据来源：Wind、开源证券研究所

### 3.3、配对相关性：配对相关性能有效衡量行业内个股的同涨同跌现象

配对相关性衡量了个股之间涨跌的联动性，个股之间的更高联动通常意味着行情共性开始占据主导，个股特性开始被忽略。共性占据主导，说明行业内个股被无差别买入或卖出，行业处于过热或者被抛售状态的概率加大。

在计算行业内个股的配对相关性时，需要对个股收益率的长度进行选取，此处设置为 122 个交易日，在标准化过程中，与前文保持一致，采用回溯 122 个交易日。从结果来看，高配对相关性组在测试期内有一定的反向预测能力，但主要是 2021 年以后。我们猜测，配对相关性表现不好可能由于本身这个指标没有方向性，高配对相关性既可以对应行业的整体上涨，亦可以是行业无差别下跌。但是两种方向的运动对策略的影响却截然不同。

图7：高配对相关性组有一定的反向预测能力，但波动较大



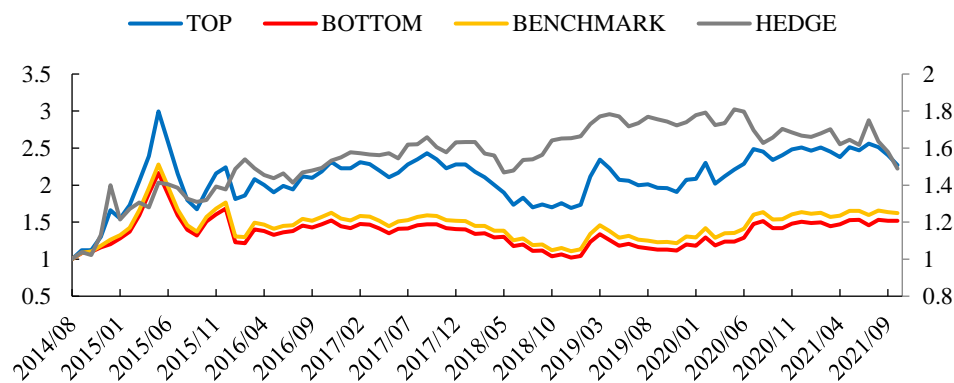
数据来源：Wind、开源证券研究所

### 3.4、成交量乖离率：高成交量乖离率或许是新增资金入场的信号

乖离率能够较好地刻画当前值与历史均值的偏离程度，当偏离度逐渐增大时，通常表明市场情绪处于过热状态，后续均值回复的概率加大。相比于价格乖离，我们使用成交量乖离来衡量行业拥挤状态。量在价前，成交量是否更具有代表性？

从测试结果来看，在2020年前，成交量乖离率反而是做多的优秀指标，其内在逻辑可能在于当行业呈现放量状态时，表明有新增资金不断介入，反而提升了未来股价上涨的概率。

图8：成交量乖离率并不适合用来刻画行业拥挤状态

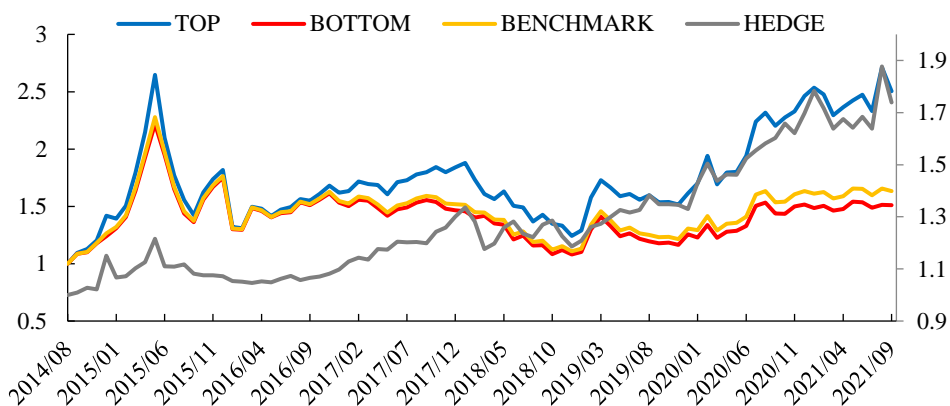


数据来源：Wind、开源证券研究所

### 3.5、相对估值：高估值并不是行业未来下跌的充分理由

行业之间的估值差异我们以PB进行衡量。估值水平中基本面信息的变动是个相对较缓慢的过程，在去量纲过程中，我们将回溯期设置为122个交易日。从结果来看，相对估值并不是衡量行业是否拥挤的合适指标，反而高估值行业始终受到市场的追捧。在A股市场，贵并不是股价下跌的充分理由，行业面的整体偏贵通常对应着行业高景气，如近两年市场上火热的新能源行业。另一个层面，高估值通常是高动量的外在表现，导致以高估值为代表的拥挤预警失灵。

