

分析师：  
徐寅  
[xuyinsh@xyzq.com.cn](mailto:xuyinsh@xyzq.com.cn)  
S0190514070004

研究助理：  
郑兆磊  
[zhengzhao lei@xyzq.com.cn](mailto:zhengzhao lei@xyzq.com.cn)

## 基于专利分类的科技动量因子研究

2019 年 06 月 25 日

### 报告关键点

我们构建的科技动量因子在行业中性化后，因子 IC\_IR 达到 0.34，T 值达到 2.87，分位数测试各组别严格单调；另外，该因子与传统的价值、成长、动量反转、分析师情绪、以及授权发明专利数因子相关系数很低，具有稳定信息增量。进一步构建的基于科技动量与单月反转的选股策略同样表现优异。

### 相关报告

团队成员：

### 投资要点

- 科创板目前正如火如荼，对一个企业的科技创新能力的衡量成为越来越多的投资人关注重点。继专利数据领域的第一篇报告《从中国心到中国芯-由贸易战引发的专利选股有效研究》之后，我们在第二篇深度研究中，以专利 IPC 分类数据为基础，通过引入**科技关联度**这一重要概念，构建了**科技动量因子**，选股效果颇佳。
- 针对于科技动量因子的**测试结果表明**，在行业中性化后，因子年化 IC\_IR 可达 1.18，T 值达到 2.87，各分位数组别收益率严格单调；**相关性方面**：该因子与传统的价值、成长、动量反转、分析师情绪以及授权发明专利数因子相关性基本为 0；Fama-MacBeth 回归测试验证该因子的信息增量十分显著，加入上面提到的 5 类因子后，综合回归的 T 值仍然高达 3.71。
- 从逻辑上讲，科技动量因子捕捉的是个股对与其具有潜在科技关联的公司的股价反应不充分、表现相对滞后所带来的机会，而反转则表现了市场对公司信息的过度反应。两者的结合，首先帮我们筛选出了股价可能被错杀的个股（前期大幅下跌，反转效应强），而后再从这其中挑选科技动量强劲的标的。该策略在 2013 年至 2018 年的 6 年里显著跑赢了单一反转策略，效果非常稳定。

**风险提示：**本报告模型及结论全部基于对历史数据的分析，当市场环境变化时，存在模型失效风险。



## 目 录

1、专利因子研究回顾 .....	- 3 -
1.1、研究背景 .....	- 3 -
1.2、过往研究回顾 .....	- 4 -
2、科技动量因子构建 .....	- 6 -
2.1、专利 IPC 分类 .....	- 6 -
2.2、从科技关联度到科技动量 .....	- 9 -
3、科技动量因子有效性验证 .....	- 12 -
3.1、因子有效性测试的方法 .....	- 12 -
3.2、科技动量因子测试 .....	- 13 -
3.3、科技动量因子特异性分析 .....	- 14 -
4、基于科技动量因子的选股策略研究 .....	- 15 -
5、总结 .....	- 17 -
6、参考文献 .....	- 18 -
7、附录 .....	- 19 -
图表 1、历届《政府工作报告》中提及“创新”的次数 .....	- 3 -
图表 2、中国大陆历年专利申请量(单位：万件) .....	- 4 -
图表 3、不同行业统计截止日前所有有效发明专利总数统计 .....	- 5 -
图表 4、部分专利因子定义 .....	- 6 -
图表 5、原始专利因子以及市值中性化以后的表现 .....	- 6 -
图表 6、专利因子与传统因子的 FamaMacBeth 回归分析 .....	- 6 -
图表 7、第一部的部分分类展示 .....	- 7 -
图表 8、专利 IPC 分类原始数据示意 .....	- 7 -
图表 9、不同部的专利总量 .....	- 8 -
图表 10、不同部的专利总量时间序列的变动趋势 .....	- 8 -
图表 11、滚动五年有效授权发明专利总数时序图 .....	- 9 -
图表 12、单个专利映射到二级分类的平均数量时序图 .....	- 9 -
图表 13、中国石化与中国化学在 20121231-20181231 期间的科技关联度 .....	- 10 -
图表 14、中国石化与中国化学 IPC 二级分类的分布情况 .....	- 11 -
图表 15、中国石化与中国化学在 20121231-20181231 期间的股价走势 .....	- 12 -
图表 16、Rank IC 计算过程 .....	- 13 -
图表 17、分位数组合法 .....	- 13 -
图表 18、科技动量因子 IC 表现 .....	- 13 -
图表 19、科技动量因子分位数表现 .....	- 14 -
图表 20、科技动量因子多头及多空净值曲线 .....	- 14 -
图表 21、科技动量因子与传统因子的相关性 .....	- 15 -
图表 22、基于各大类因子的 Fama-MacBeth 测试 T 值 .....	- 15 -
图表 23、流通市值加权下基于科技动量因子构建的选股策略表现 .....	- 16 -
图表 24、流通市值加权下基于科技动量因子构建的选股策略净值曲线 .....	- 16 -
图表 25、等权情况下基于科技动量因子构建的选股策略表现 .....	- 16 -
图表 26、等权情况下基于科技动量因子构建的选股策略净值曲线 .....	- 16 -
图表 27、专利因子定义 .....	- 19 -
图表 28、专利重要概念汇总 .....	- 19 -

