

证券研究报告



分析师:

任瞳

[rentong@xyzq.com.cn](mailto:rentong@xyzq.com.cn)

S0190511080001

徐寅

[xuyinsh@xyzq.com.cn](mailto:xuyinsh@xyzq.com.cn)

S0190514070004

研究助理:

郑兆磊

[zhengzhaolei@xyzq.com.cn](mailto:zhengzhaolei@xyzq.com.cn)

## 雪球知股乎系列二：百万投资组合里的宝贝

2017年06月19日

### 报告关键点

### 相关报告

《雪球知股乎系列之一：和关注度因子有个约定》 2017-05-08

### 投资要点

- 作为“猎金系列之十五”，我们继续深耕雪球网的文本信息价值，研究用户创建投资组合带来的选股信息增量。
- 从抓取的结果来看，时间窗为2014年10月16日至2017年3月20日(抓取截至日)，雪球网共计有443,043位用户创建了989,886个投资组合。
- 通过投资组合包含的股票仓位求出相应的价值，构建月度价值变动因子。对投资组合进行限制，只考虑截面上包含10只股票以上的组合；同时从月度价值变动因子中剥离价格变动产生的影响，构建有效价值变动因子。
- IC以及分位数组合测试验证了因子选股的有效性。研究其和传统的价值、成长、反转、分析师情绪以及《雪球系列一》中的关注度因子五类因子之间的相关性，结果显示和上述因子相关性较低。且Fama-MacBeth回归分析也表明价值变动因子可以带来显著的信息增量。
- 基于价值变动因子构建选股策略。以中证500作为对冲基准，策略年化超额收益为22.25%，夏普比率为3.27，最大回撤为6.25%。



## 目 录

1、雪球网数据维度概述.....	- 3 -
1.1、雪球平台数据优势.....	- 3 -
1.2、雪球网“创建投资组合”功能介绍.....	- 3 -
1.3、创建投资组合在基金构建中的应用.....	- 5 -
1.4、创建投资组合数据抓取.....	- 6 -
2、用户创建投资组合数据分析.....	- 7 -
2.1、投资组合数目时间维度分析.....	- 7 -
2.2、投资组合对 A 股覆盖度.....	- 9 -
3、基于投资组合构建选股因子.....	- 9 -
3.1、仓位量变动的回填期限确定.....	- 10 -
3.2、组合包含的股票数目研究及限定.....	- 11 -
3.3、价值变动因子构建流程.....	- 13 -
3.4、价值变动因子有效性测试.....	- 14 -
3.5、与传统选股因子对比分析，不一样的烟火.....	- 15 -
4、基于价值变动因子的选股策略构建.....	- 16 -
4.1、价值变动因子选股策略构建.....	- 16 -
5、结语.....	- 17 -
图表 1、用户创建投资组合过程.....	- 4 -
图表 2、用户调仓细节.....	- 4 -
图表 3、用户调仓行为的结果.....	- 5 -
图表 4、用户创建投资组合的最终界面.....	- 5 -
图表 5、基于雪球网数据产生的大数据基金.....	- 6 -
图表 6、用户界面下的投资组合入口.....	- 6 -
图表 7、用户调仓行为的后台展示.....	- 7 -
图表 8、调仓行为对应的数据明细.....	- 7 -
图表 9、雪球网用户创建投资组合分布.....	- 8 -
图表 10、用户创建投资组合&取消投资组合时间维度分布.....	- 8 -
图表 11、用户调仓行为的结果占比.....	- 9 -
图表 12、投资组合对 A 股覆盖度分析.....	- 9 -
图表 13、调仓行为天数间隔的 T 统计值.....	- 10 -
图表 14、调仓行为天数间隔分布.....	- 11 -
图表 15、包含不同股票数的投资组合占比分析.....	- 11 -
图表 16、组合包含的股票数目相应统计值.....	- 12 -
图表 17、不同类别的组合包含的股票数目加总统计.....	- 12 -
图表 18、月度价值变动因子的 IC 检验.....	- 14 -
图表 19、行业中性化以后的月度价值变动因子的 IC 检验.....	- 14 -
图表 20、月度价值变动因子的 IC 趋势图.....	- 14 -
图表 21、月度价值变动因子分位数组合测试结果.....	- 15 -
图表 22、月度价值变动因子分位数测试结果.....	- 15 -
图表 23、价值变动因子与传统因子秩相关性.....	- 16 -
图表 24、Fama-Beth 回归分析.....	- 16 -
图表 25、价值变动因子策略选股策略表现.....	- 17 -
图表 26、价值变动因子策略在不同年度的表现.....	- 17 -
图表 27、改进价值选股因子的对冲净值曲线.....	- 17 -

## 1、雪球网数据维度概述

### 1.1、雪球平台数据优势

现有的量化选股因子和体系已经被过度挖掘和开发，虽然不断的推陈出新，但推陈出新的难度不断加大。更令人担忧的是，由于大家的研究方法和数据等同质性非常高，风险相对集中。今年以来的市场表现更是印证了这一点，小盘股效应不见，大盘效应敲打着投资者的神经和耐性，日复一日。那么如果有更多的量化投资视角、有更全面的量化数据源，或许能让模型考虑更多的风险和因素，从而变得更加强壮，进而可以有效的应对市场风险。

诚如你我所见、耳闻：我们不知不觉的发现精准定位出现在生活的每一个角落，你我的每一次行为；个人的数据性、功能性概念在不断放大，私密性、自我性在逐渐变得奢侈。这就是大数据时代的直接体现，这就是人工智能所带来的震撼。而随着大数据和人工智能不断的成熟和普及，在量化界开花结果是一种必然。我们在《雪球知股乎系列一》中详细介绍了互联网文本挖掘在量化领域的应用。目前来看，分析师研报、财经新闻、贴吧的用户帖子成为量化投资领域新的数据源，而贴吧帖子实际上正是 UGC (User Generated Content, 用户原创内容) 的最佳体现，专业研究人员、财经媒体、众多散户的观点可以通过这些素材得到淋漓尽致的体现。

在众多互联网平台中，我们选择雪球网作为我们进入这个领域的敲门砖。从最早专门讨论美股投资的社区“i 美股”算起，雪球的整体运行时间已接近 6 年。在创始人方三文的一次采访中，他提及：有超过 600 万人注册了这个中国最大的投资交流平台，每天活跃着 100 万用户。除了量级大以外，这个平台本身的 UGC 内容也有较高的质量，你会不经意发现，身边的一些大咖有可能就是雪球网的用户。他们会在上面阐述自己的投资心得、分享自己的观点。

因此，接下来，我们仍然会基于雪球网已经抓取的数据，深度挖掘有效信息，构建选股、择时策略，最大化雪球网的投资价值。

### 1.2、雪球网“创建投资组合”功能介绍

在《雪球知股系列一》中，我们对抓取的方法、抓取的结果进行了详细的阐述，这些结果也吻合雪球网创始人方总在采访中提到的数据量级。我们知道雪球网从对象上来看，有用户、股票/其他标的、投资组合、帖子这样四大维度，而用户和股票是两大主线，用户的基本行为有“发帖”、“关注”、“创建投资组合”三类基本动作。根据对象不同，用户基本行为会得到不同的展示。这里，我们对整体抓取以及抓取结果不做赘述，开始聚焦本系列的研究重点：用户创建投资组合。

“创建投资组合”功能是雪球网在2014年底推出的新功能。本质上来讲，是为了让管理投资组合这种行为大众化。进入雪球网首页→个人账户→创建新组合，进入到创建投资组合的界面。然后选择“组合名字”→“市场范围”→“添加成分股”。这样就能在雪球网上创建一个投资组合。