图9：高相对估值组稳定跑赢基准，与预期相反



数据来源：Wind、开源证券研究所

## 4、拥挤度代理变量融合

在前文，我们按照去量纲后数值的大小，每期选出四个行业作为高拥挤行业并予以剔除，发现绝大部分指标的表现都乏善可称。同时，当整个市场成交都非常清淡的时候，可能不存在高拥挤行业，但是以上测试方法还是会每期对部分行业打上高拥挤的标签，不太符合拥挤度作为极端情形下风险提示的设定。

他山之石，可以攻玉。在本节，我们尝试借鉴文献《Crowded Trades - Implications For Sector Rotation And Factor Timing》中的思路，将不同行业根据所选代理变量的大小划分为四种状态，分别代表股价变动的不同阶段：非泡沫期（泡沫开始前）、泡沫膨胀期、泡沫破裂期和非泡沫期（泡沫破裂后）。在原始文献中，作者提出了资产集中度和相对价值两个指标，分别用来代表市场的交易热度和识别泡沫膨胀到破裂的拐点。首先，我们尝试复现资产集中度和相对价值两个代理变量在A股的表现情况。

### 4.1、资产集中度可识别泡沫存在与否

在参考文献中，资产集中度用来衡量板块或行业的交易热度。在板块或行业处于交投活跃状态时，其波动会不断放大，资产集中度会不断提升。值得注意的是，虽然资产集中度能够判断板块或行业是否存在泡沫，但是无法预测泡沫何时破裂。资产集中度的主要计算步骤如下：

- （1）提取各板块的历史收益率，对其进行加权，权重为每个板块市值的平方根，同时采用指数加权对近期的数据赋予更高的权重，半衰期设为1年
- （2）计算各板块（行业）收益率的协方差矩阵
- （3）对协方差矩阵进行主成分分析，得到特征向量
- （4）根据所选特征向量的值，计算吸收比率