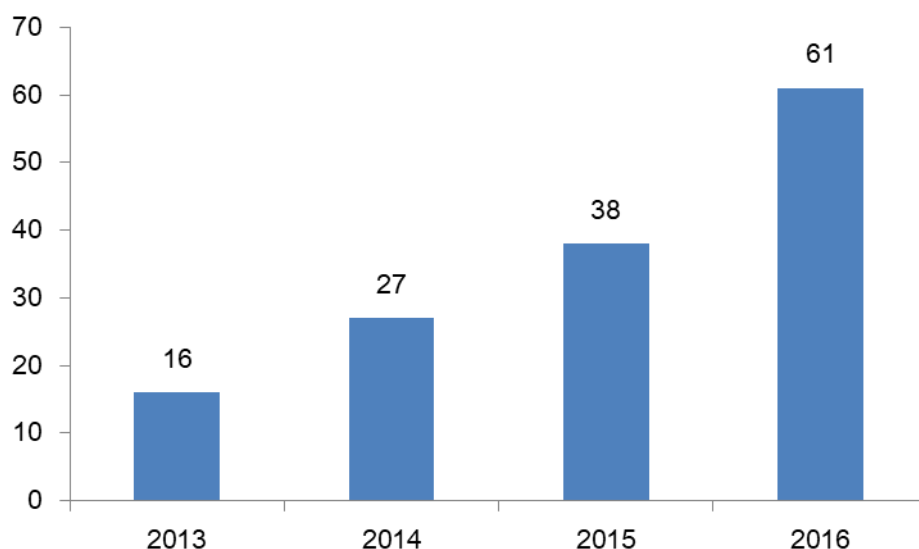
## 报告正文

## 1、专利因子研究回顾

## 1.1、研究背景

从 1978 年改革开放至今，中国经济增长在很大程度上是通过低成本、高素质的劳动力比较优势和资源、环境要素禀赋驱动的，也因而有了传统的三驾马车的发展模式（投资、消费、出口），同时出现了中国制造这样的全球化现象。2008 年全球金融危机之后，伴随着内部和外部约束因素的增强，出口、投资这两具引擎对中国经济增长的边际贡献呈现逐年下降的趋势，中国经济的增长开始聚焦于创新驱动的发展战略之上，希望逐渐由“中国制造”过渡到“中国创造”。当前，创新已经成为了中国的国家战略发展方向，相关利好政策不断出台。

图表 1、历届《政府工作报告》中提及“创新”的次数



资料来源：兴业证券经济与金融研究院整理

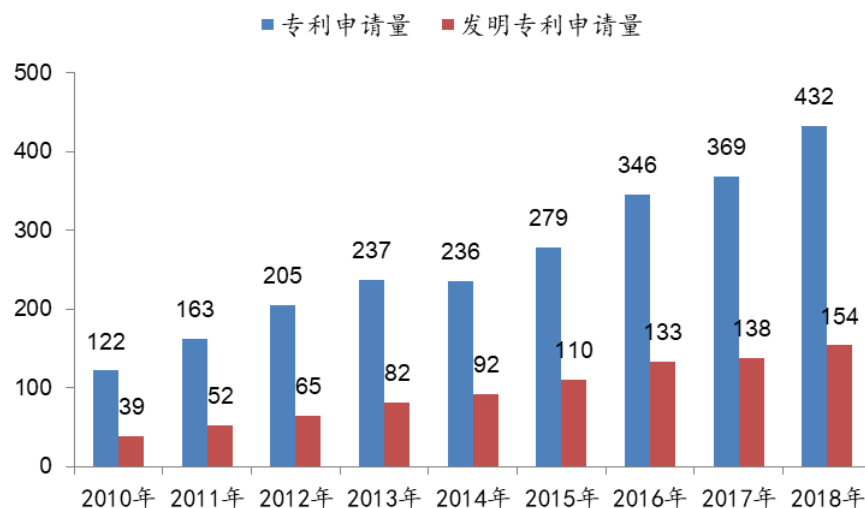
2018 年 1 月 30 日，习近平总书记在主持中央政治局会议时就提出要建设创新引领、协同发展的产业体系；发改委也表示，将加快推动新兴产业发展，组织实施一批新兴产业重大工程，大力发展“新技术、新产业、新业态、新模式”的四新经济形态；2 月 28 日证监会拟对生物科技、云计算在内的 4 个行业放宽 IPO 条件。

2018 年 4 月中旬，美国商务部宣布，今后 7 年内，将禁止该国企业向中国电信设备制造商中兴通讯出售任何电子技术或通讯元件，中美贸易摩擦升级到高科技领域。中美贸易摩擦实际上从某个侧面说明，中国的创新发展已经动了美国的“奶酪”，但面对历史的趋势，我们避无可避。

2018 年 11 月 5 日，国家主席习近平在首届中国国际进口博览会开幕式上宣布设立科创板。科创板是独立于现有主板市场的新设板块，设立科创板并试点注册制是提升服务科技创新企业能力、增强市场包容性、强化市场功能的一项资本市场重大改革举措。