图表 1、用户创建投资组合过程



资料来源：雪球网，兴业证券研究所

选择完毕之后，可以通过拖动图表-2中的仓位按钮实现仓位控制。每个投资组合的初始总净值为1。并且每一个被选中的股票仓位都必须大于0，否则创建组合失败。

图表 2、用户调仓细节



资料来源：雪球网，兴业证券研究所

那么每次调仓行为发出以后，会有四种状态：调仓成功、失败、取消调仓行为、等待行为落地（后面我们会给出这几种行为占比）。

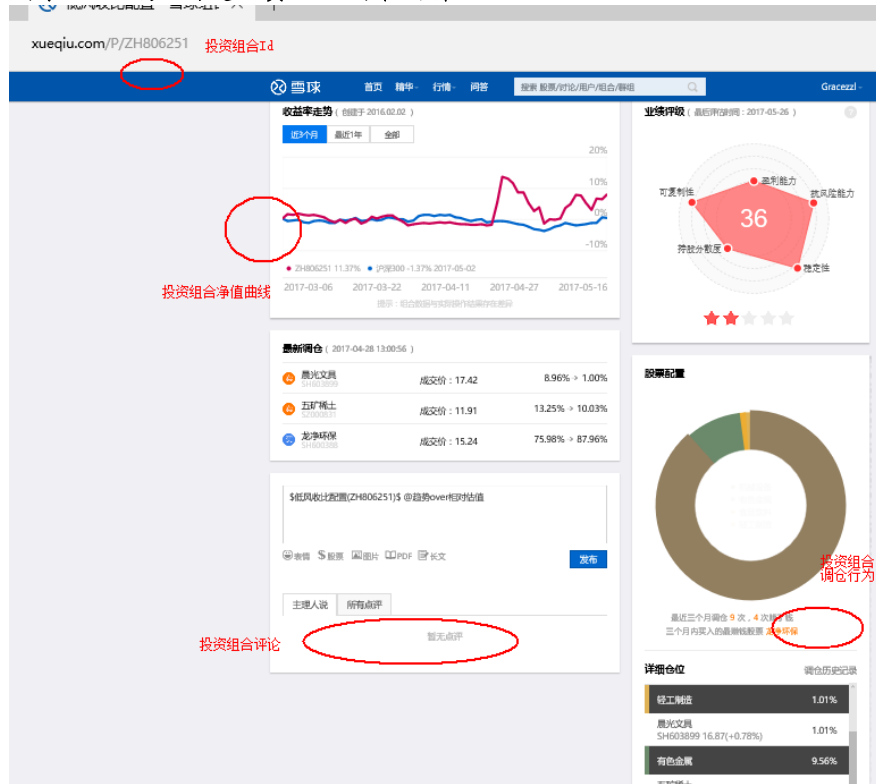
图表 3、用户调仓行为的结果

调仓状态	释义
Cancel	取消调仓行为
Fail	调仓行为失败
Success	调仓行为成功
Pending	等待调仓行为落地（发生在闭市以后）

资料来源：雪球网，兴业证券研究所

当我们完成投资组合的创建并根据自己的投资倾向做出调仓之后，这个投资组合就会生成一个单独的界面。在这个界面上可以观察到这个投资组合的收益、调仓的行为、以及其他雪球用户针对于该投资组合做的评论(或者自己的评论)。当然用户也可以取消自己创建的投资组合，一旦取消投资组合之后，在该界面上会有相应的标识。

图表 4、用户创建投资组合的最终界面



资料来源：雪球网，兴业证券研究所

而投资组合的收益净值变动以及用户的调仓行为正是我们抓取的对象，也是本报告研究的重点。

### 1.3、创建投资组合在基金构建中的应用

在《雪球知股系列一》中，我们介绍了基于雪球网平台的一些产品，都不同程度的利用了创建投资组合的功能。总体来看，主要利用了组合的活跃性以及包含股票的冷热程度两个维度。

图表 5、基于雪球网数据产生的大数据基金

名称	代码	合作公司	上市时间	对创建投资组合功能的利用
雪球 100	CSI:930759	天弘基金	2015/08/21	选取近一个月仓位变动的组合作为参考；并且股票被至少 20 人所持有
雪球精选	CSI:930762	大成基金	2015/09/15	-
雪球智选	CSI:930760	博时基金	2015/09/21	选取近一个月有仓位变动的组合

资料来源：雪球网，兴业证券研究所

### 1.4、创建投资组合数据抓取

根据前面雪球网创建投资组合界面介绍，我们构建相应界面（图表-4）的抓取方法。首先，通过特定的方法获取所有有过活跃行为的用户 ID（这一步详见《雪球系列一》），进一步在用户界面下获取投资组合的 ID（参见图表-5），然后解析每个投资组合的界面（参见图表-4），主要包括投资组合历史净值以及调仓行为的明细，而调仓行为明细是接下来的研究重点（图表-7）。利用 scrapy 完成抓取工作。我们以 ZH184251 为例，通过解析 ZH184251 这个投资组合的调仓行为明细，能得到网页端对应的数据，这里包含详细的调仓对象、时点、仓位、调仓行为是否成功、仓位价值等细节。当然，实际上底层数据端口还包含每次调仓欲达到的仓位量以及上一次调仓欲达到的仓位量，但这不是本次研究的重点，就不予以赘述。

图表 6、用户界面下的投资组合入口



组合	最新净值	总收益	操作
业绩成长	1.8063	80.63%	+ <
成长价值股投资	1.3204	32.04%	+ <
蓝筹股	1.3783	37.83%	+ <
打新组合	1.0703	7.03%	+ <

资料来源：雪球网，兴业证券研究所

图表 7、用户调仓行为的后台展示

```

安全 | https://xueqiu.com/cubes/rebalancing/history.json?cube_symbol=ZH184251&count=20&page=
应用 兴业证券 个人 数据库 量化投资 爬虫

{"count":20,"page":1,"totalCount":31,"list":
[{"id":26733965,"status":"success","cube_id":184153,"prev_bebalancing_id":18363813,
"all","created_at":1493256684095,"updated_at":1493256684095,"cash_value":7.8E-
7,"cash":0.0,"error_code":null,"error_message":null,"error_status":null,"holdings"
[{"id":148736958,"rebalancing_id":26733965,"stock_id":1025218,"stock_name":"中科创
达","stock_symbol":"SZ300496","volume":0.01688971,"price":29.63,"net_value":0.5004
00.0,"prev_target_weight":100.0,"prev_weight_adjusted":99.76,"prev_volume":0.01684
oactive":true,"created_at":1493256684095,"updated_at":1493256684095,"target_volume
nt":null,"diff":0.0,"new_buy_count":0},
{"id":18363813,"status":"success","cube_id":184153,"prev_bebalancing_id":14501837,
11","created_at":1466127046170,"updated_at":1466127046170,"cash_value":7.8E-
7,"cash":0.0,"error_code":null,"error_message":null,"error_status":null,"holdings"
[{"id":97774049,"rebalancing_id":18363813,"stock_id":1025218,"stock_name":"中科创
达","stock_symbol":"SZ300496","volume":0.01684989,"price":75.0,"net_value":1.2637,
0.0,"prev_target_weight":100.0,"prev_weight_adjusted":25.54,"prev_volume":0.004303
oactive":true,"created_at":1466127046170,"updated_at":1466127046170,"target_volume
nt":null,"diff":0.0,"new_buy_count":0},
资料来源：雪球网，兴业证券研究所
    
```