$$AR = \frac{\sum_{i=1}^n \sigma_{E_i}^2}{\sum_{j=1}^N \sigma_{A_j}^2}$$

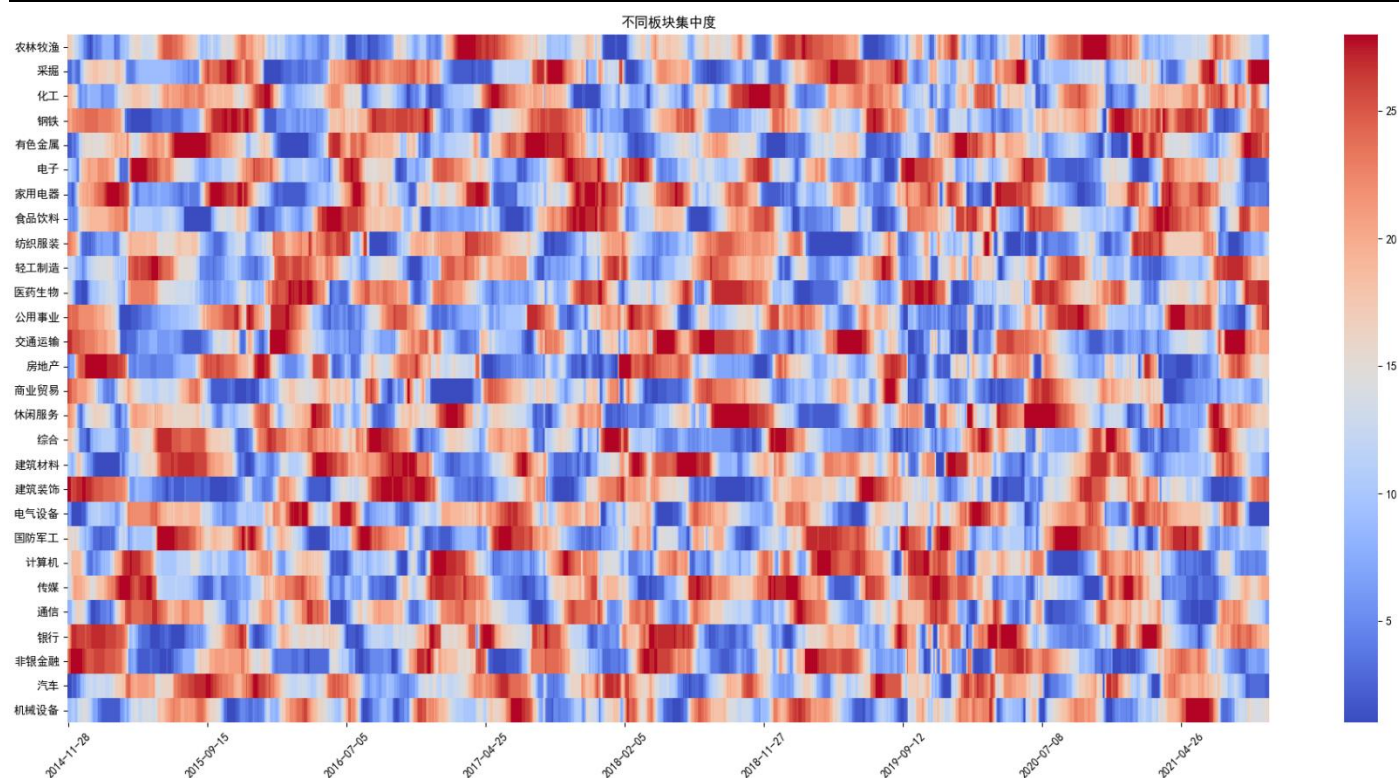
- （5）对吸收比率进行加权得到资产集中度

$$C_i = \frac{\sum_{j=1}^n \left( AR^j * \frac{|EV_i^j|}{\sum_{k=1}^N |EV_k^j|} \right)}{\sum_{j=1}^n AR^j}$$

其中， $N$ 表示资产数量， $n$ 表示用来计算吸收比率的特征向量的数量， $\sigma_{E_i}^2$ 表示第 $i$ 个特征向量的方差， $\sigma_{A_j}^2$ 表示第 $j$ 个资产的方差， $C_i$ 表示资产 $i$ 的集中度， $AR^j$ 表示第 $j$ 个特征向量的吸收比率， $EV_i^j$ 表示第 $i$ 个资产在第 $j$ 个特征向量中的暴露值。

根据以上规则，我们测算了 A 股不同行业的资产集中度变化。区别于参考文献中展示的美股不同行业资产集中度具有相对长期的稳定性，A 股不同行业资产集中度变化更加频繁。侧面表明 A 股的行业轮动的节奏更快，行情持续性相对更弱。

图10: A 股不同行业资产集中度变动更加频繁



数据来源: Wind、开源证券研究所

## 4.2、相对价值能够判别泡沫从膨胀到破裂的拐点

由于资产集中度指标只能识别是否存在泡沫，无法判断泡沫破裂的拐点，作者提出了相对价值指标，用来识别泡沫膨胀到泡沫破裂之间的拐点。相对价值指标在泡沫膨胀期不断上升，但是在泡沫破裂期，会呈现下降态势，与泡沫变化的走势近似。原文阐述了以行业 PB 值为基础的相对价值指标的构建方法，具体步骤如下：

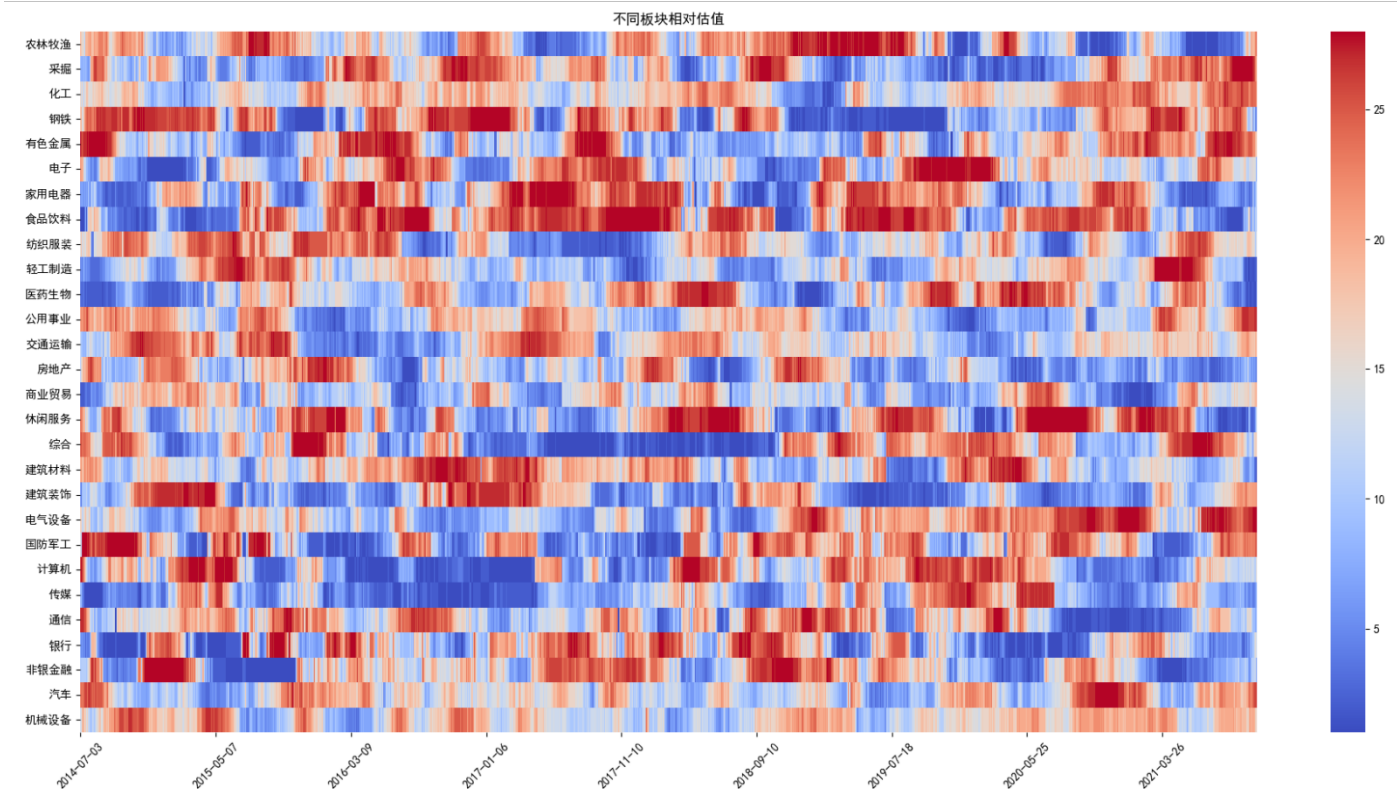
- (1). 在每个交易日，计算所有板块的日度 PB 值；
- (2). 将每日的 PB 除以过去一段时间 PB 的均值，得到标准化 PB 值；
- (3). 将每个行业标准化后的 PB 值除以剩余其他行业 PB 值的均值，得到每个行业 PB 值。



