

2021 年 02 月 23 日

北上资金行业配置的双轮驱动力

金融工程研究团队

——开源量化评论（20）

魏建榕（首席分析师）

邮箱：weijianrong@kysec.cn

证书编号：S0790519120001

张翔（分析师）

邮箱：zhangxiang2@kysec.cn

证书编号：S0790520110001

傅开波（分析师）

邮箱：fukaibo@kysec.cn

证书编号：S0790520090003

高鹏（分析师）

邮箱：gaopeng@kysec.cn

证书编号：S0790520090002

苏俊豪（研究员）

邮箱：sujunhao@kysec.cn

证书编号：S0790120020012

胡亮勇（研究员）

邮箱：huliangyong@kysec.cn

证书编号：S0790120030040

王志豪（研究员）

邮箱：wangzhihao@kysec.cn

证书编号：S0790120070080

相关研究报告

《开源量化评论（10）-北上资金行为视角下的行业轮动》-2020.10.28

《开源量化评论（16）-行业量化配置：在重要行业上的应用》-2021.01.28

《开源量化评论（18）-行业量化配置：在沪深 300 增强上的应用》-2021.02.17

魏建榕（分析师）

weijianrong@kysec.cn

证书编号：S0790519120001

张翔（分析师）

zhangxiang2@kysec.cn

证书编号：S0790520110001

胡亮勇（联系人）

huliangyong@kysec.cn

证书编号：S0790120030040

●北上资金交易行为因子可以从历史偏好、定价权和边际变化三大维度展开

当前北上资金通过陆股通配置 A 股的每日明细，港交所的官方网站披露易次日都会进行详细展示，这为我们增进对北上资金投资行为的了解提供了很好的窗口。北上资金在不同行业上的成交金额、净流入金额、持仓规模等信息，归纳起来可以分别用来代表北上资金对某个行业的历史偏好、定价权和边际变化。

以 2020 年 12 月 31 日的截面数据为例，按照北上资金持仓占比偏离行业实际占比幅度大小排序，家用电器和休闲服务高配幅度最大，分别高配 1.77 倍和 1.39 倍，纺织服装和国防军工低配幅度最明显，配置比例仅占行业实际占比的 14% 和 36%。

北上资金在银行、食品饮料和家用电器行业整体成交金额占比处于最高一档，具有更强的定价能力，通信、综合和纺织服装成交金额占比则处于相对后排的位置，定价参与度不足。

当前北上资金持有 A 股规模达到一定体量，且追踪北上资金调仓行为的潜在资金不断增多的背景下，北上资金边际偏好的变化对市场的影响亦在逐渐放大。最显著的例子莫过于北上资金分钟净流入与分钟市场涨跌的 R^2 在 2020 年 7 月后开始跃升，近期虽有回落，但依然保持在高位。

●历史偏好代表的存量信息和边际变化代表的增量信息是北上资金行业配置能力的双轮驱动力

历史偏好因子和定价权因子具有较高的关联度，二者相关系数达到 0.8 以上，而带有择时特征的边际变化因子，与其他因子整体相关性较低，普遍低于 0.3。

通过回溯北上资金在不同因子层面的持仓，其中食品饮料、家用电器、医药生物、电子和银行是北上资金相对青睐的行业，而这些高频出现的行业亦是 2017 年以来市场上表现相对亮眼的行业。

当使用成交金额对持仓规模进行截面回归得到成交金额残差因子时，其超额收益能力下降明显，说明成交金额本身的超额收益来源与持仓规模高度相关。3 个月净流入金额因子在截面上对持仓规模因子进行回归后的残差因子，净值表现出现改善，说明净流入因子是区别于持仓规模因子的增量信息点。

合成因子的夏普比率从存量因子的 2.1 上升到 2.4，最大回撤相比增量因子从 -6% 降低到 -4.4%，胜率无论是相对增量因子还是存量因子，均提升明显，达 74%。

概括地讲，北上资金在行业轮动上超额收益的主要来源是：以持仓规模为代表的历史偏好类因子所蕴含的存量信息，和以净流入金额为代表的边际变化因子所蕴含的增量信息。而存量因子和增量因子的联合又能够相互取长补短，提供更高夏普比率的多空对冲曲线，二者构成了北上资金行业配置的双轮驱动力。

●风险提示：模型测试基于历史数据，市场未来可能发生变化。

目 录

1、 交易行为因子构建	4
2、 探索性数据分析	6
2.1、 历史偏好	6
2.2、 定价权	7
2.3、 边际变化	9
3、 交易行为因子分析	10
3.1、 因子相关性	10
3.2、 因子测试	11
3.2.1、 历史偏好：持仓规模 vs. 持仓规模占比	11
3.2.2、 定价权：成交金额 vs. 成交金额占比	11
3.2.3、 边际变化：净流入金额 vs. 净流入金额占比	12
3.3、 交易行为因子的多维对比	12
3.3.1、 不同因子的多指标比较	12
3.3.2、 收益来源再探索	13
3.3.3、 参数敏感性测试	14
3.4、 因子合成	15
4、 风险提示	16

图表目录

图 1： 北上资金在食品饮料、家用电器和休闲服务行业上高配比例显著	6
图 2： 北上资金在不同行业上的持仓规模变动（亿元）	7
图 3： 北上资金成交金额占比稳定提升	8
图 4： 估算成交金额与实际成交金额具有高度相关性（亿元）	8
图 5： 不同行业成交金额占比箱型图分布	9
图 6： 北上资金分钟净流入与分钟市场涨跌的 R^2 在 2020 年 7 月后跃升	9
图 7： 2020 年北上资金在不同行业净流入规模估算（亿元）	10
图 8： 不同因子的秩相关性	10
图 9： 持仓规模多空收益	11
图 10： 持仓规模占比多空收益	11
图 11： 成交金额分组收益	12
图 12： 成交金额占比分组收益	12
图 13： 净流入金额分组收益	12
图 14： 净流入金额占比分组收益	12
图 15： 不同因子的多头收益对比	13
图 16： 不同回溯时长下的因子表现对比	15
图 17： 成交金额回归持仓规模的残差表现	15
图 18： 净流入金额回归持仓规模的残差表现	15
图 19： 合成因子多空收益率曲线夏普比率更高	16

表 1： 因子选取及构建方式介绍	5
------------------------	---

表 2: 定价权类因子整体表现相对突出	13
表 3: 不同因子持仓行业次数统计	14

2017 年以来，随着北上资金对 A 股的影响力逐步提升，不同维度下针对北上资金持仓的分析研究如火如荼。北上资金持仓超额收益来源可以拆解到三个方面：择时（大级别下的择时）、选股（相对优秀的个股选择）、行业配置（相对基准的主动行业偏离）。北上资金的流入是一个长期的大趋势，因此北上资金的择时能力更多体现在其反向预警的作用，即北上资金大幅流出通常意味着市场未来的下行风险上升；选股能力则体现在持有北上资金的重仓股，从 17 年以来能显著跑赢主流宽基指数；在北上资金持股偏好已经被投资者反复挖掘的背景下，北上资金在行业配置上所带来的超越基准的回报也逐渐受到大家的关注。循着这个思路，本篇报告，我们试图探讨北上资金在行业配置上的超额收益来源，其在行业上的主动偏离能带来超额收益的内在驱动力是什么。