2019年4月，美国倾国家之力，全面封杀华为，这种逆全球化的手段虽不得人心，但究其原因无外乎在于核心专利技术所映射的大国科技霸权。中国崛起不仅仅需要一个华为，我们要千千万万个华为。幸运的是，我们已经在这条道路上披荆斩棘、不断前行。从2010年开始，国内的专利申请量突破了100万件，同时也成为全球范围内专利申请量最多的国家。

图表 2、中国大陆历年专利申请量(单位：万件)



资料来源：国家专利局、兴业证券经济与金融研究院整理

从上述的一系列事件中我们不难看出，科技创新将是未来一个时期中国经济发展的重要驱动力，也是大国之间竞争的主战场。资本市场，特别是以股票交易为核心的二级市场，本来就是服务实体经济的重要抓手，因此把和创新相关的信息纳入到我们的股票投资分析框架之中，将是未来投资管理人必然的选择。本文正是从反映创新成果的重要维度——专利信息入手，尝试从量化的角度将其融入股票组合的筛选过程之中，增强投资组合的表现。

## 1.2、过往研究回顾

在专利数据系列研究的第一篇报告《从中国心到中国芯--贸易战引发的专利选股有效性研究》中，我们详细的介绍了专利数据的特征，因子的构建方式，并对专利因子的选股能力进行了全面的测试。这里，我们将从以上三个方面对之前的研究做一个简单的回顾。

### ➤ 专利数据特征

我们研究所使用的专利数据由深圳德高行知识产权数据技术有限公司提供，德高行是一家专业的知识产权解决方案供应商，其原始数据来自于国家专利局。我们对专利数据的分析主要从**专利分类**、**专利所处状态**以及**具体指标**三个维度展开：

1. 专利分类：从类别上来看，专利一共有三类：
  - a) 发明专利：是指对产品、方法或者其改进所提出的新的技术方案；
  - b) 实用新型专利：是指对产品的形状、构造或者其结合所提出的适于实用

的新的技术方案；

- c) 外观设计专利：是指对产品的形状、图案或者其结合以及色彩与形状、图案的结合所作出的富有美感并适于工业应用的新设计。
2. 专利状态：一个专利的获得大致要经历专利申请、专利公开、实质审查、专利授权几个流程（注意新型和外观专利没有实质审查，仅有形式审查）。
3. 具体指标：落脚到对专利具体的描述，目前已有的数据包括：专利数量、说明书总字数、权利要求总项数、独权总项数、附图总数、摘要总字数、IPC分类号总数、寿命加总、专利审查期加总、前引、后引、非专利方面引用、同族专利共计 13 种类型，具体释义参见附录。

由于我们对专利数据的主要应用场景为 A 股市场投资策略的开发，因此对数据的行业分布特征是非常关注的。我们以授权发明专利总数为研究对象，分析了专利的行业分布特点。从统计结果来看，不同行业的专利数量有很大的差异，专利较多且数量稳定的行业有 10 个：家电、电子元器件、汽车、机械、基础化工、通信、电力设备、计算机、医药、建筑，我们称之为专利属性行业。

**图表 3、不同行业统计截止日前所有有效发明专利总数统计**

	2011 年	2013 年	2015 年	2018 年
电子元器件	2,332	4,891	11,700	30,535
家电	1,597	4,025	8,855	29,083
通信	9,389	14,940	21,049	27,609
石油石化	5,228	8,877	14,694	26,743
机械	2,366	5,069	10,775	22,614
汽车	2,250	4,634	8,202	18,178
计算机	2,469	4,928	8,295	16,461
建筑	1,020	2,817	6,231	15,818
电力设备	1,361	2,978	5,325	12,017
医药	2,441	4,512	7,492	11,910
基础化工	1,394	3,164	5,570	11,706
钢铁	1,999	3,729	5,388	8,519
国防军工	252	1,093	2,618	6,893
有色金属	1,208	1,911	2,793	4,791
电力及公用事业	160	432	1,478	3,503
建材	438	904	1,542	3,037
食品饮料	220	828	1,574	2,422
农林牧渔	246	829	1,625	2,306
轻工制造	285	559	898	1,809
纺织服装	277	530	933	1,645
煤炭	105	372	708	1,327
传媒	40	118	265	1,247
综合	73	183	479	1,186
交通运输	297	421	522	905
银行	107	197	282	674
非银行金融	51	57	238	509
房地产	90	200	220	499
商贸零售	69	79	123	206
餐饮旅游	-	-	7	8

资料来源：德高行，兴业证券经济与金融研究院整理

### ➤ 选股因子构建

根据前面提到的专利分类、专利状态和具体指标，在理论上我们可以定义出  $3*4*13=156$  个专利因子（专利有 3 种类型、4 种状态，13 个具体指标）。例如，



我们可以定义公开发明专利数、授权发明专利数或者授权发明专利前引数等因子。但需要注意的是，并不是所有的由排列组合得到的因子都有意义。在实践中，我们一共构建了 33 个专利类的选股因子用于研究，部分因子的定义可以参见图表 4，完整定义请参见附录。