在 A 股市场,由于行业轮动节奏更快,我们对原文中的初始参数做了一定调整。在计算相对价值指标时,我们选取过去 122 个交易日的数据进行时序标准化运算。

可以看到,相对价值能够较好地反应市场上各行业的冷热状态。以几个具体的例子进行说明,农林牧渔行业在 2018 年~2019 年之间由于猪周期具有亮眼的表现,其相对价值水平亦有明显的提升,热度图呈现深红色;以采掘、化工和钢铁为代表的周期行业,在 2020 年 3 月至 2021 年 9 月之间,股价快速上涨,相对价值水平亦逐渐走高,尤以采掘为甚;前期相对价值较高的家用电器、食品饮料和医药生物行业相对价值水平在 2021 年则回落较多,热度图逐渐从前期的红色转为蓝色。

图11: 相对价值指标能够较好地反应市场上各行业的冷热状态变化



数据来源: Wind、开源证券研究所

#### 4.3、象限切分后不同区域代表行情的不同阶段

根据不同行业的资产集中度和相对价值大小,我们对 28 个一级行业进行状态划分,总共分为四个象限,具体如下:

- (1) 低资产集中度 + 低相对价值: 非泡沫期 A
- (2) 高资产集中度 + 低相对价值: 泡沫膨胀期 B
- (3) 高资产集中度 + 高相对价值: 泡沫破裂期 C
- (4) 低资产集中度 + 高相对价值: 非泡沫期 D

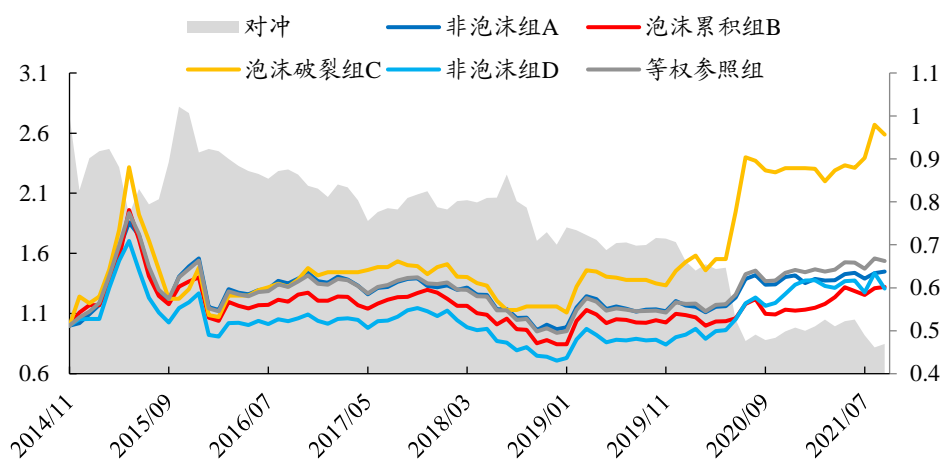
根据原文观点,泡沫膨胀组通常对应股价迅猛上升的阶段,累积收益率最高,而泡沫破裂组则是股价快速下跌的阶段,累积收益率最低。我们尝试在 A 股市场上验证基于资产集中度和相对价值指标的象限切分是否有符合预期的表现。