前面提及，每个投资组合的初始净值为 1。买入的占比作为资金量，而量纲化时点价格之后得到相应标的的仓位量。对这个数据结构进行整理，我们列举样本的部分关键指标。从中可以发现，主要包含三大块：基本信息（包括组合 Id、调仓日期、调仓对象、调仓是否成功等）、此次调仓对应的标的的净值（调仓前的净值、调仓后的净值、此次调仓欲达到的目标）、此次调仓对应的仓位量信息（具体和净值结构一致）。注意，一旦股票发生高送转，那么仓位量会相应的出现变动。

图表 8、调仓行为对应的数据明细

	变量	变量释义	样本 1	样本 2
基本信息	port_id	组合 id	ZH121557	ZH121557
	action_id	调仓行为 id	24616489	24593227
	action_dt	调仓日期	2017-1-19 9:45	2017-1-18 9:58
	stock_name	股票	中金黄金	中国铝业
	stock_symbol	股票代码	SH600489	SH601600
	status	调仓行为的结果	success	success
此次调仓 对应标的 净值	prev_weight(%)	调仓前的权重	39.91	6.33
	target_weight(%)	此次调仓欲达到的权重	8	0
	Weight(%)	调仓之后的权重	8	0
此次调仓 对应标的 仓位量大 小	prev_volume	调仓前的仓位量	5.8%	2.4%
	target_volume	调仓欲达到的仓位量	1.2%	0.0%
	volume	仓位量	1.2%	0.0%

资料来源：雪球网，兴业证券研究所

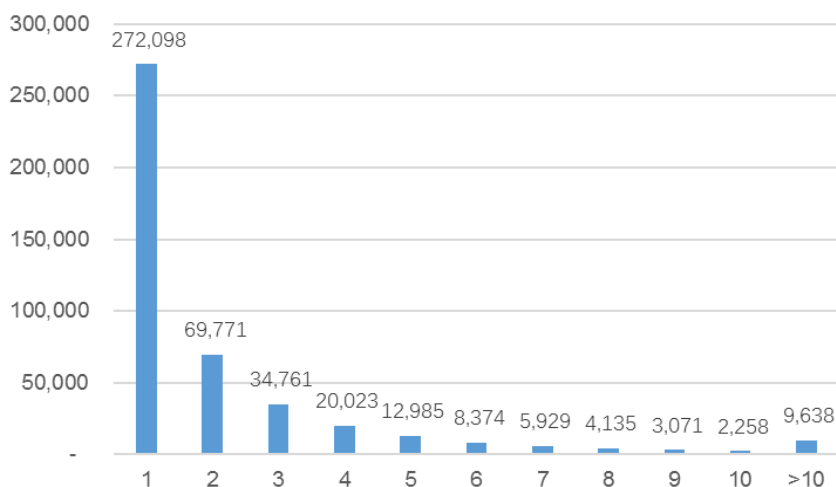
## 2、用户创建投资组合数据分析

### 2.1、投资组合数目时间维度分析

抓取的时间截点为 2017 年 3 月 20 日。截至 17 年 3 月 20 日，有 443,043 位用户创建了 989,886 个投资组合，人均创建 2.23 个组合，创建组合个数为 1 的用

户占比 61.4%。创建投资组合的日期自 2014 年 10 月份开始，而这与雪球网开启该功能较晚有关。

**图表 9、雪球网用户创建投资组合分布**

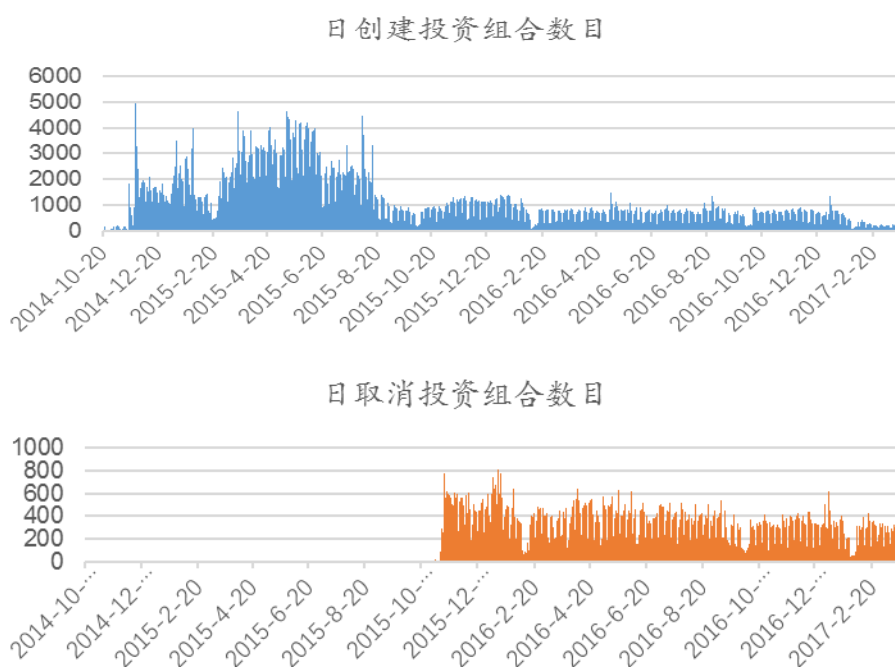


注：图表的意义是：曾经创建过N个投资组合的用户数分布，如横坐标为2表示：创建过2个投资组合的用户有69,771个；

资料来源：雪球网，兴业证券研究所

从用户创建&取消投资组合的时间维度分布来看，15年年中创建投资组合达到高峰，而后来爆减并趋于稳定。目前来看，日创建投资组合在200左右。而从取消投资组合的维度来看，自2015年10月开始（该日是雪球网开始推出取消功能的大致时间点）取消行为进入顶峰期，目前雪球网日取消投资组合的数目在300左右。截止至2017年3月20日，共计产生了98.9万投资组合，而关闭的投资组合数目达到16.6万，占比16.6%。

**图表 10、用户创建投资组合&取消投资组合时间维度分布**





资料来源：雪球网，兴业证券研究所

## 2.2、投资组合对 A 股覆盖度

前面提及，调仓结果共计有 4 种，“成功”、“失败”、“取消”、“等待”。而根据雪球网的功能设计机制，实际上除了状态为“success”的之外，其他功能绝大部分发生时间都在闭市之后。而从抓取的实际结果来看，状态为“success”占比高达 86.2%。因此，我们后来选股因子的构建将基于状态为调仓成功的行为结果。

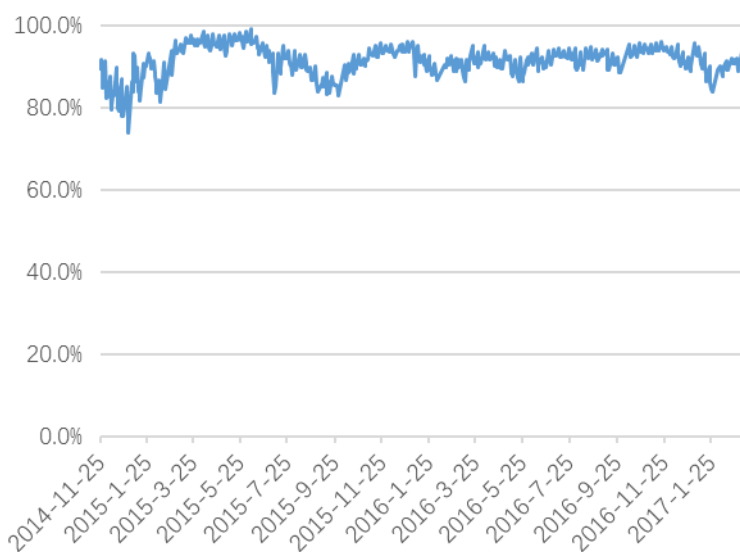
**图表 11、用户调仓行为的结果占比**

调仓状态	释义	行为数目	比例
Cancel	取消调仓行为	3,525,509	9.7%
Fail	调仓行为失败	1,471,400	4.1%
Success	调仓行为成功	31,265,420	86.2%
Pending	等待调仓行为落地	6,585	0.0%