**图表 4、部分专利因子定义**

分类	概念	定义
发明专利	TP010401	统计截止日前三年内新生成的有效发明授权的专利数
	TP010408	统计截止日前三年内新生成的有效发明授权的寿命加总
实用新型	TP020401	统计截止日前三年内新生成的有效实用新型的专利数
	TP020408	统计截止日前三年内新生成的有效实用新型的寿命加总
外观设计	TP030401	统计截止日前三年内新生成的有效外观设计的专利数
	TP030408	统计截止日前三年内新生成的有效外观设计的寿命加总

资料来源：德高行，兴业证券经济与金融研究院整理

### ➤ 因子选股能力研究

授权发明类因子整体都呈现出比较稳定的选股能力。以行业市值中性化后的授权发明专利数因子为例，其 IC 均值达到了 0.024，t 统计量为 6.0，预测能力非常显著，具体结果参见图表-5。

**图表 5、原始专利因子以及市值中性化以后的表现**

	原始因子表现				行业市值中性化以后的表现			
	均值	Std	IC_IR	T 值	均值	Std	IC_IR	T 值
授权发明专利数	0.007	0.09	0.08	0.8	0.024	0.04	0.64	6.0

资料来源：德高行，Wind，兴业证券经济与金融研究院整理

相关性测试结果显示，授权发明专利数因子与价值、成长、分析师情绪的相关性仅为 0.17、-0.01、0.0035，可见专利因子与传统因子的相关性较低。而 FamaMacBeth 回归测试结果显示：即便在加入价值、成长、分析师情绪后，专利因子的选股有效性依然非常强，信息增量特征明显。

**图表 6、专利因子与传统因子的 FamaMacBeth 回归分析**

	TP0104	ValueFactor	SentimentFactor	GrowthFactor
T 统计量（综合回归）	2.49	2.58	7.60	6.90
T 统计量（单独回归）	5.14	4.27	8.19	6.26

资料来源：德高行，Wind，兴业证券经济与金融研究院整理

应该说，专利系列研究的第一篇报告为我们接下来的工作开了一个好头，同时也引发了我们的一些思考。除了一些简单的特征描述指标，我们还能够从哪些角度对专利数据进行分析呢？本文将尝试从专利分类入手，通过定义科技关联度这一核心概念，进而聚焦于科技动量因子的研究。

## 2、科技动量因子构建

### 2.1、专利 IPC 分类

《国际专利分类表》（IPC 分类）是根据 1971 年签订的《国际专利分类斯特拉斯堡协定》编制的，是目前国际通用的专利文献分类和检索工具。国际专利分类系统按照技术主题设立类目，把整个技术领域分为 5 个不同等级：部、大类、小类、大组、小组，分别对应着 8、145、670、3,000+、10,000+ 个类别。某种程

请务必阅读正文之后的信息披露和重要声明

度上 IPC 分类可以理解作为一种全新的行业分类（部分概念释义参见图表-7）。从数据来源角度讲，一个专利的 IPC 分类的原始数据是一个 JPG 格式的文件，从中进行解析即可得到 IPC 分类相应的数据，具体参见图表-8。

图表 7、第一部的部分分类展示

分部：农业	6
A01 农业；林业；畜牧业；狩猎；诱捕；捕鱼	6
A01B 农业或林业的整地；一般农业机械或农具的部件、零件或附件（用于播种、种植或施厩肥的开挖沟穴或覆盖沟穴入 A01C 5/00；收获根作物的机械入 A01D；可变换成整地设备或能够整地的割草机入 A01D 42/04；与整地机具联合的割草机入 A01D 43/12；工程目的的整地入 E01, E02, E21）	6
A01C 种植；播种；施肥（与一般整地结合的入 A01B 49/04；农业机械或农具的部件、零件或附件一般入 A01B 51/00 至 A01B 75/00）	14
A01D 收获；割草	16
A01F 脱粒（联合收割机入 A01D 41/00）；禾秆、干草或类似物的打捆；将禾秆、干草或类似物形成捆或打捆的固定装置或手动工具；禾秆、干草或类似物的切碎；农业或园艺产品的储藏（与收割有关的制作或设置堆垛的设备入 A01D 85/00）	26
A01G 园艺；蔬菜、花卉、稻、果树、葡萄、啤酒花或海藻的栽培；林业；浇水（水果、蔬菜、啤酒花等类植物的采摘入 A01D 46/00；通过组织培养技术的植物再生入 A01H 4/00；洋葱或花球茎的去顶或剥皮装置入 A23N 15/08；繁殖单细胞藻类入 C12N 1/12；植物细胞培养入 C12N 5/00）	29