在陆股通系列第三篇《北向资金行为视角下的行业轮动》中，我们简略地讨论了北上资金托管机构层面的行业轮动能力差异，并发现净流入因子具有较稳定的行业轮动表现。本篇报告，我们希望在此基础上，对其内容进行完善处理，从更长时间、更多维度来剖析北上资金在行业配置上超额收益的来源。

1、交易行为因子构建

当前北上资金通过陆股通配置 A 股的每日明细，港交所的官方网站披露易次日都会进行详细展示，这为我们增进对北上资金投资行为的了解提供了很好的窗口。基于北上资金持仓信息，我们可以计算陆股通持有个股的成交金额、净买入金额和持仓规模，将个股数据在行业上聚合后便能得到行业因子值。北上资金在不同行业上的成交金额、净流入金额、持仓规模等信息，归纳起来可以分别用来代表北上资金对某个行业的历史偏好、定价权和边际变化。

（1）持仓规模

持仓规模表示北上资金在行业上的持有的流通市值，其刻画了北上资金的历史偏好，是北上资金用脚投票的长期结果。持仓规模的大小是资金持续流入的累积变化，而资金的持续流入是股价表现优秀的必要条件。

（2）持仓规模占比

持仓规模占比是持仓规模市值与该行业流通市值的比例。持仓规模占比是对行业市值偏离的修正，避免大市值行业天然具有高持仓规模特征对策略的影响，以便合理地对比北上资金在不同行业之间的历史偏好。

（3）成交金额

北上资金在特定行业的成交金额由该行业内陆股通标的个股成交额加总得到，个股成交额则由前后两个交易日持仓变化量的绝对值乘以当日均价得到，注意此处未考虑日内交易，因此估算成交金额理论上会低于实际成交金额。

成交金额类因子可以理解为投资者的定价权，成交金额越高，理论上对该股票价格的变动影响力越大。

（4）成交金额占比

成交金额占比是北上资金在该行业内所有陆股通标的个股成交金额之和占所在行业实际总成交金额的比例。其中，行业实际总成交金额由该行业内所有个股的成交金额加总得到。

(5) 净流入金额

净流入金额刻画了北上资金对某个行业偏好上的边际变化，净流入金额越大，表明北上资金近期对某个行业越看好。短期内的边际变化可以理解为北上资金在行业上的择时。

(6) 净流入金额占比

净流入金额占比是净流入金额与所属行业流通市值的比值。不同行业的体量是存在明显差异的，比如银行和休闲服务行业，流通市值差异显著，银行业高净流入金额并不必然代表高净流入金额占比。

以上六个因子，主要从**历史偏好**、**定价权**、**边际变化**三个维度进行了简单的介绍，具体计算公式如表 1 所示。

表1: 因子选取及构建方式介绍

因子名称	计算公式
持仓规模	$holding_mv = \sum_{j=1}^m close_j * holding_j$
持仓规模占比	$holding_ratio = \frac{holding_mv}{industry_liquid_mv}$
成交金额	$trade_amount = \sum_{j=1}^m \sum_{i=1}^n vwap_j * abs(\Delta holdings_{j,i})$
成交金额占比	$trade_ratio = \frac{trade_amount}{industry_trade_amount}$
净流入金额	$net_inflow = \sum_{j=1}^m vwap_j * \Delta holdings_j$
净流入金额占比	$inflow_ratio = \frac{net_inflow}{industry_liquid_mv}$

资料来源：开源证券研究所

其中， $vwap_j$ 表示第 j 只股票的成交均价， $\Delta holdings_{j,i}$ 表示该行业内的第 j 只股票在第 i 个托管机构的持仓股票变化数， n 表示持有该股票的托管机构数， m 表示该行业内的股票个数， $holding_j$ 表示第 j 只股票的持仓量， $\Delta holdings_j$ 表示第 j 只股票的持仓变动， $close_j$ 表示第 j 只股票的收盘价， $industry_liquid_mv$ 表示行业流通市值， $industry_trade_amount$ 表示行业成交金额。以上因子为了降低随机扰动的影响，均采用过去一个月的数值进行平滑处理。具体而言，历史偏好类因子采用最近一个月持仓规模的均值，定价权类因子采用最近一个月成交金额的均值，边际变化类因子采用最近一个月净流入金额之和。

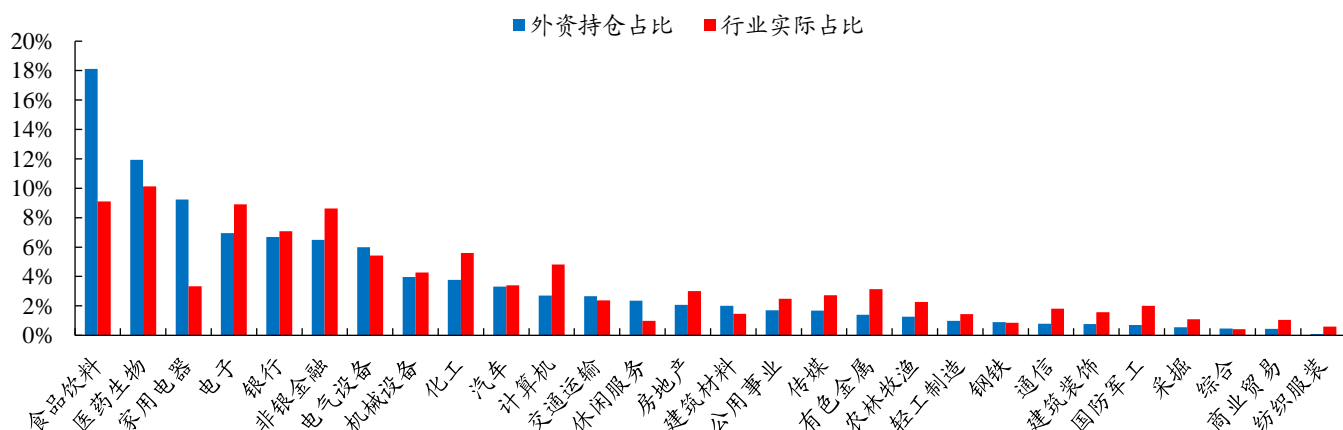
2、探索性数据分析

在因子构建部分，我们基于北上资金的持仓及其变动构建了交易行为类相关因子，代表了投资交易的三个维度。为了更清晰地展示不同交易行为因子在行业上的分布及变化情况，我们尝试对不同因子进行分析对比，以期直观地了解北上资金的交易行为特征。