资料来源：雪球网，兴业证券研究所

统计创建投资组合对 A 股覆盖度的具体情况。从结果来看，自 2014 年 10 月至 2017 年 3 月 17 日，投资组合日调仓行为包含的股票对 A 股的覆盖度非常高，均值达到 91.3%，且稳定性非常高。而从月度角度来看，自 2015 年初开始，月度覆盖度为 100%。

**图表 12、投资组合对 A 股覆盖度分析**



资料来源：雪球网，兴业证券研究所

## 3、基于投资组合构建选股因子

每个雪球网用户创建的投资组合实际上就是一个投资产品。而目前摆在我们面前的是有近 100 万个基金经理管理的投资组合，只不过这些基金经理是来自五湖四海，来自大众。而这正是互联网文本挖掘所需要的，也正是雪球网所想致力

于实现的目标。扎实的用户基础、较高的用户质量保证了这部分数据的有效性。因此，我们利用创建的投资组合的仓位价值变动作为有效信息，将所有的投资组合重构成一个大的投资组合，在重构的过程之中着重考虑组合的有效性，删除关闭的投资组合，进而构建选股因子。

市场本身是一场零和博弈。众多专业投资者想要通过利用自己的专业投资能力获取超额收益，但结果往往事与愿违。部分投资者可能实现了这一目标，但是只是短暂的。想要长期实现超额正收益、创造巴菲特式奇迹实属不易。市场的投资者总有一种“后知后觉”的现象。当市场对某只股票的态度保持一种不变的状态的时候，市场的机会可能就此来临。而我们拥有 100 万左右的投资组合，那么这些投资组合的价值变动大致上可以代替、反应这种趋势和规律。所以，我们构建价值变动因子，当因子变化幅度小的时候，就是价值来临的时刻；而该因子变动较大往往意味着机会的消失。

### 3.1、仓位量变动的回填期限确定

我们希望通过用户的调仓行为得到仓位量以及仓位价值的具体分布。根据雪球网抓取的数据特征，我们得到的是调仓行为相应时间截面上的数据，这个数据维度不是连续性的，比如张三的某个组合在 2017 年 1 月 1 日发生了调仓（下次调仓是 17 年 3 月 1 日），调仓标的是 SZ000001。那么根据抓取的数据结果，我们得到这一天的 SZ000001 具体仓位量，但 1 月 2 日至 3 月 1 日之前是没有记录的。所以我们需要将这个变化导致的最终仓位量回填到下一次调仓行为发生之前。但是有些用户可能由于个人原因（比如不太活跃、甚至直接放弃组合，但是没有将其下线而已），导致两次调仓行为时间间隔特别久远，我们需要确定一个适合的回填期限。回填期限的确定需要从用户调仓的活跃性以及价值投资的折中性两个角度考虑。如果周期过短，那么实际上是否认了价值投资的意义。如果周期过长，那么实际上是没有充分考虑用户换仓的这种行为本身的意义。

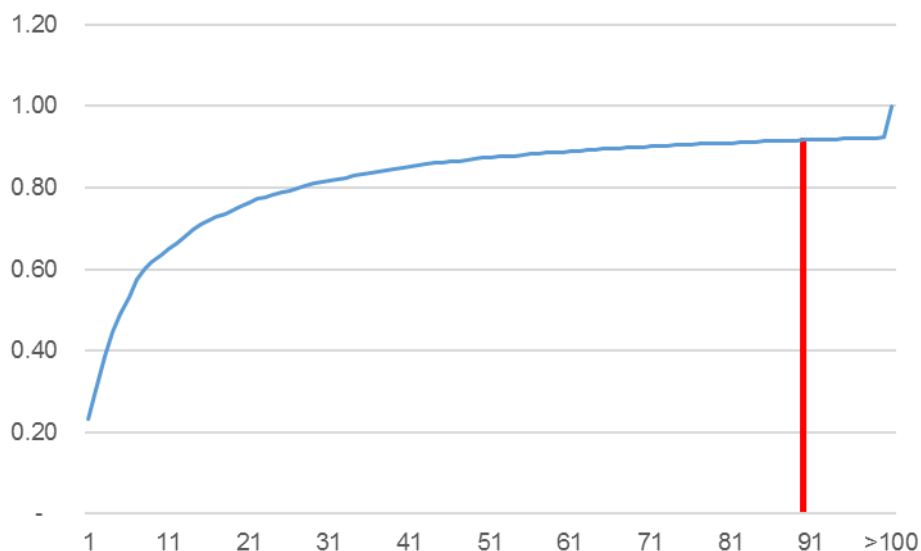
我们研究了用户调仓的时间间隔（下次调仓与本次调仓的时间间隔），结果发现，用户调仓行为的时间间隔平均为 28.7 天，75%中位数为 20 天，而 90%的调仓间隔发生在 90 天以内。这说明用户的调仓行为还是较为活跃的。而另一方面说明，如果我们将回填周期 N 设置相对较长并不会因为用户行为产生较大的影响。所以这里为了更好的体现价值投资的理念，我们设置回填周期 N 为 3 个月。

**图表 13、调仓行为天数间隔的 T 统计值**

	25%中位数	50%中位数	75%中位数	均值
相应值	1 天	6 天	20 天	28.7 天

资料来源：雪球网，兴业证券研究所

图表 14、调仓行为天数间隔分布



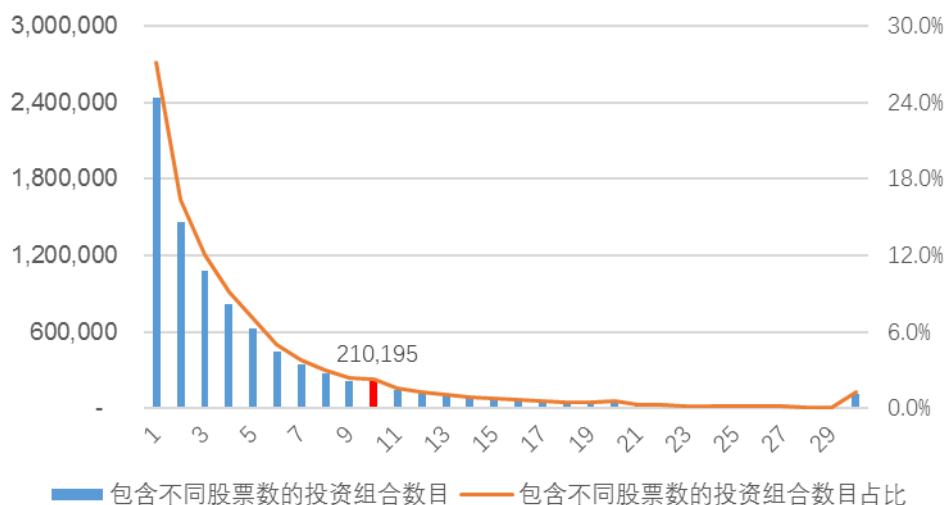
资料来源：雪球网，兴业证券研究所

### 3.2、组合包含的股票数目研究及限定

不同组合包含的股票数目可能不同，我们统计组合包含的股票数目。这里需要注意，我们统计的是截面数据，即以月度为窗口看所有月份月末仍然活跃的投资组合包含的股票数目，并对这个数目进行加总。以 ZH121557 这个投资组合为例，假设其寿命有 2 个月。那么对于这个投资组合来说，将包含 2 个截面样本。之所以选择月度的时间窗与我们后面因子构建时间周期有关（我们构建的是月度因子）。