资料来源：兴业证券经济与金融研究院整理

图表 8、专利 IPC 分类原始数据示意

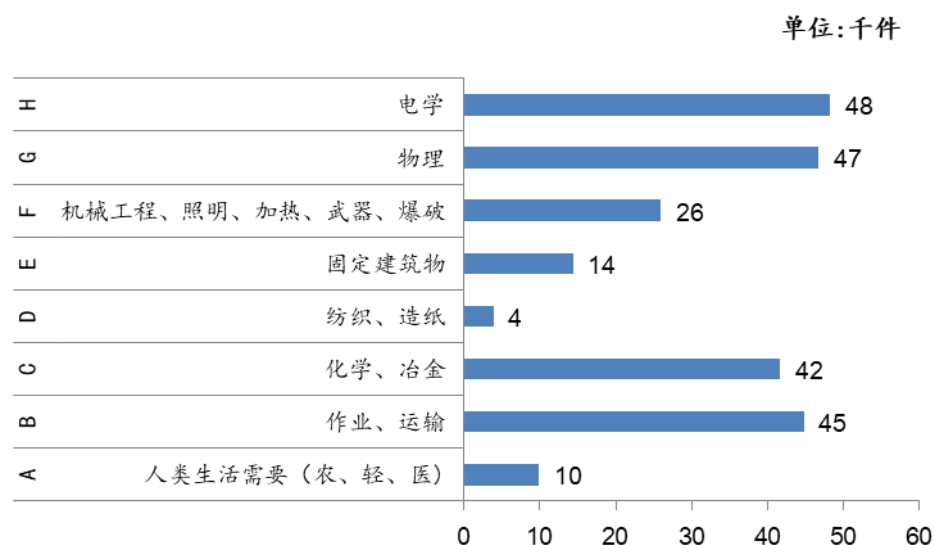
[19] 中华人民共和国国家知识产权局		[51] Int. Cl. <sup>7</sup>	
		C02F 9/04	
		C02F 9/02 A47J 31/00	
		///(C02F9/04,1: 2	
		8,1: 68)	
[12] 实用新型专利说明书			
[21] ZL 专利号 01246490.2			
[45] 授权公告日 2002 年 5 月 22 日		[11] 授权公告号 CN 2492553Y	
[22] 申请日 2001.7.11		[21] 申请号 01246490.2	
[73] 专利权人 蔡忠春		[74] 专利代理机构 浙江翔隆专利事务所	
地址 315411 浙江省余姚市丈亭镇宁波中联电		代理人 白洪长	
器有限公司			
[72] 设计人 蔡忠春			
		权利要求书 1 页 说明书 4 页 附图页数 3 页	
[54] 实用新型名称 饮水机净水矿化装置			