在每个自然月末，计算资产集中度和相对价值指标的月度均值。对于 28 个一级行业，我们分别按照资产集中度和相对价值数值的大小，分为高低两组，高低两组包含的行业个数按照 1:3 进行配置，即高分组包括 7 个行业，低分组包括 21 个行业。由于各个组别划分得到的象限是交集关系，所以各组入选的行业个数存在差异，且在时序上处于变化状态。高低组行业个数设置为不同数量的原因在于，拥挤度本身刻画的是一个阈值概念，只有超过特定范围的值才更有意义，反映到行业个数分配上，便是高低两组的行业个数的差异化。

从测试结果来看，泡沫膨胀组 B 表现平平，但泡沫破裂组 C 却在全区间内具有显著的超额收益。适用于国外成熟市场的代理变量并不完全适用于 A 股这个弱有效市场。当下的 A 股对估值容忍度较高，在行业交投活跃的背景下，相对价值提升反而是市场资金持续流入的反应。与前文对估值指标测试结果相一致，估值相对越高的行业未来一个月通常表现越好。

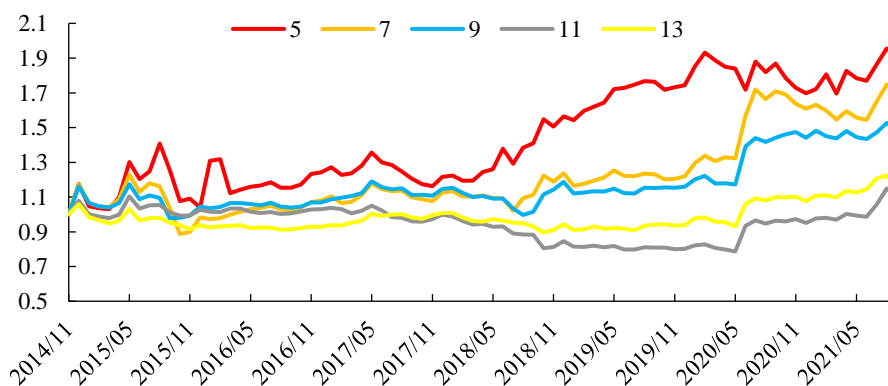
图12: 资产集中度和相对价值无法有效地捕获 A 股泡沫



数据来源: Wind、开源证券研究所

为避免偶然性，我们对高分组包含的行业个数进行敏感性测试。可以看到，高分组纳入的行业个数越少，泡沫破裂组 C 未来对应的超额收益越高，进一步证实了以上两个代理变量并不适合用来刻画 A 股行业的拥挤情况。鉴于此，我们尝试寻找更加符合 A 股特征的指标进行测试。

图13: 高分组纳入的行业数量越少，未来超额收益越高



数据来源: Wind、开源证券研究所

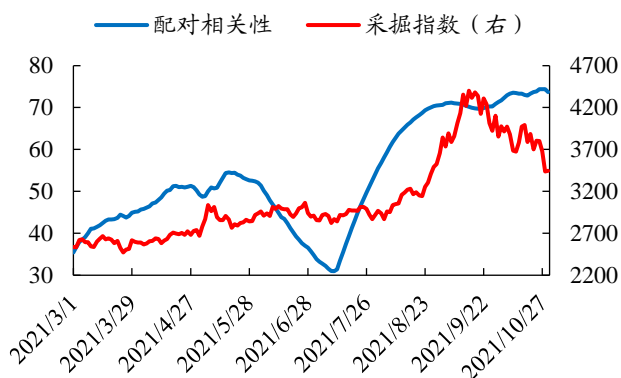
#### 4.4、代理变量的选择

前文中我们测试了资产集中度和相对价值指标对 A 股行业泡沫识别的表现，从测试结果来看，二者的融合指标并不能对 A 股行业泡沫破裂阶段进行有效识别，反而作者定义的泡沫破裂阶段是收益表现最好的一组。拆解代理变量失效的原因，我们认为资产集中度无法很好地刻画 A 股市场的交易活跃度，相对价值也偏离 A 股对估值相对不敏感的行为特征。

我们尝试针对 A 股市场的特性选取相应的指标，来更好地参与到泡沫膨胀阶段，并提前退出泡沫破裂阶段。根据第三节的测试结果，换手率和配对相关性具有相对较优的风险预测能力，在测试期内呈现相对稳定的负超额，且二者的变化比较符合原始文献中作者对泡沫膨胀和泡沫破裂时的特征设定。