而从统计结果来看，组合包含股票数目的均值为 5.18，50%中位数来看，为 5，组合包含的股票数目具有一定得分散性。

图表 15、包含不同股票数的投资组合占比分析



资料来源：雪球网，兴业证券研究所

图表 16、组合包含的股票数目相应统计值

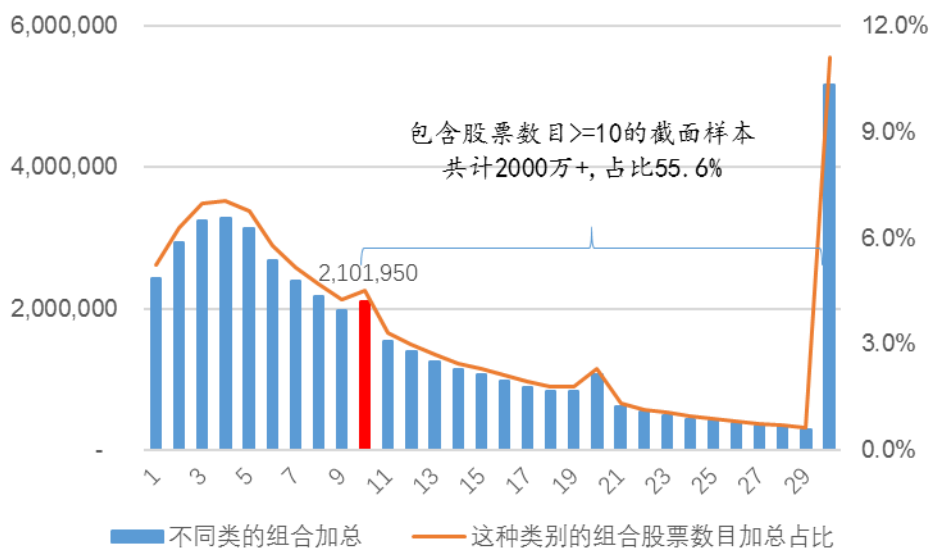
	25%中位数	50%中位数	75%中位数	均值
相应值	1	3	6	5.18

资料来源：雪球网，兴业证券研究所

作为互联网平台，实际上有些用户的行为是尝试性的。有时候用户可能仅仅是出于好奇心对功能进行了尝试，这样该组合包含的股票数目就特别少。这部分用户的投资组合在相应时点截面上市没有意义的。那么为了保证投资组合这种行为本身的意义，我们需要对截面上组合包含的股票数目进行限定。具体做法：在横截面上观察每个有效组合包含的股票数目，如果股票数目小于 N，那么我们认为该截面组合是无效的。这里我们确定  $N=10$ 。这意味着如果横截面上组合包含的股票数目  $\leq 10$ ，那么在该时点不考虑这个组合。

可能读者会疑问，从图表-15 的统计结果来看， $N=10$  意味着很多截面上的组合将无效。但是从另外一个角度来看，我们最终是看截面上选择的股票数目，而不是看截面的组合数（注意这点区别非常重要）。那么，我们将组合包含的股票数目作为研究对象，看看这种对象下总的包含的股票数目。在图表-15 中，当  $x=10$  时  $y$  值为 210,195，代表着有 210,915 截面组合是包含了 10 个股票。而在图表-17 中当  $x=10$  时，对应  $y$  值为 2,101,950，表示 210,195 个截面组合包含的股票样本数，而这才是我们选股因子关注的最终对象（图表-17 中  $y$  轴）。而当我们把  $N$  设置为 10 的时候，我们共计得到 2,000+万的样本数，占截面股票样本量的 55.6%。巨大的样本量完全支持我们去构建选股因子。

图表 17、不同类别的组合包含的股票数目加总统计



资料来源：雪球网，兴业证券研究所

### 3.3、价值变动因子构建流程

根据前面的数据结构，我们构建时间周期为 T 的价值变动因子，因子构建的数据准备如下：

- 1、针对于某个投资组合，T 时刻发生调仓后，对仓位量进行回填，回填期限为  $\min[\text{下次调仓日}-\text{当前调仓日}+1 \text{ 天}, 90 \text{ 天}]$ ，一旦用户调仓行为时间间隔超过 3 个月，那么 [T+3 个月，下次调仓行为发生时刻] 期间该组合的标的无效。
- 2、如果在每个月末，组合包含的股票数目  $\leq 10$ ，那么本月的组合无效。
- 3、我们前面提及，有将近 16% 的组合被下线/关闭，如果组合在某时刻组合关闭，那么认为该组合以后永远无效。
- 4、取每个投资组合 T 末的仓位量，将所有未关闭的有效投资组合关于某股票的仓位量进行加总；
- 5、根据股票收盘价计算仓位量相应的价值大小。用 T 期仓位价值变动/T<sub>-1</sub> 仓位价值得到仓位价值变动因子（注意，这里用的是收盘价而没有做复权处理，前面提及，一旦发生高送转，那么相应标的仓位量会相应变动）；

$$\text{仓位价值变动因子} = \frac{Vol_{T_0} * P_{T_0}}{Vol_{T_{-1}} * P_{T_{-1}}} - 1 \dots \dots (1)$$

T <sub>0</sub>	… … … … 当前周期
T <sub>-1</sub>	… … … … 上一周期
Vol	… … … … 仓位量
P	… … … … 复权收盘价

实际上通过上述方式计算的价值变动因子有一个严重的问题：价值的变动到底是因为用户持仓量的变动所致还是由持仓标的的价格变动所致，如果是后者那么实际上意味着这种方式和反转因子的相关性很高。所以，为了剥离价格变动产生的影响，我们进一步优化因子的计算方式，那么将步骤 5 替换成步骤 5’：

5’、T 期仓位量变动\*T 期标的价格因子/T<sub>-1</sub> 仓位价值计算价值变动因子；

$$\text{仓位价值变动因子} = \frac{\left( Vol_{T_0} - Vol_{T_{-1}} \right) * P_{T_0}}{Vol_{T_{-1}} * P_{T_{-1}}} \dots \dots (2)$$

实际上我们欲剥离的价格变动的的影响，即在原来价值变动因子上-价格变动：

$$\text{仓位价值变动因子} = \frac{Vol_{T_0} * P_{T_0}}{Vol_{T_{-1}} * P_{T_{-1}}} - 1 - \frac{\left( P_{T_0} - P_{T_{-1}} \right)}{P_{T_{-1}}} \dots \dots (3)$$

公式 (3) 与公式 (2) 是等价的, 为了便利, 我们启用公式 (2) 计算。

### 3.4、价值变动因子有效性测试

价值变动因子定义: 具体参见公式 (2);

个股数据: 全市场 A 股, 剔除交易日当天为 ST、停牌、涨跌停股票;

行业分类: 中信一级行业;

数据区间: 2015 年 1 月 1 日至 2017 年 3 月 17 日;

构建月度频率的价值变动因子, 并对其做 IC 检验。检测结果显示, IC 均值为 0.038, 风险调整 IC 为 0.562。而从 IC 历史走势来看, 大于 0 的占比达到 70%, 证明了该因子的有效性。