2.1、历史偏好

北上资金在行业上的主动偏离反映了北上资金对不同行业的相对偏好程度。以 2020 年 12 月 31 日的截面数据为例，我们对比了北上资金在不同行业上的持仓规模占比和行业实际流通市值规模占比，可以发现，北上资金在食品饮料、医药生物和家用电器行业上存在高配现象，三者合计持仓规模达到总持仓规模的 39.27%，显著高于该三大行业的实际流通市值占比（22.58%），剩余相对高配行业还有电气设备、交通运输和休闲服务。按照北上资金持仓占比偏离行业实际占比幅度大小排序，家用电器和休闲服务高配幅度最大，分别高配 1.77 倍和 1.39 倍，纺织服装和国防军工低配幅度最明显，配置比例仅占行业实际占比的 14% 和 36%。

图1：北上资金在食品饮料、家用电器和休闲服务行业上高配比例显著

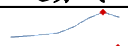





















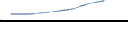



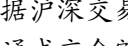
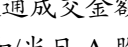


数据来源：Wind、开源证券研究所

数据日期：2020.12.31

按照持仓规模来看，2017 年以来，北上资金在不同行业上的持仓规模虽然有所波动（市场阶段性表现导致的持仓市值变动），但总体的趋势是向上的（北上资金的持续增配促进了整体向上趋势）。根据热度图的颜色可知，北上资金在不同行业上持仓规模的大小的顺序非常稳定，比如，食品饮料始终是所有行业中的持仓规模第一，纺织服装、商业贸易和综合行业一直不受北上资金青睐，整个行业的持仓规模甚至不如特定行业单个股票的持仓规模。

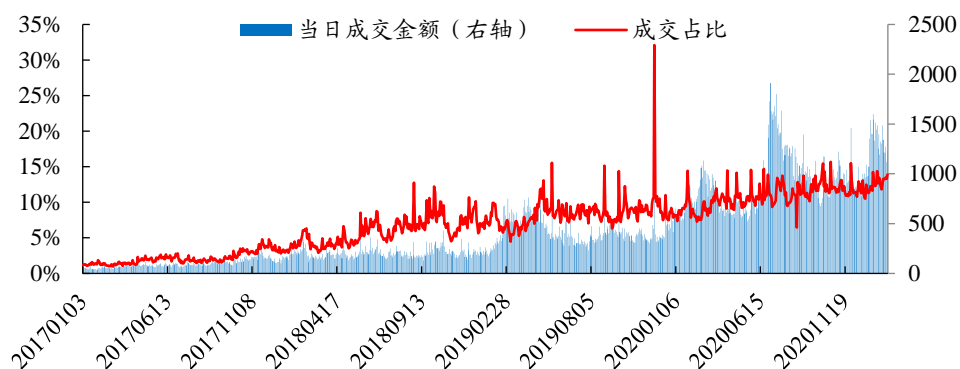
图2：北上资金在不同行业上的持仓规模变动（亿元）

行业名称	201706	201712	201806	201812	201906	201912	202006	202012	趋势线
农林牧渔	19.8	36.8	52.7	67.5	152.6	263.6	349.6	283.1	
采掘	14.3	26.9	39.4	50.0	82.4	110.7	74.4	130.5	
化工	64.5	93.8	140.3	156.2	178.4	285.6	394.9	883.1	
钢铁	20.5	47.8	73.0	91.5	77.2	105.2	93.0	199.4	
有色金属	25.8	41.5	64.8	54.7	81.9	153.1	204.8	342.6	
电子	312.2	499.2	605.7	391.2	488.3	864.0	1280.0	1553.5	
家用电器	373.1	749.7	792.7	602.8	982.1	1601.9	1557.3	2083.0	
食品饮料	628.4	960.3	1367.0	1272.4	2157.5	2465.7	3109.8	4315.2	
纺织服装	5.8	5.3	12.6	14.9	21.4	26.7	17.9	18.9	
轻工制造	19.7	25.2	57.9	33.6	50.7	114.8	164.0	229.5	
医药生物	305.2	492.0	806.4	552.2	926.4	1398.0	2378.8	2882.6	
公用事业	147.6	211.1	241.4	261.3	324.4	373.5	372.0	371.5	
交通运输	189.6	275.4	331.5	328.0	469.1	521.3	469.7	614.8	
房地产	47.5	82.5	87.1	160.1	274.9	486.9	439.2	434.5	
商业贸易	10.9	23.8	42.7	35.7	57.4	71.9	131.8	96.2	
休闲服务	37.7	103.0	160.4	165.8	237.8	246.3	415.6	624.9	
综合	2.1	5.4	10.7	12.2	37.7	62.2	69.8	110.1	
建筑材料	56.9	56.5	118.1	142.9	248.4	425.9	394.8	451.6	
建筑装饰	39.5	40.2	46.8	77.4	100.3	158.8	151.3	151.7	
电气设备	35.1	70.2	87.7	105.1	194.4	355.0	648.5	1592.4	
国防军工	24.2	22.0	25.6	33.0	54.5	93.0	124.9	172.6	
计算机	25.3	41.7	73.0	69.4	158.7	321.8	635.2	652.8	
传媒	45.5	40.0	78.2	83.4	121.7	229.7	353.0	365.8	
通信	23.9	52.8	35.8	54.4	94.3	163.2	201.0	169.3	
银行	272.5	374.9	412.4	581.7	887.6	1378.7	1106.7	1482.8	
非银金融	363.5	460.0	502.2	558.0	978.0	1256.1	1101.9	1491.1	
汽车	226.8	323.5	350.4	278.3	313.4	381.5	387.6	759.3	
机械设备	37.4	50.7	68.3	119.4	230.7	377.3	484.2	936.4	