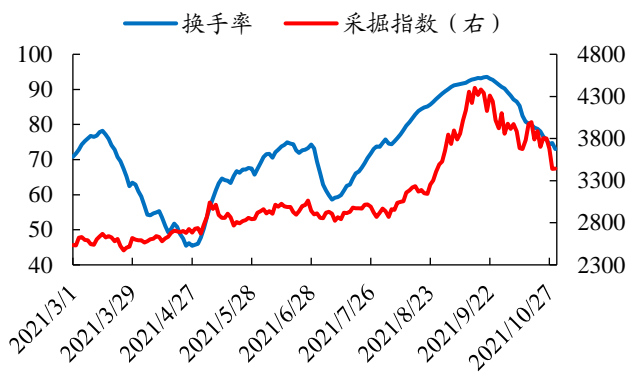
我们以 2021 年的采掘(煤炭)行业为例，观察配对相关性和换手率的变化情况。在行业股价处于快速上行的过程中，配对相关性会处于不断上升的状态，且在行业股价转为快速下跌时，短期内依然能够维持在高位(图 14)。因此，我们尝试用配对相关性来替换资产集中度来表征市场交易热度。换手率在股价快速上行时亦会快速提升，在股价快速下跌时，由于处置效应，换手率则会快速衰减，比较符合泡沫破裂时的特征(图 15)。

图14：配对相关性可用来识别泡沫是否存在



数据来源：Wind、开源证券研究所

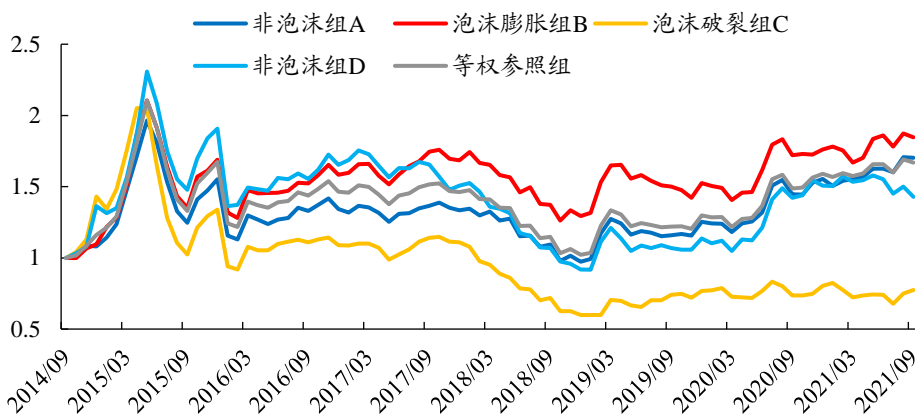
图15：换手率可用来识别泡沫何时破裂



数据来源：Wind、开源证券研究所

我们根据配对相关性和换手率指标将所有行业切分到四个象限，可以看到调整后的结果有所改善，且整体结论符合预期：泡沫破裂组 C 明显跑输其他组别，泡沫膨胀组 B 跑赢其他所有组别。

图16: 配对相关性+换手率能够较好地识别行业泡沫膨胀和破裂阶段



数据来源: Wind、开源证券研究所

根据测试结果, 泡沫破裂组 C 在年化收益率层面呈现负值, 低于任意分组, 且最大回撤达到 64.4%。相比之下, 泡沫膨胀组 B 表现较好, 年化收益率相比于等权参照组提升 2.2%, 最大回撤层从 51.5% 下降到 42.3%。

表1: 泡沫破裂组 C 显著跑输等权参照组

分组	年化收益率	年化波动率	夏普比率	最大回撤	月度胜率
非泡沫组 A	7.7%	24.2%	0.32	-49.0%	57.6%
泡沫膨胀组 B	9.3%	23.5%	0.39	-42.3%	51.8%
泡沫破裂组 C	-3.5%	29.3%	-0.12	-64.4%	45.9%
非泡沫组 D	5.6%	27.2%	0.21	-60.4%	50.6%
等权参照组	7.1%	24.3%	0.29	-51.5%	55.3%

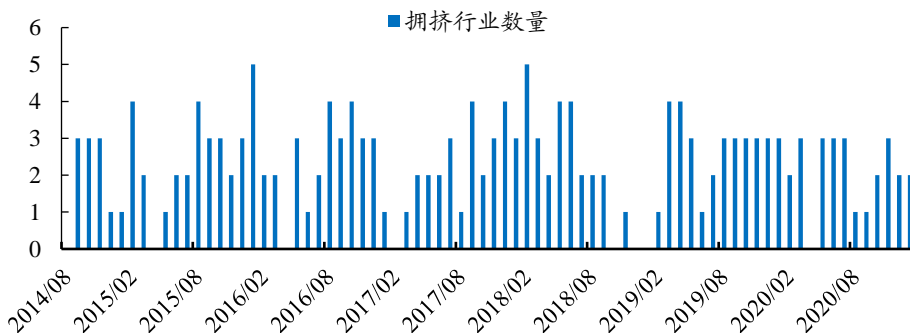
数据来源: Wind、开源证券研究所

## 5、拥挤度的实践

### 5.1、基于拥挤度的行业剔除

根据泡沫破裂组的定义, 我们每个月选出的拥挤行业数量如图 17 所示。在样本期内, 每个月给出拥挤信号的行业个数中位数为 3 个, 特定时间段拥挤行业个数为 0, 单月拥挤行业个数最多为 5 个。