**图表 18、月度价值变动因子的 IC 检验**

	平均值	标准差	最小值	最大值	风险调整的 IC	t 统计量
月度价值变动因子	0.038	0.068	-0.089	0.227	0.562	2.92

资料来源: 雪球网, 兴业证券研究所

我们进一步测试了该因子与行业以及市值的关系。近两年来小市值效应比较明显, 那么月度价值变动因子的有效性是否因为踏中了“小市值效应”的节奏呢? 我们研究月度价值变动因子与市值的关系。结果发现流通市值与价值变动因子相关性较弱, 月度价值变动因子与流通市值秩相关性均值达到 0.005。这意味着, 该因子与市值相关性非常低。

进一步检测了行业中性化以后的价值变动因子表现, IC 测试显示, 行业中性化以后, IC 水准并没有下降, 而稳定性有所大幅提升。

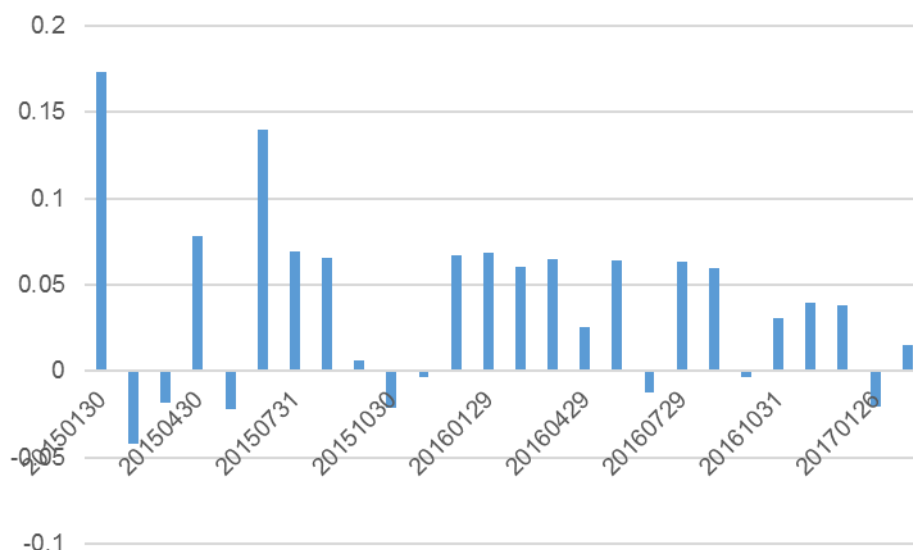
后面如无特殊强调, 价值变动因子均是指行业中性化以后的因子。

**图表 19、行业中性化以后的月度价值变动因子的 IC 检验**

	平均值	标准差	最小值	最大值	风险调整的 IC	t 统计量
月度价值变动因子	0.037	0.048	-0.043	0.174	0.750	3.920

资料来源: 雪球网, 兴业证券研究所

**图表 20、月度价值变动因子的 IC 趋势图**



资料来源: 雪球网, 兴业证券研究所

而从分位数组别的测试结果来看, 组别之间的单调性一般, 但是首尾组别的

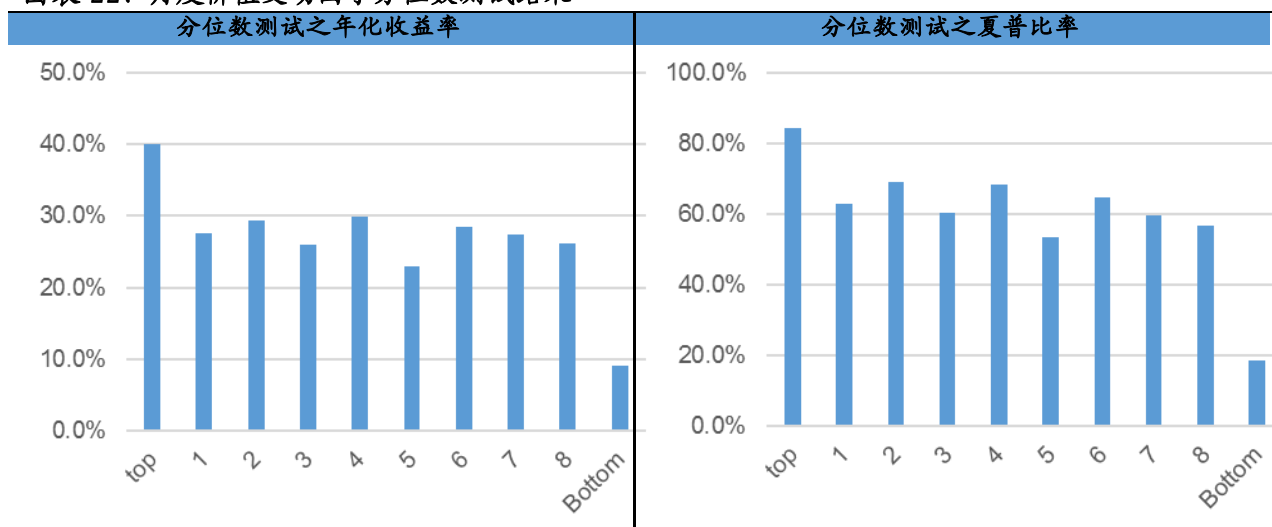
差异性非常大，首组表现出强势的选股能力，年化收益率达到 40.38%，夏普率为 0.86。

图表 21、月度价值变动因子分位数组合测试结果

	总收益率	年化收益率	Sharpe 比率	换手率	最大回撤率	年化超额收益率	跟踪误差	信息比率	胜率
top	112.47%	40.38%	0.86	1.83	39.10%	11.60%	4.40%	2.64	75.00%
1	72.96%	27.96%	0.64	1.78	37.80%	0.32%	2.84%	0.11	50.00%
2	76.88%	29.26%	0.70	1.76	35.06%	0.57%	4.88%	0.12	46.43%
3	68.64%	26.52%	0.63	1.75	39.08%	-1.31%	3.17%	-0.41	46.43%
4	79.70%	30.19%	0.70	1.73	37.29%	2.00%	2.38%	0.84	57.14%
5	62.47%	24.41%	0.57	1.74	41.51%	-2.86%	4.09%	-0.70	39.29%
6	74.97%	28.63%	0.67	1.76	38.85%	0.62%	3.37%	0.18	50.00%
7	72.75%	27.90%	0.62	1.78	40.25%	0.90%	3.03%	0.30	60.71%
8	63.01%	24.60%	0.53	1.81	42.49%	-1.27%	4.66%	-0.27	42.86%
bottom	25.63%	10.81%	0.22	1.84	49.67%	-11.54%	7.26%	-1.59	14.29%
市场	70.64%	27.19%	0.62		39.47%				
L-S	65.18%	25.34%	3.34		2.67%				