数据来源：Wind、开源证券研究所

2.2、定价权

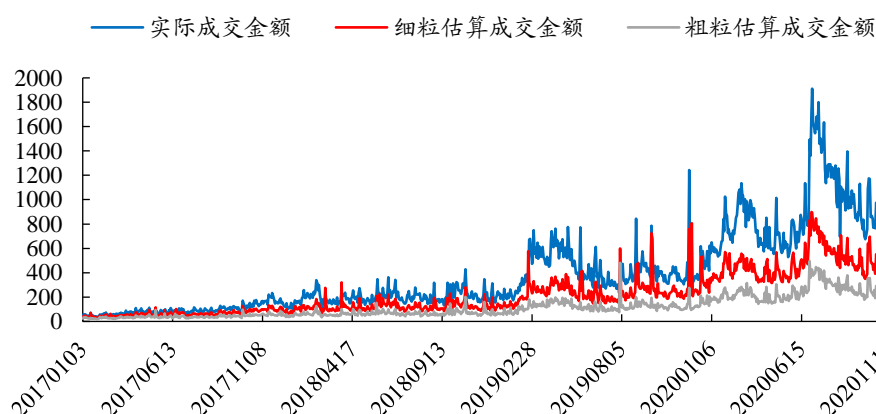
定价权衡量的是北上资金对特定个股、行业或板块的掌控程度。根据沪深交易所每日公布的数据可知，北上资金成交占比逐年提升，2017 年以来陆股通成交金额占比从初始仅约 1% 的水平逐步提升到当前约 14% 的水平（买卖金额之和/当日 A 股成交金额），以成交金额（占比）刻画的北上资金定价权稳定提升。

图3：北上资金成交金额占比稳定提升


数据来源：Wind、开源证券研究所

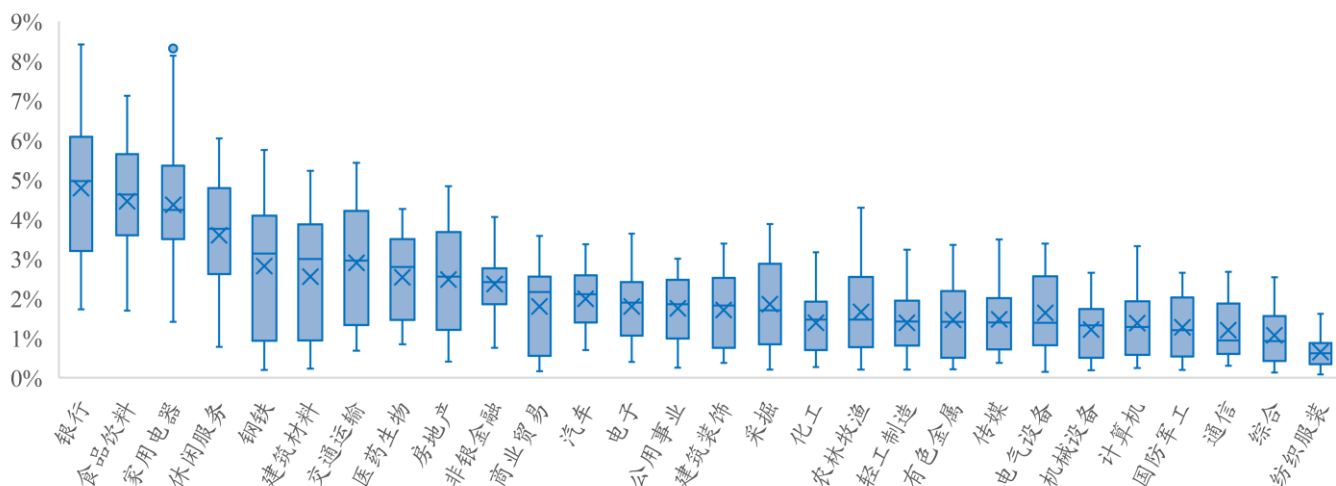
虽然交易所公布了陆股通每日成交的总额，但是并没有公布各个行业的成交金额，为此我们需要对各行业每日成交金额及其占比进行估算。行业成交金额的估算有两种方式，第一直接基于某个股票每日持仓变动的绝对值乘以当日均价来估算该股票当日成交金额，并在行业上进行聚合；第二基于更细分颗粒度的托管机构来进行估算，不同托管机构持有某个标的的持仓变动绝对值之和乘以当日均价得到该股票的当日成交金额，并在行业上聚合。

图3展示了两种估算方式下的成交金额走势，可以发现，两种计算方式得到的数值均与实际成交金额的走势高度相关，且两两之间的历史相关性均达到0.95以上。由于无法获取北上资金在不同行业上的实际成交金额，根据相似性原理，这里假设北上资金在各行业的估算成交金额与该行业实际成交金额亦有高相关性，那么我们可以用估算的行业成交金额来代替实际成交金额。为更加贴近实际，下文成交金额基于细粒估算方法得到。

图4：估算成交金额与实际成交金额具有高度相关性（亿元）


数据来源：港交所、Wind、开源证券研究所

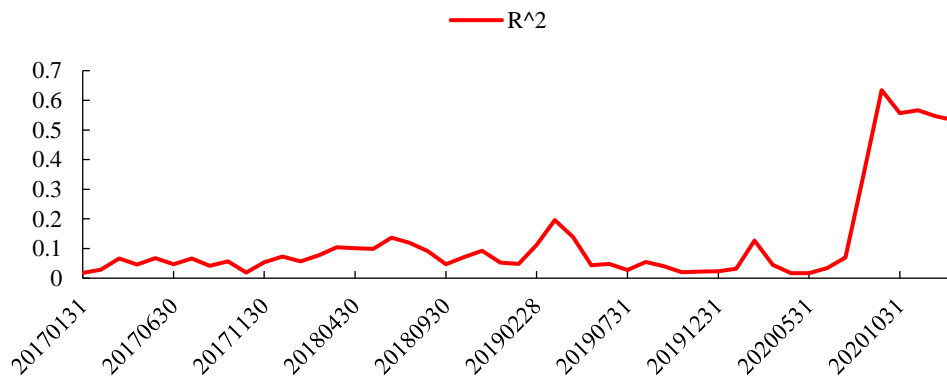
基于以上介绍方法，我们绘制了不同行业成交金额占比的箱型图分布。根据下图所示，银行、食品饮料和家用电器行业整体成交金额占比处于最高一档，通信、综合和纺织服装成交金额占比则处于相对后排的位置。

图5：不同行业成交金额占比箱型图分布


数据来源：港交所、Wind、开源证券研究所

2.3、边际变化

随着北上资金每日净流入金额与当日 A 股主流宽基指数的相关性越来越高，越来越多的投资者开始关注北上资金边际偏好的变化对个股及行业的影响。尤其在当前北上资金持有 A 股规模达到一定体量，且追踪北上资金调仓行为的潜在资金不断增多的背景下，北上资金边际偏好的变化对市场的影响亦在逐渐放大。最显著的例子莫过于北上资金分钟净流入与分钟市场涨跌的 R^2 在 2020 年 7 月后开始跃升，近期虽有回落，但依然保持在高位。

图6：北上资金分钟净流入与分钟市场涨跌的 R^2 在 2020 年 7 月后跃升


数据来源：港交所、开源证券研究所

以 2020 年为例，按月统计北上资金在不同行业的净流入规模，可以发现化工和电气设备行业全年仅 3 月份流出，其他月份均为净流入，食品饮料行业则是全年净流出规模最多的行业，不同行业上的边际变化存在明显的差异度。