图17: 基于配对相关性和换手率的每月拥挤行业个数中位数为 3 个

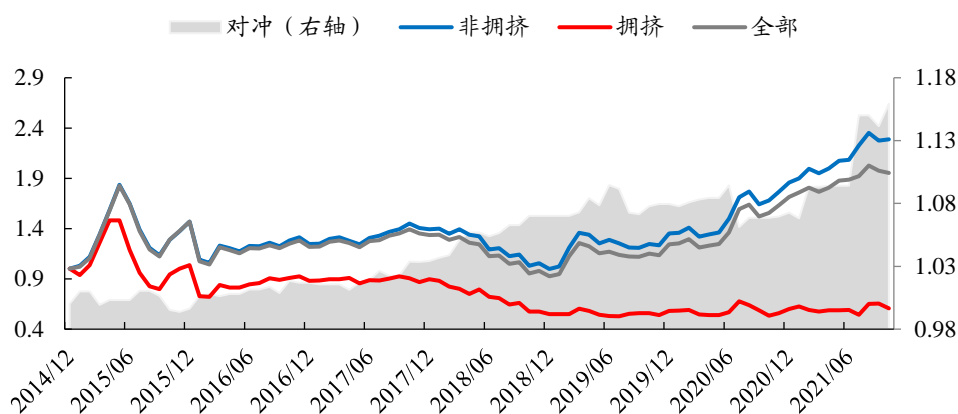


数据来源: Wind、开源证券研究所

为了检验泡沫破裂组 C 在实际行业轮动信号中的改善效果，我们将拥挤信号融入到既有模型中，观察剔除拥挤行业后的剩余行业表现情况。值得注意的是，在实际应用过程中，所选行业轮动模型本身并不是每期均含有拥挤状态的行业，所以拥挤行业剔除的操作并不必然每期都会触发。以开源金融工程团队的行业配置体系为例（详情参见报告《行业配置的量化解决方案》），测试结果如下：

测试期内，累计调仓 82 次，剔除拥挤行业的操作触发了 58 次，占比 71%，剔除拥挤行业后当月超额收益为正的次数为 38 次，占比约 66%，测试期内累积超额收益率提高约 16%，年化超额约 2%。

**图18：不含拥挤行业的轮动模型表现年化收益率提升约 2%**



数据来源：Wind、开源证券研究所

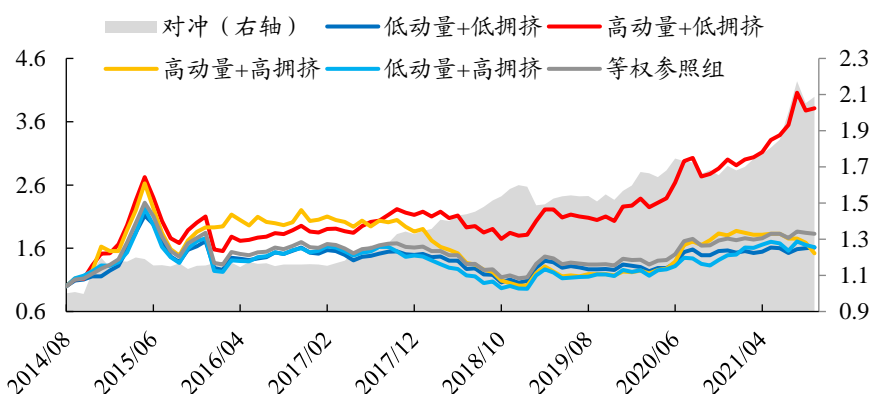
我们亦尝试在全市场 28 个行业中每期剔除拥挤行业，观察剩余行业相对全体行业的改善幅度，但效果并不好。究其原因，主要在于拥挤行业权重占比过小导致对结果影响偏低。举例来说，全市场有 28 个行业，在等权配置的前提下，每个行业权重占比仅约 3.6%，若当期选出两个拥挤行业，则相当于将 7.2% 的权重对应的未来收益率赋值为 0，当所选的两个行业未来跑输基准 1%，则剩余非拥挤行业等权收益率仅提升 0.077%，年化收益率提升不足一个百分点。

## 5.2、纳入拥挤度的行业轮动

基于此，我们认为单纯的基于行业拥挤状态进行行业剔除不是拥挤度的最佳实践方式。参考不同指标下的象限切分方法，我们考虑将拥挤度作为选择行业的一个维度，在此基础上纳入其他维度信息，比如动量维度，观察二者的结合能否有效提升所选行业的表现。