资料来源：雪球网，兴业证券研究所

图表 22、月度价值变动因子分位数测试结果



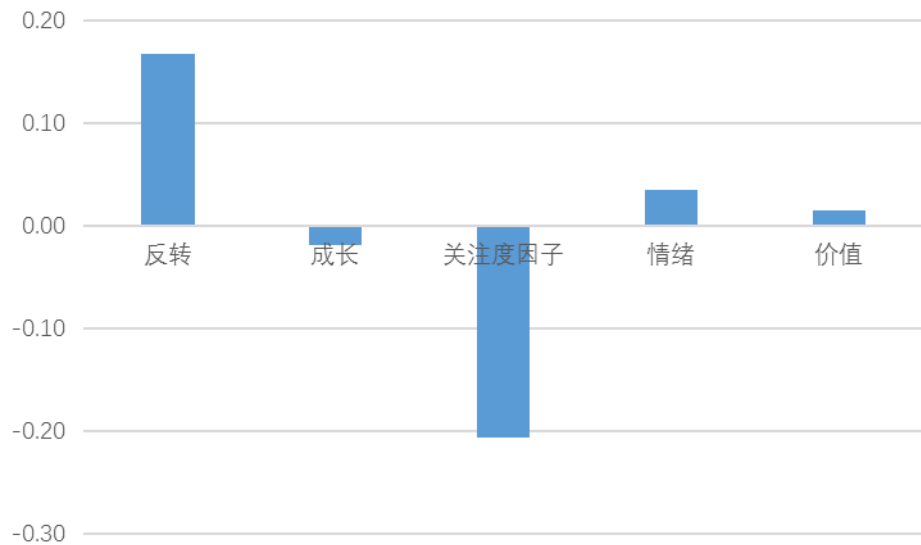
资料来源：Wind，兴业证券研究所

### 3.5、与传统选股因子对比分析，不一样的烟火

前面的测试证明了行业中性化之后的价值变动因子依然有较强的选股能力。那么，和传统的价值、成长、反转、分析师情绪因子相比，其效果如何呢？在经过这些因子的“剥削”之后，价值变动因子是否依然是那个“不一样的烟火”？我们从相关性分析以及 Fama-Beth 回归角度进行了验证。

测试结果显示，价值变动因子与传统因子以及《雪球系列一》中的关注度因子相关性较小。与《雪球系列一》中的关注度因子相关性较小意味着，虽然两个因子的数据源相同，均来自于雪球网，但是很好的保证了特异性。

图表 23、价值变动因子与传统因子秩相关性



资料来源：雪球网，兴业证券研究所

进一步我们对价值变动因子进行 Fama-MacBeth 回归分析，将价值、成长、反转、分析师情绪、《雪球系列一》中的关注度这 5 大风格因子以及价值变动因子作为解释变量，结果如下表所示。在我们加入各类风格因子之后，价值变动因子对股票收益的解释能力依然强悍。因而，相对于传统的几大类 Alpha 因子而言，该因子确实带来了信息增量，也确实是“不一样的烟火”。

图表 24、Fama-Beth 回归分析

		价值变动因子	关注度	成长	反转	分析师情绪	价值
收益率	综合回归	0.03	0.05	0.07	0.11	0.05	0.00
	单个因子回归	0.02	0.04	0.03	0.01	0.02	0.04
T 统计量	综合回归	2.66	2.77	4.87	3.55	3.98	0.01
	单个因子回归	2.98	2.98	2.37	2.73	5.15	1.26

资料来源：雪球网，兴业证券研究所

## 4、基于价值变动因子的选股策略构建

选股基本条件：

个股数据：全市场 A 股，剔除交易日当天为 ST、停牌、涨跌停股票；

行业分类：中信一级行业；

数据区间：2015 年 1 月 1 日至 2017 年 3 月 17 日；

### 4.1、价值变动因子选股策略构建

进一步利用价值变动因子，在全市场选股并用中证 500 对冲，选股流程如下：每个月最后一个交易日结束后进行组合的筛选和调整，按照中信行业对价值变动因子排序、选择每个行业因子值最低的 10% 比例股票建仓，按中证 500 权重配权。

从对冲策略的结果来看，全样本选股策略年化超额收益为 22.25%，夏普比率为 3.27，最大回撤为 6.25%。



图表 25、价值变动因子策略选股策略表现

	总收益率	年化收益率	年化波动率	Sharpe 比率	胜率	最大回撤
多头	82.21%	31.13%	36.56%	0.85	58.10%	50.88%
基准	21.80%	9.32%	35.88%	0.26	56.98%	54.35%
对冲	56.02%	22.25%	6.80%	3.27	62.57%	6.43%

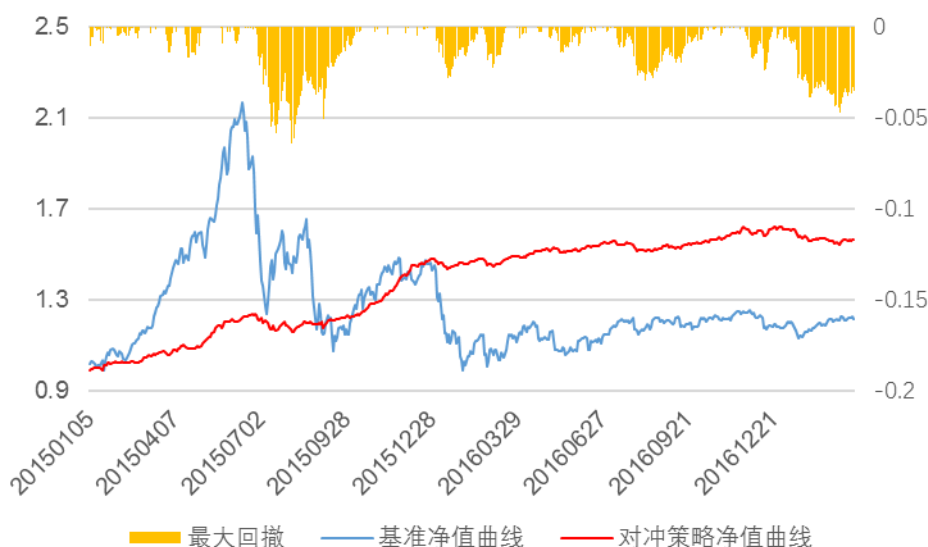
资料来源：雪球网，兴业证券研究所

图表 26、价值变动因子策略在不同年度的表现

	策略多头	基准	对冲
2015 年	102.80%	43.12%	46.69%
2016 年	-10.34%	-17.78%	9.63%
2017 年	0.52%	3.51%	-2.87%

资料来源：雪球网，兴业证券研究所

图表 27、改进价值选股因子的对冲净值曲线



资料来源：雪球网，兴业证券研究所

## 5、结语

本文是兴业定量研究团队“猎金”系列之十五。在本篇中，我们继续挖掘雪球网的信息价值，抓取用户创建的投资组合。每个投资组合相当于一个基金产品，那么在百万个投资组合中寻求能带来信息增量的选股因子。最终，通过定义有效投资组合，根据投资组合的价值变动，定义价值变动因子。基于该因子在全市场选股，并用中证 500 作为对冲基准，测试结果验证了该因子良好的选股能力。

见微知著，通过“猎金系列十四”我们构建了关注度因子，而本研究构建了价值变动因子，这些为我们随后的文本分析工作打下坚实基础。

扬帆起航，接下来，我们将基于雪球网的文本信息，开拓情绪分析类选股、择时研究。争取，最大化的让兴业金工力量服务大家、让“文本力量”为大家所用。

投资评级说明

**行业评级** 报告发布日后的 12 个月内行业股票指数的涨跌幅度相对同期恒生指数的涨跌幅为基准,投资建议的评级标准为:

- 推 荐: 相对表现优于市场;
- 中 性: 相对表现与市场持平
- 回 避: 相对表现弱于市场

**公司评级** 报告发布日后的 12 个月内公司的涨跌幅度相对同期恒生指数的涨跌幅为基准,投资建议的评级标准为:

- 买 入: 相对大盘涨幅大于 15% ;
- 增 持: 相对大盘涨幅在 5% ~ 15%之间
- 中 性: 相对大盘涨幅在-5% ~ 5%;
- 减 持: 相对大盘涨幅小于-5%