图7：2020 年北上资金在不同行业净流入规模估算（亿元）

2020年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	柱形图
医药生物	73.4	67.4	18.6	88.9	199.4	114.3	-6.1	-31.9	-36.7	-45.4	15.0	116.9	
电气设备	26.7	42.9	-41.4	21.2	50.0	59.8	19.3	3.0	36.9	17.7	86.0	159.9	
电子	11.5	-32.7	-49.8	63.2	77.6	213.9	43.6	7.3	-54.4	10.6	88.3	3.5	
化工	29.6	7.7	-27.0	17.3	23.5	51.9	31.4	30.9	26.1	17.4	78.4	35.4	
计算机	52.7	42.2	-19.7	18.6	51.1	35.0	26.6	-18.4	-6.3	9.6	-11.6	64.4	
机械设备	4.2	16.7	-41.9	30.7	31.3	27.1	61.3	16.2	-1.6	1.4	47.5	27.9	
汽车	25.7	1.6	-42.2	20.6	8.5	29.4	42.3	8.0	16.8	6.3	32.7	17.1	
非银金融	4.9	-7.9	-73.4	11.4	-1.5	50.3	88.5	-24.2	-64.3	-2.6	79.0	54.6	
传媒	20.1	1.2	-27.6	15.8	27.8	56.9	2.1	17.8	4.5	0.1	0.4	-12.4	
钢铁	4.5	-8.3	-6.3	5.1	4.7	9.1	9.2	1.5	10.4	7.0	37.1	20.2	
有色金属	11.2	-5.7	3.0	1.1	5.4	37.4	11.2	-18.0	-0.4	-5.0	34.4	15.1	
交通运输	-9.3	12.0	-30.4	2.9	-1.5	30.2	13.0	6.8	-0.8	-16.4	39.1	33.8	
轻工制造	10.8	3.9	-11.7	6.1	13.4	11.6	2.0	16.2	-6.1	-2.2	2.3	19.0	
银行	-12.9	-32.7	-10.4	20.6	-7.7	-29.1	-1.9	21.2	-25.7	40.2	60.9	35.6	
采掘	-5.5	-4.6	-3.7	4.6	-0.7	0.0	5.4	9.9	-4.1	-10.8	25.7	16.9	
通信	22.5	-18.7	-8.1	-0.2	19.1	9.4	28.2	11.8	-2.7	-11.0	-15.5	-2.2	
商业贸易	6.9	4.6	0.7	14.1	7.5	-1.3	12.3	-10.7	-3.0	5.2	-8.9	1.1	
房地产	12.5	-6.7	-13.6	19.2	1.2	16.4	14.8	13.3	-2.2	-15.9	19.2	-30.4	
农林牧渔	3.5	4.5	6.8	-1.7	-5.8	62.2	-2.6	-5.9	-14.6	-21.8	-9.9	5.9	
建筑装饰	17.5	15.0	5.4	6.5	-19.5	-5.3	11.1	4.7	5.2	-7.4	7.7	-26.0	
建筑材料	4.5	-24.6	-38.9	17.8	-5.8	5.7	25.5	2.7	-2.5	53.2	-22.1	-4.2	
综合	1.8	-9.5	-5.9	5.5	-0.5	-1.7	2.6	0.3	2.1	4.0	0.5	4.8	
国防军工	6.0	4.7	-5.8	8.8	7.7	5.5	8.7	-2.7	-14.9	-0.9	-15.2	-4.4	
纺织服装	1.5	-0.9	-1.4	0.4	0.8	-3.5	-0.8	0.1	0.9	-0.4	-2.2	1.2	
家用电器	31.3	-37.0	-59.7	31.0	44.3	22.9	-1.3	-12.5	-33.0	40.4	19.7	-62.6	
公用事业	1.9	-5.5	-9.3	-5.6	4.8	10.1	0.3	8.8	-13.6	-11.2	-6.4	4.4	
休闲服务	15.2	28.3	-15.8	-4.9	4.9	-10.8	-79.5	-21.7	-2.9	-14.1	30.1	1.7	
食品饮料	21.3	79.5	-150.6	135.5	28.0	-8.3	-160.1	-14.1	-128.3	14.7	-23.9	81.1	

数据来源：Wind、开源证券研究所

3、交易行为因子分析

因子分析主要从因子相关性、因子分组收益以及因子衡量指标对比等多个方面展开。测试日期从2017年1月开始，截止到2021年1月，共计49期数据，按照月频调仓，每月底选取因子值靠前的7个行业作为下个月的持仓组合。

3.1、因子相关性

下图展示了不同因子之间的秩相关性，历史偏好因子和定价权因子具有较高的关联度，二者无论是绝对值还是占比因子相关系数均达到0.8以上，而带有择时特征的边际变化因子，其与其他类因子整体相关性较低，普遍低于0.5。

图8：不同因子的秩相关性

秩相关性	持仓规模	持仓规模占比	成交金额	成交金额占比	净流入金额
持仓规模占比	0.85				
成交金额	0.89	0.66			
成交金额占比	0.76	0.86	0.68		
净流入金额	0.32	0.20	0.31	0.21	
净流入金额占比	0.19	0.21	0.17	0.22	0.93

数据来源：港交所、Wind、开源证券研究所

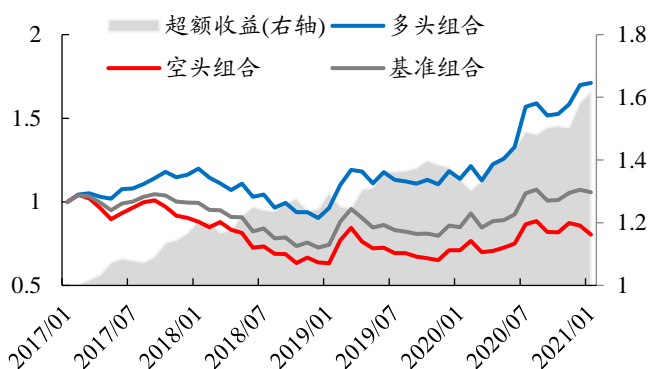
3.2、因子测试

3.2.1、历史偏好：持仓规模 vs. 持仓规模占比

持仓规模是北上资金对各个行业用脚投票的结果，持仓规模的大小除了受行业本身规模的影响之外，亦受到北上资金对行业本身偏好的影响。北上资金对某个行业持仓规模（占比）越高，通常表明其对该行业未来取得超额收益的能力越看好。