资料来源：德高行，兴业证券经济与金融研究院整理

我们将所有上市公司过去 5 年的有效授权发明专利映射到 IPC 一级分类，并分别统计每个类别包含的数量。从结果来看，不同部之间的差异较大，B、C、G、

H 四大部占绝对优势；从不同部的变动趋势来看，呈现稳定增加的态势。

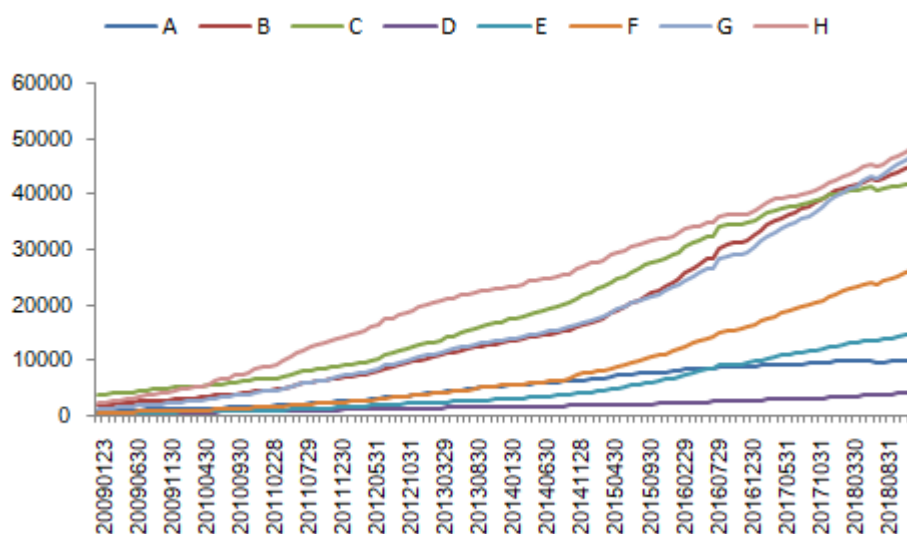
图表 9、不同部的专利总量



统计时点：2018 年 12 月 31 日

资料来源：德高行，兴业证券经济与金融研究院整理

图表 10、不同部的专利总量时间序列的变动趋势

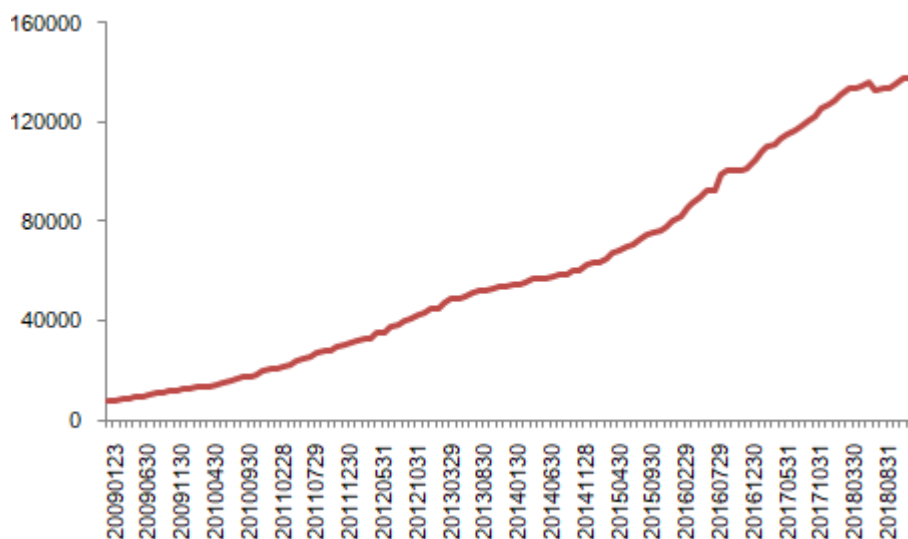


资料来源：德高行，兴业证券经济与金融研究院整理

一个专利可以映射到同一个类别下的多个维度（比如如果考察部这个级别，那么一个专利即可以映射到 A、也可以映射到 B），而这是审查员（专利局）的职权。不同领域、不同时期的审查员的水平可能不一样，IPC 的映射质量可能也会有一定的波动。这里，我们观察单个专利映射到二级分类的平均数量。从趋势来看，平均分类数呈现稳定增加的态势。

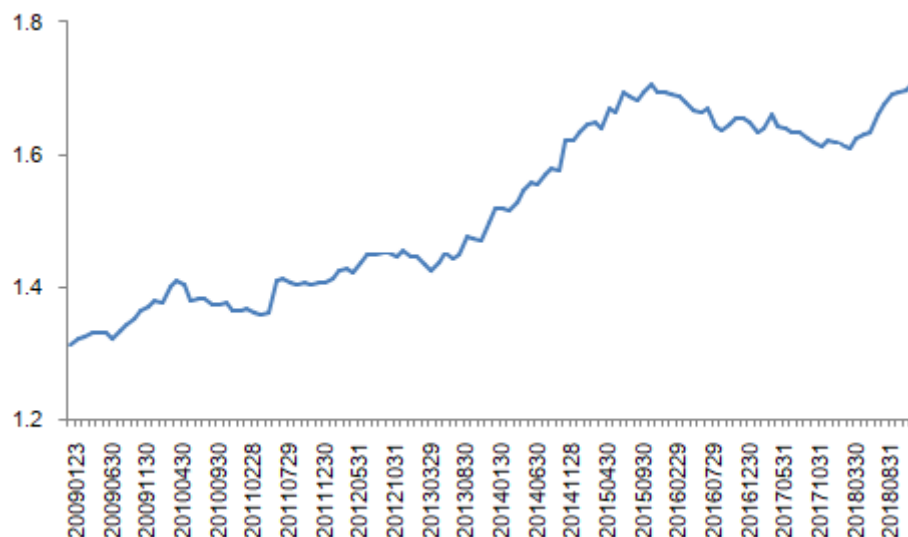


图表 11、滚动五年有效授权发明专利总数时序图



资料来源：德高行，兴业证券经济与金融研究院整理

图表 12、单个专利映射到二级分类的平均数量时序图



资料来源：德高行，兴业证券经济与金融研究院整理

## 2.2、从科技关联度到科技动量

一般来说,我们认为同一个行业内的公司,其股价往往也会呈现相似的走势。所以在传统的选股框架下,从宏观配置(宏观经济层面)、中观把控(行业层面),再到微观选择(个股层面),是一套相对成熟且在被无数次验证和使用的方法论。但与此同时我们也发现,现如今的公司业务线非常复杂,很多时候两家公司的主营业务收入来源可能并不一样,但却在产品研发和创新领域有着深层次的关联,而这种关联很有可能在企业未来的发展中扮演越来越重要的角色。如何在传统行业分类方式之外对这种科技领域之间的联系进行有效刻画?这正是接下来我们将要构建的**科技关联度**指标所擅长的。

科技关联度的定义离不开专利的 IPC 分类。前面提及,IPC 分类共有 5 层:

请务必阅读正文之后的信息披露和重要声明

部、大类、小类、大组、小组。我们以第二层——大类（共计 145 个大类）为例来阐述科技关联度的定义如下：

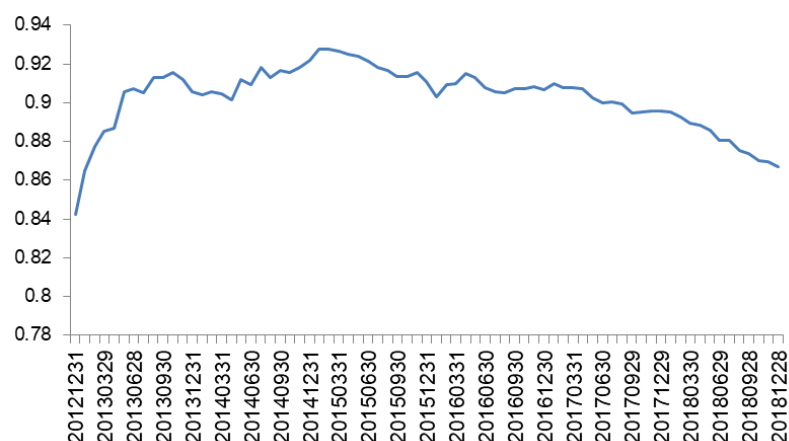
$$tech_{ij,t} = \frac{T_{i,t} \cdot T_{j,t}'}{(T_{i,t} \cdot T_{i,t}')^{\frac{1}{2}} (T_{j,t} \cdot T_{j,t}')^{\frac{1}{2}}}$$