机构销售经理联系方式

<b>机构销售负责人</b>		邓亚萍	021-38565916	dengyp@xyzq.com.cn	
<b>上海地区销售经理</b>					
姓名	办公电话	邮 箱	姓名	办公电话	邮 箱
盛英君	021-38565938	shengyj@xyzq.com.cn	冯诚	021-38565411	fengcheng@xyzq.com.cn
顾超	021-20370627	guchao@xyzq.com.cn	杨忱	021-38565915	yangchen@xyzq.com.cn
			王溪	021-20370618	wangxi@xyzq.com.cn
王立维	021-38565451	wanglw@xyzq.com.cn	李远帆	021-20370716	liyuanfan@xyzq.com.cn
			胡岩	021-38565982	huyan@xyzq.com.cn
姚丹丹	021-38565778	yaodandan@xyzq.com.cn	曹静婷	18817557948	caojt@xyzq.com.cn
地址: 上海浦东新区长柳路 36 号兴业证券大厦 12 层 (200135) 传真: 021-38565955					
<b>北京地区销售经理</b>					
姓名	办公电话	邮 箱	姓名	办公电话	邮 箱
郑小平	010-66290223	zhengxiaoping@xyzq.com.cn	朱圣诞	010-66290197	zhusd@xyzq.com.cn
			刘晓洲	010-66290220	liuxiaoliu@xyzq.com.cn
肖霞	010-66290195	xiaoxia@xyzq.com.cn	陈扬	010-66290197	chenyang@xyzq.com.cn
袁博	15611277317	yuanb@xyzq.com.cn	吴磊	010-66290190	wulei@xyzq.com.cn
陈殊宏	15117943079	chenshuhong@xyzq.com.cn	王文凯	010-66290197	wangwenkai@xyzq.com.cn
地址: 北京西城区锦什坊街 35 号北楼 601-605 (100033) 传真: 010-66290220					
<b>深圳地区销售经理</b>					
姓名	办公电话	邮 箱	姓名	办公电话	邮 箱
朱元贱	0755-82796036	zhuyuan@xyzq.com.cn	杨剑	0755-82797217	yangjian@xyzq.com.cn
李昇	0755-82790526	lisheng@xyzq.com.cn	邵景丽	0755-23836027	shaojingli@xyzq.com.cn
王维宇	0755-23826029	wangweiyu@xyzq.com.cn			
地址: 福田区中心四路一号嘉里建设广场第一座 701 (518035) 传真: 0755-23826017					
<b>国际机构销售经理</b>					
姓名	办公电话	邮 箱	姓名	办公电话	邮 箱
刘易容	021-38565452	liuyirong@xyzq.com.cn	徐皓	021-38565450	xuhao@xyzq.com.cn
张珍岚	021-20370633	zhangzhenlan@xyzq.com.cn	陈志云	021-38565439	chanchiwan@xyzq.com.cn
马青岚	021-38565909	maql@xyzq.com.cn	曾雅琪	021-38565451	zengyaqi@xyzq.com.cn
申胜雄	021-20370768	shensx@xyzq.com.cn	陈俊凯	021-38565472	chenjunkai@xyzq.com.cn
俞晓琦	021-38565498	yuxiaoqi@xyzq.com.cn	蔡明珠	13501773857	caimzh@xyzq.com.cn
地址: 上海浦东新区长柳路 36 号兴业证券大厦 12 层 (200135) 传真: 021-38565955					
<b>私募及企业客户负责人</b>		刘俊文	021-38565559	liujw@xyzq.com.cn	
<b>私募销售经理</b>					
姓名	办公电话	邮 箱	姓名	办公电话	邮 箱
徐瑞	021-38565811	xur@xyzq.com.cn	杨雪婷	021-20370777	yangxueting@xyzq.com.cn
唐恰	021-38565470	tangqia@xyzq.com.cn	韩立峰	021-38565840	hanlf@xyzq.com.cn
李桂玲	021-20370658	ligl@xyzq.com.cn	施孜琪	021-20370837	shzq@xyzq.com.cn
王磊	021-20370658	wanglei1@xyzq.com.cn			
地址: 上海浦东新区长柳路 36 号兴业证券大厦 12 层 (200135) 传真: 021-38565955					

### 港股机构销售服务团队

机构销售负责人			丁先树	18688759155	dingxs@xyzq.com.hk
姓名	办公电话	邮箱	姓名	办公电话	邮箱
王文洲	18665987511	petter.wang@xyzq.com.hk	郑梁燕	18565641066	zhengly@xyzq.com.hk
陈振光	13818288830	chenzg@xyzq.com.hk	周围	13926557415	zhouwei@xyzq.com.hk
孙博轶	13902946007	sunby@xyzq.com.hk			

地址: 香港中环德辅道中 199 号无限极广场 32 楼 3201 室 传真: (852) 3509-5900

#### 【信息披露】

兴业证券股份有限公司(“本公司”) 在知晓的范围内履行信息披露义务。客户可登录 [www.xyzq.com.cn](http://www.xyzq.com.cn) 内幕交易防控栏内查询静默期安排和关联公司持股情况。

#### 【分析师声明】

本人具有相关监管机构所须之牌照。本人确认已合乎监管机构之相关合规要求, 并以勤勉的职业态度, 独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因, 不因, 也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

#### 【法律声明】

本报告由兴业证券股份有限公司(已具备证券投资咨询业务资格)制作。

本报告由受香港证监会监察的兴证国际证券有限公司(香港证监会中央编号: AYE823)于香港提供。香港的投资者若有任何关于本报告的问题请直接联系兴证国际证券有限公司的销售交易代表。

本报告将依据其他国家或地区的法律法规和监管要求于该国家或地区提供本报告。

本公司不会因接收人收到本报告而视其为客户。客户应当认识到有关本报告的短信提示、电话推荐等只是研究观点的简要沟通, 需以本公司 <http://www.xyzq.com.cn> 网站刊载的完整报告为准, 本公司接受客户的后续问询。

本公司的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。

本公司的资产管理部、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告并非针对或意图发送予或为任何就发送、发布、可得到或使用此报告而使本公司违反当地的法律或法规或可致使本公司受制于相关法律或法规的任何地区、国家或其他管辖区域的公民或居民, 包括但不限于美国及美国公民(1934年美国《证券交易所》第15a-6条例定义为本「主要美国机构投资者」除外)。

本报告可能附载其它网站的地址或超级链接。对于本报告可能涉及到本公司网站以外的资料, 本公司未有参阅有关网站, 也不对它们的内容负责。提供这些地址或超级链接的目的, 纯粹为了收件人的方便及参考, 连结网站的内容不构成本报告的任何部份。收件人须承担浏览这些网站的风险。

本公司系列报告的信息均来源于公开资料, 本公司对这些信息的准确性和完整性不作任何保证, 也不保证所包含的信息和建议不会发生任何变更。本公司已力求报告内容的客观、公正, 但文中的观点、结论和建议仅供参考, 报告中的信息或意见并不构成所述证券的买卖出价或征价, 投资者据此做出的任何投资决策与本公司和作者无关。

在法律许可的情况下, 兴业证券股份有限公司可能会持有本报告中提及公司所发行的证券头寸并进行交易, 也可能为这些公司提供或争取提供投资银行业务服务。因此, 投资者应当考虑到兴业证券股份有限公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

若本报告的接收人非本公司的客户, 应在基于本报告作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。未经授权的转载, 本公司不承担任何转载责任。