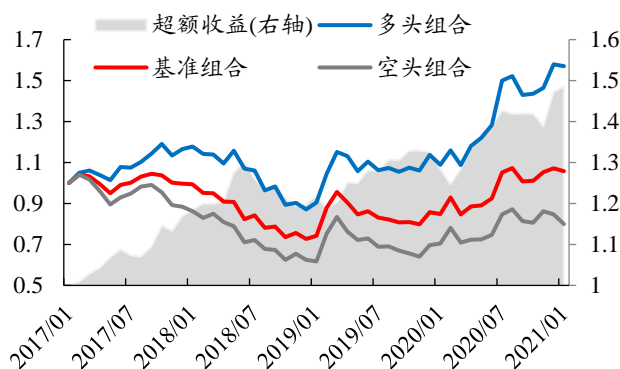
在每月月底，选取本月持仓规模（占比）最高的 7 个行业作为下一个月的多头组，持仓规模（占比）最低的 7 个行业作为空头组，28 个行业每月底再平衡作为基准组，结果如下：在回测期内，持仓规模因子和持仓规模占比因子多头累积净值分别达到 1.71 和 1.57，二者超额收益分别为 1.62 和 1.49。

图9：持仓规模多空收益



数据来源：Wind、开源证券研究所

图10：持仓规模占比多空收益

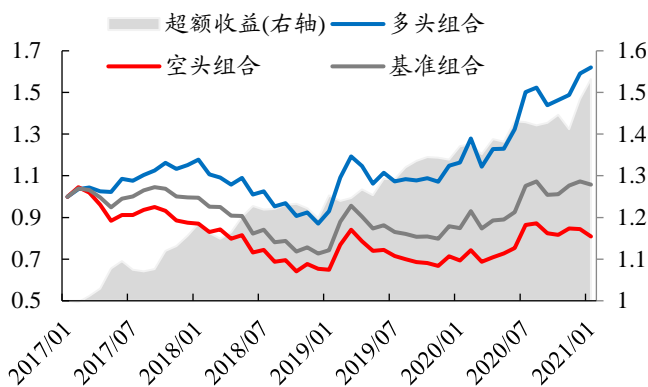


数据来源：Wind、开源证券研究所

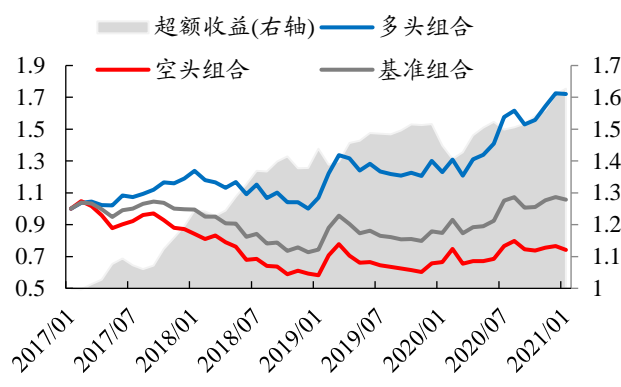
3.2.2、定价权：成交金额 vs. 成交金额占比

成交金额或成交金额占比因子刻画了北上资金对特定行业的定价权深度，一般而言，当北上资金在某个行业的成交金额（占比）越高，其对该行业的收益走势控制能力越强。其与历史偏好类因子潜在的区别在于，若北上资金中配置型资金持有了特定行业较大的规模，但由于其在该行业上的低换手特性，其在该行业上的成交金额并不会表现出与持仓规模成正比的特征。

在每月月底，选取本月成交金额（占比）最高的 7 个行业作为下一个月的多头组，成交金额（占比）最低的 7 个行业作为空头组，28 个行业每月底再平衡作为基准组，结果如下：在回测期内，成交金额因子和成交金额占比因子多头部分均大幅跑赢了基准组合，多头净值分别达到 1.62 和 1.72，但成交金额因子的超额收益相对更加稳定，回撤幅度更小。

图11: 成交金额分组收益


数据来源：港交所、Wind、开源证券研究所

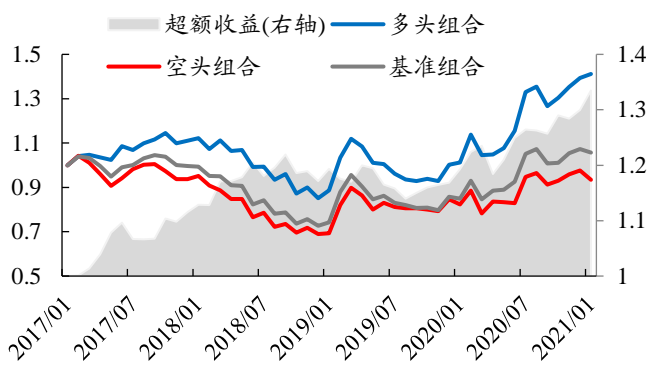
图12: 成交金额占比分组收益


数据来源：港交所、Wind、开源证券研究所

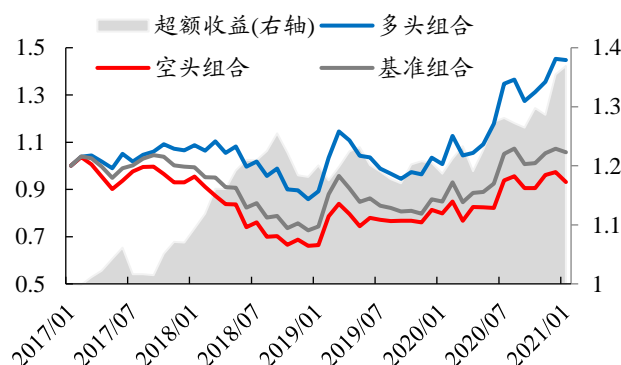
3.2.3、边际变化：净流入金额 vs. 净流入金额占比

北上资金在行业上的边际变化是当下偏好漂移的体现，其与持仓规模是两个不同维度的信息，分别代表着信息的增量部分和存量部分。净流入金额（占比）代表的边际改变，其内在推动力或许来源于某个行业景气度的改善，或许基于持续对某个行业投入的动量，亦或是特定行业的性价比凸显，无论是出于何种目的，净流入金额（占比）均能反映北上资金对该行业近期所持有的积极态度。

在每月月底，选取本月净流入金额（占比）最高的7个行业作为下一个月的多头组，成交金额（占比）最低的7个行业作为空头组，28个行业每月底再平衡作为基准组，结果如下：在回溯期内，净流入金额因子和净流入金额占比因子多头净值分别为1.41和1.45，相比定价权因子，表现稍显乏力，仔细观察可以发现边际变化类因子在2018年9月到2020年4月期间，超额收益整体处于横盘震荡状态。