$T_{i,t}$  为公司 i 在 t 时刻的 IPC 二级分类， $tech_{ij,t}$  公司 i 和公司 j 在 t 时刻的科技关联度值。

从上述定义公式不难看出，科技关联度本质上表达的是两家公司的专利在类别上的相似程度。指标值越大，代表两个公司在申请的专利类别上布局越相似、创新研发方向也越趋于一致。这种相似性的衡量方式，已经跨域了传统财报以及行业分类所能提供的信息，在一定程度上反映了公司未来的发展战略方向。这里我们通过一个具体的例子体会一下。

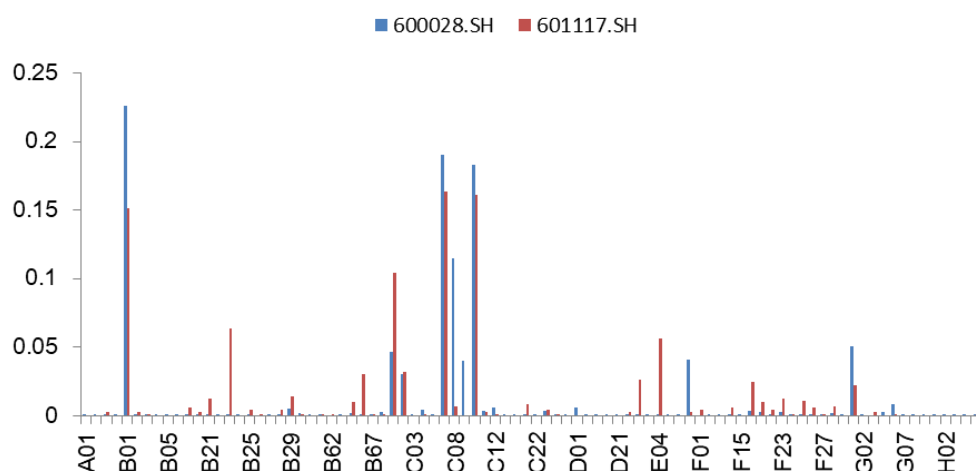
中国石化（600028.SH）与中国化学（601117.SH）两个公司分属于石油石化与建筑这两个不同的行业，在主营业务层面两者的关系并不大。但通过计算，我们发现两者的科技关联度一直维持在一个较高的水平（具体参见图表 13）。以两家公司在 2018 年 12 月 31 日的 IPC 二级分类分布为例，我们发现两家公司在大部分专利类别中都呈现出较高的一致性（具体参见图表-14）。

图表 13、中国石化与中国化学在 20121231-20181231 期间的科技关联度



资料来源：德高行，兴业证券经济与金融研究院整理

图表 14、中国石化与中国化学 IPC 二级分类的分布情况



注：横坐标代表相应公司某一个二级分类占该二级分类总数的比例

资料来源：德高行，兴业证券经济与金融研究院整理

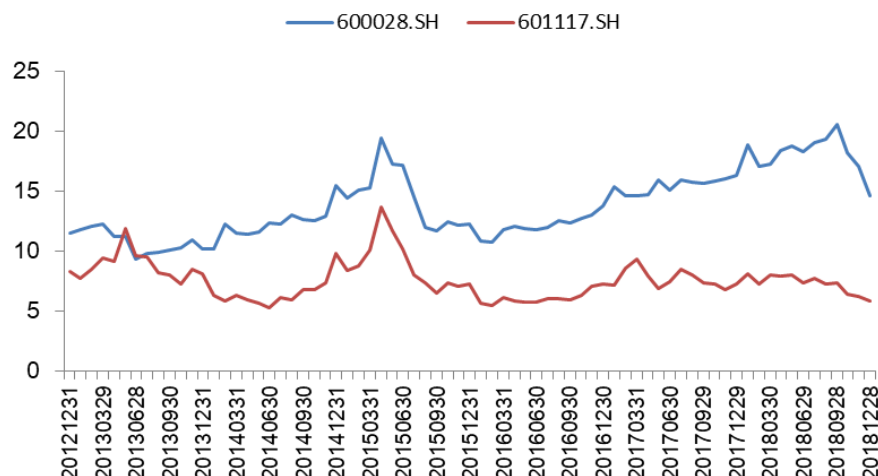
正因为科技关联度高的公司在业务方向和创新研发上具有更高的相似性，所以我们也有理由推测，其二级市场股价的联动性也将更为显著。有鉴于此，我们利用科技关联度和股票涨跌幅开发了科技动量因子，定义如下式：

$$tech\_ret_{i,t} = \frac{\sum_{j \neq i} tech_{ij,t} \cdot ret_{j,t}}{\sum_{j \neq i} tech_{ij,t}},$$