动量因子这里做简化处理，选取过去一个月各行业的涨跌幅作为行业动量因子。在第三小节，换手率分位点具有较稳定的负向超额，我们尝试作为拥挤因子的代理变量，测试动量加拥挤的双维度行业筛选表现。这里高低分组依然保持倾斜配置比例，按月调仓。从结果来看，高动量+低拥挤的行业表现最好，符合预期。在 2017 年之前，高动量+高拥挤的行业具有较好的表现，但后续逐渐被市场抛弃。

图19: 高动量+低拥挤组表现优异，超额稳定性较好



数据来源: Wind、开源证券研究所

具体来看，高动量+低拥挤组合年化收益率达到 18.8%，最大回撤仅 42.9%，改善效果显著。且对冲效果较稳定，年化超额收益率达 10.7%，最大回撤为 7.1%。

表2: 高动量+低拥挤组合能有效改善策略表现

分组	年化收益率	年化波动率	夏普比率	最大回撤	胜率
低动量+低拥挤	5.6%	23.9%	0.23	-50.1%	55.8%
高动量+低拥挤	18.8%	25.9%	0.72	-42.9%	65.1%
高动量+高拥挤	4.4%	28.1%	0.16	-61.6%	43.0%
低动量+高拥挤	5.2%	26.6%	0.20	-56.9%	51.2%
等权参照组	7.3%	24.4%	0.30	-51.5%	55.8%
对冲	10.7%	9.5%	1.13	-7.1%	65.1%

数据来源: Wind、开源证券研究所

## 6、风险提示

模型测试基于历史数据，市场未来可能发生变化。



## 特别声明

《证券期货投资者适当性管理办法》、《证券经营机构投资者适当性管理实施指引（试行）》已于2017年7月1日起正式实施。根据上述规定，开源证券评定此研报的风险等级为R3（中风险），因此通过公共平台推送的研报其适用的投资者类别仅限定为专业投资者及风险承受能力为C3、C4、C5的普通投资者。若您并非专业投资者及风险承受能力为C3、C4、C5的普通投资者，请取消阅读，请勿收藏、接收或使用本研报中的任何信息。因此受限于访问权限的设置，若给您造成不便，烦请见谅！感谢您给予的理解与配合。

## 分析师承诺

负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映分析人员的个人观点。负责准备本报告的分析师获取报酬的评判因素包括研究的质量和准确性、客户的反馈、竞争性因素以及开源证券股份有限公司的整体收益。所有研究分析师或工作人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中具体的推荐意见或观点有直接或间接的联系。

## 股票投资评级说明

	评级	说明
证券评级	买入（Buy）	预计相对强于市场表现 20%以上；
	增持（outperform）	预计相对强于市场表现 5% ~ 20%；
	中性（Neutral）	预计相对市场表现在 - 5% ~ + 5%之间波动；
	减持（underperform）	预计相对弱于市场表现 5%以下。
行业评级	看好（overweight）	预计行业超越整体市场表现；
	中性（Neutral）	预计行业与整体市场表现基本持平；
	看淡（underperform）	预计行业弱于整体市场表现。
备注：评级标准为以报告日后的 6~12 个月内，证券相对于市场基准指数的涨跌幅表现，其中 A 股基准指数为沪深 300 指数、港股基准指数为恒生指数、新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）、美股基准指数为标普 500 或纳斯达克综合指数。我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。		

## 分析、估值方法的局限性说明

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。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

## 法律声明

开源证券股份有限公司是经中国证监会批准设立的证券经营机构，已具备证券投资咨询业务资格。

本报告仅供开源证券股份有限公司（以下简称“本公司”）的机构或个人客户（以下简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。本报告是发送给开源证券客户的，属于机密材料，只有开源证券客户才能参考或使用，如接收人并非开源证券客户，请及时退回并删除。

本报告是基于本公司认为可靠的已公开信息，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他金融工具的邀请或向人做出邀请。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、价值及投资收入可能会波动。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。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，不应视本报告为做出投资决策的唯一因素。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。本公司未确保本报告充分考虑到个别客户特殊的投资目标、财务状况或需要。本公司建议客户应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。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。在任何情况下，本公司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。若本报告的接收人非本公司的客户，应在基于本报告做出任何投资决定或就本报告要求任何解释前咨询独立投资顾问。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的开源证券网站以外的地址或超级链接，开源证券不对其内容负责。本报告提供这些地址或超级链接的目的纯粹是为了客户使用方便，链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

开源证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。开源证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。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。

## 开源证券研究所

### 上海

地址：上海市浦东新区世纪大道1788号陆家嘴金控广场1号楼10层  
邮编：200120  
邮箱：research@kysec.cn

### 深圳

地址：深圳市福田区金田路2030号卓越世纪中心1号楼45层  
邮编：518000  
邮箱：research@kysec.cn

### 北京

地址：北京市西城区西直门外大街18号金贸大厦C2座16层  
邮编：100044  
邮箱：research@kysec.cn

### 西安

地址：西安市高新区锦业路1号都市之门B座5层  
邮编：710065  
邮箱：research@kysec.cn