图13: 净流入金额分组收益


数据来源：Wind、开源证券研究所

图14: 净流入金额占比分组收益


数据来源：Wind、开源证券研究所

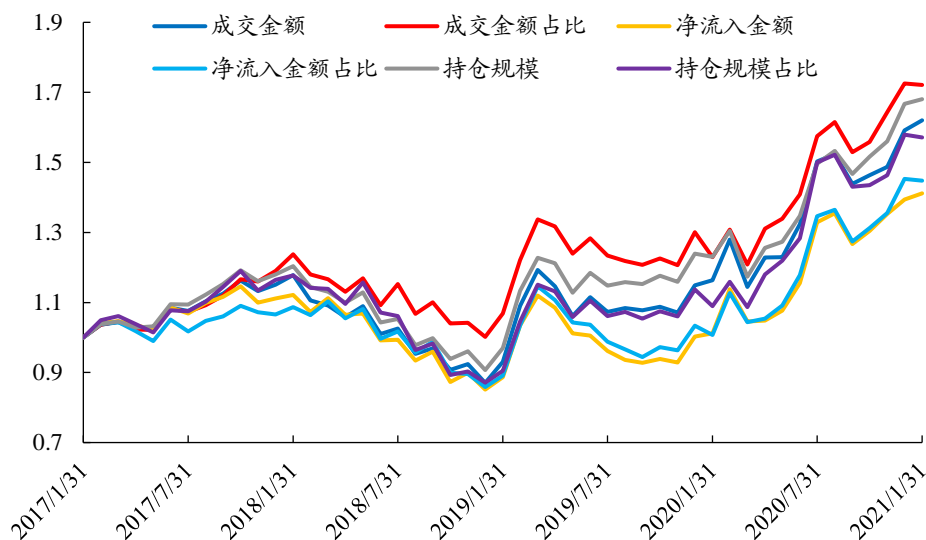
3.3、交易行为因子的多维对比

3.3.1、不同因子的多指标比较

通过对比不同策略的多头收益，可以发现历史偏好和定价权具有更佳的策略收益，而边际偏好次之。北上资金历史偏好是资金持续流入的结果，而资金本身又是推动股价变化的关键因素；随着北上资金的持续流入导致的持仓规模提升，持仓规模与成交金额的高相关性使得北上资金在部分行业的定价权随之上升，定价权反过

来又一定程度上能够决定行业的未来表现；边际变化表现相对中庸，可能在于边际变化更多是对当下市场行业偏好的纠偏和助推，本身不具有持续性，换句话说边际变化类似于择时，而择时是相对困难的一件事。

图15：不同因子的多头收益对比



数据来源：Wind、开源证券研究所

分指标来看，成交金额占比因子年化收益最高（14.22%），夏普比率最大（0.82），持仓规模占比因子胜率最高（66.67%），净流入金额因子整体表现最差，具有低收益、高波动、低胜率和高换手的特性。换手率方面，历史偏好类因子总体换手率最低（低于5%），表明北上资金的偏好具有持续性，其次为定价权类因子，每期换手率均值在20%以下，而代表边际变化的净流入因子，当回溯期为一个月时，每期换手达到50%左右的高水平。

表2：定价权类因子整体表现相对突出

因子名称	年化收益率	年化波动率	夏普比率	最大回撤	胜率	换手率
持仓规模	13.56%	17.73%	0.76	-24.65%	62.50%	1.19%
持仓规模占比	11.70%	19.44%	0.60	-26.79%	66.67%	4.46%
成交金额	12.54%	19.34%	0.65	-26.09%	64.58%	11.31%
成交金额占比	14.22%	17.42%	0.82	-19.05%	62.50%	15.77%
净流入金额	8.81%	19.09%	0.46	-25.72%	58.33%	49.40%
净流入金额占比	9.49%	19.22%	0.49	-22.18%	60.42%	56.85%

数据来源：Wind、开源证券研究所

3.3.2、收益来源再探索

通过回溯北上资金在不同因子层面的持仓，可以发现：北上资金在不同行业之间存在明显的偏好差异。从历史偏好角度，食品饮料、医药生物、家用电器、电子和休闲服务出现次数较高；从定价权角度，食品饮料、医药生物、银行及家用电器出现次数较高；从边际变化角度，家用电器、医药生物和食品饮料位列出现频率前三。综合而言，食品饮料、家用电器、医药生物、电子和银行是北上资金相对青睐的行业，而这些高频出现的行业亦是2017年以来市场上表现相对亮眼的行业。

表3: 不同因子持仓行业次数统计

行业名称	持仓规模	持仓规模占比	成交金额	成交金额占比	净流入金额	净流入金额占比
食品饮料	49	49	49	48	27	22
非银金融	49	3	49	17	29	11
医药生物	49	39	49	26	35	27
电子	45	16	46	2	27	14
银行	49	12	37	49	28	16
家用电器	49	49	28	49	32	32
计算机	5	0	23	3	16	12
化工	0	0	20	0	19	4
电气设备	7	2	12	0	17	15
汽车	14	31	11	10	11	11
有色金属	0	0	7	0	3	5
房地产	0	0	7	14	11	11
交通运输	27	48	2	33	12	17
农林牧渔	0	0	1	1	6	12
传媒	0	0	1	0	6	7
机械设备	0	0	1	0	13	6
综合	0	6	0	0	1	8
建筑材料	0	36	0	22	11	21
建筑装饰	0	0	0	0	4	6
休闲服务	0	49	0	41	5	22
国防军工	0	0	0	0	1	2
采掘	0	0	0	0	3	10
公用事业	0	2	0	1	8	8
通信	0	0	0	0	6	9
轻工制造	0	0	0	0	3	9
纺织服装	0	0	0	0	2	2
钢铁	0	1	0	27	4	19
商业贸易	0	0	0	0	3	5

数据来源: 港交所、Wind、开源证券研究所

3.3.3、参数敏感性测试

以上因子在构建时为了避免随机扰动,对数值进行了平滑处理,平滑区间选取了常规的一个月,对于存量型因子(历史偏好和定价权)这种方法并没有特别明显的瑕疵,但是对于增量型因子(边际变化),其回溯时长会是比较敏感的参数。下图结果亦证实了这个猜想,净流入金额(占比)因子在回溯时长为3个月时,策略年化收益率最高,夏普比率最大,相比于选取回溯一个月效果改进明显。

图16: 不同回溯时长下的因子表现对比

	年化收益率					夏普比率				
	1M	3M	6M	9M	12M	1M	3M	6M	9M	12M
持仓规模	13.6%	12.8%	12.5%	11.9%	11.6%	0.76	0.72	0.69	0.64	0.59
持仓规模占比	11.7%	11.6%	11.7%	11.5%	12.2%	0.60	0.59	0.58	0.55	0.57
成交金额	12.5%	12.7%	12.4%	11.4%	11.9%	0.65	0.66	0.63	0.57	0.58
成交金额占比	14.2%	13.4%	12.5%	11.6%	11.9%	0.82	0.74	0.67	0.60	0.60
净流入金额	8.8%	12.5%	10.2%	8.5%	10.0%	0.46	0.68	0.54	0.45	0.49
净流入金额占比	9.5%	12.7%	12.2%	9.6%	9.1%	0.49	0.68	0.62	0.51	0.47

数据来源: 港交所、Wind、开源证券研究所

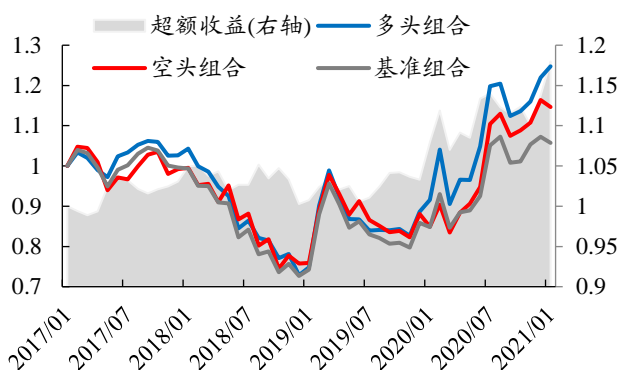
3.4、因子合成

在前文, 我们讨论了不同交易行为因子的表现, 从结果来看, 历史偏好和定价类因子整体表现稳定, 边际变化因子对平滑时长较敏感, 当回溯时长选取为三个月时, 历史表现大幅改善。值得注意的是, 前文我们在做因子相关性检测时, 发现不同交易行为因子之间本身是有一定的关联度, 尤其是历史偏好类因子和定价权类因子, 相关系数达到 0.8, 为此我们尝试在不同因子之间进行截面回归以消除彼此的潜在影响。下文根据参数敏感性测试的结果, 选取最优参数进行验证。为了简便, 仅选取每类因子中的一个进行测试说明。

当使用成交金额对持仓规模进行截面回归得到成交金额残差因子时, 其超额收益能力下降明显, 说明成交金额本身的超额收益来源与持仓规模高度相关。但是细看超额收益曲线, 其在 2020 年以后开始有一定的向上趋势, 表明代表定价权的成交金额因子开始包含了持仓规模所不涵盖的信息量。

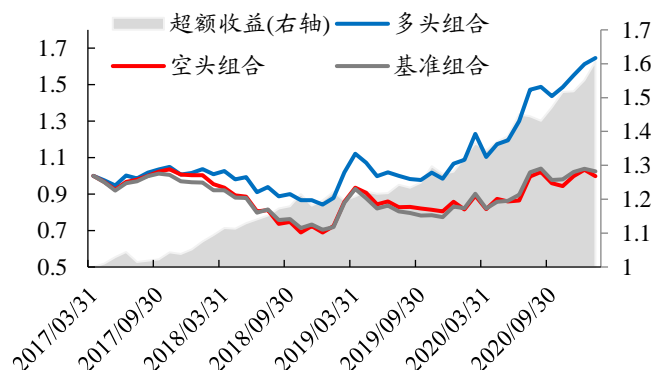
值得注意的是, 3 个月净流入金额因子在截面上对持仓规模因子进行回归后的残差因子, 表现出现改善。说明净流入因子是区别于持仓规模因子的增量信息点, 这与前文我们提到持仓规模代表存量信息, 净流入金额代表增量信息的说明亦相符合。净流入金额在截面上对持仓规模进行回归, 本质上是调整净流入金额因子在不同行业上持仓规模的偏差。

图17: 成交金额回归持仓规模的残差表现



数据来源: 港交所、Wind、开源证券研究所

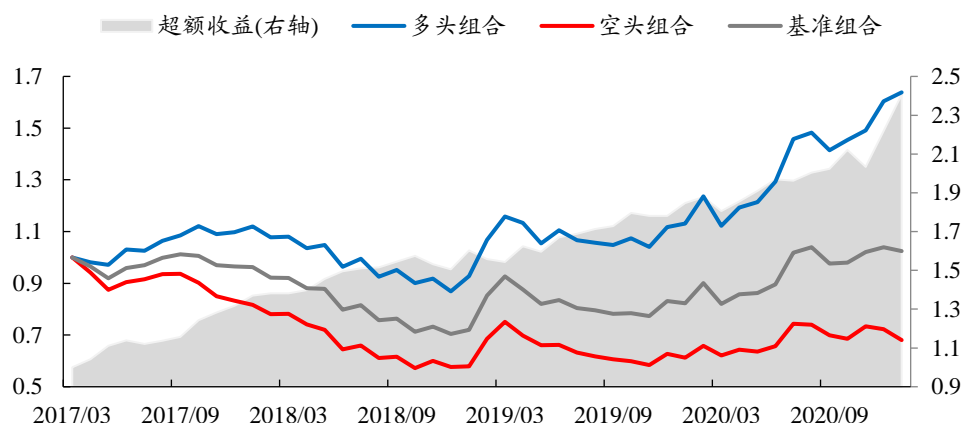
图18: 净流入金额回归持仓规模的残差表现



数据来源: Wind、开源证券研究所

基于净流入金额截面回归持仓规模后得到的残差因子表现稳定，我们将其与持仓规模因子进行合成，得到一个既包括存量信息亦考虑增量信息的新因子。每期分别对持仓规模和残差因子进行截面标准化，然后将标准化后的数值进行等权相加得到合成因子值，最后选择分值最高的 7 个行业作为持仓多头，分值最低的 7 个行业作为持仓空头。纳入双维度信息的合成因子多空策略在夏普比率、最大回撤和胜率方面都有一定的提升。其中，合成因子的夏普比率从存量因子的 2.1 上升到 2.4，最大回撤相比增量因子从 -6% 降低到 -4.4%，胜率无论是相对增量因子还是存量因子，均提升明显，达 74%。

图19：合成因子多空收益率曲线夏普比率更高



	因子名称	年化收益率	年化波动率	夏普比率	最大回撤	胜率
多空对冲	存量因子	22.9%	10.8%	2.1	-4.3%	67.4%
	增量因子	13.6%	9.4%	1.4	-6.0%	65.2%
	合成因子	25.1%	10.5%	2.4	-4.4%	73.9%
多头基准对冲	存量因子	12.2%	6.8%	1.8	-4.6%	65.2%
	增量因子	12.9%	5.9%	2.2	-2.9%	76.1%
	合成因子	12.7%	6.2%	2.1	-3.2%	71.7%

数据来源：Wind、开源证券研究所

行文至此，概括地讲，北上资金在行业轮动上超额收益的主要来源是：以持仓规模为代表的历史偏好类因子所蕴含的存量信息，和以净流入金额为代表的边际变化因子所蕴含的增量信息。而存量因子和增量因子的联合又能够相互取长补短，提供更高夏普比率的多空对冲曲线，二者构成了北上资金行业配置的双轮驱动力。

4、风险提示

模型测试基于历史数据，市场未来可能发生变化。测量过程存在误差。

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的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