$ret_{j,t}$  代表公司  $j$  在  $t$  时刻月度收益率， $tech_{ij,t}$  表示公司  $i$  和公司  $j$  在  $t$  时刻科技关联度， $tech\_ret_{i,t}$  代表公司  $i$  在  $t$  时刻的科技动量因子。简单来说，我们在每个月底，计算了除目标公司外的其他公司在科技关联度下的加权平均月收益率，而该收益率就是我们所谓的科技动量。

我们也是在一些观察经验的基础上作出上述推测的，例如前面提到的中国石化（600028.SH）与中国化学（601117.SH）这两只股票，观察两者的股价，我们发现其走势一致性非常强。那么这种规律是否能够在更多的股票以及更长的时段上成立呢？我们将在第三章中给出详细的测试结果。

图表 15、中国石化与中国化学在 20121231-20181231 期间的股价走势



资料来源：Wind，兴业证券经济与金融研究院整理

### 3、科技动量因子有效性验证

在正式开始对科技动量因子进行测试之前，有几点需要明确：

1. 后续所有的研究时间窗口为 2013 年 1 月-2018 年 12 月；
2. 在上一篇专利数据研究报告中，我们发现专利数据具有较强的行业属性，在银行、房地产、非银等金融行业中覆盖度比较低。所以后续的研究中，我们将剔除大金融板块（房地产、银行、非银）；
3. 后续的研究是均基于 IPC 二级分类展开。

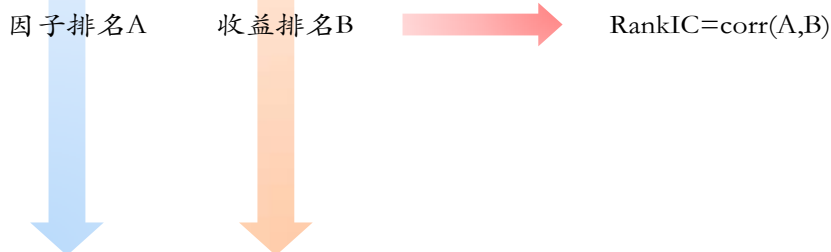
#### 3.1、因子有效性测试的方法

为了对因子的表现进行检验，我们采用 Rank IC (Spearman 秩相关系数) 和分位数组合测试两种方法。两种方法的逻辑参见后续图文说明（图表-16、17）。

##### 1. Rank IC 测试

如果一个因子对股票的预期收益具有预测作用，那么股票当期的因子值与下期股票的收益之间就会存在一定的相关性，我们可以用相关系数来刻画二者之间相关性，从而反映该因子对收益的预测效果。如果计算两者之间的 Pearson 线性相关系数，那么某些异常值可能会对结果产生很大影响。因此，在这里我们将采用更为稳健的 Spearman 秩相关系数来衡量因子的有效性。如下图所示，按照所有股票当期因子值大小进行排序，得到名次序列 A，按照所有股票下期收益大小进行排序，得到名次序列 B，计算序列 A、B 之间的相关系数，即为 Rank IC。Rank IC 的绝对值越大，表明该因子对股票收益的预测能力越强。一般地，我们会统计样本区间内 Rank IC 的均值、标准差和 T-统计量，从预测能力的显著性、稳定性等多个角度分析因子的表现。

图表 16、Rank IC 计算过程

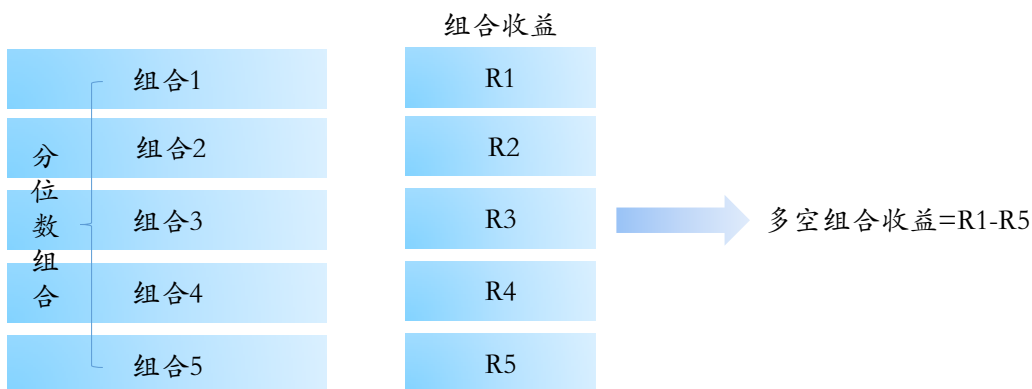


资料来源：兴业证券经济与金融研究院整理

## 2. 分位数组合测试

分位数组合测试是一种较为常用的用来衡量因子有效性的方法，如图表-17所示。首先，按照当期因子值的大小将股票分为 1-5 五个等权的分位组合，组合下期的收益分别为 R1, R2, ..., R5，我们根据多空组合的收益 R1-R5 来判断因子的有效性，如果多空组合收益显著异于零则表明该因子是有效的，组合夏普比率越高，表明因子越有效。但是值得注意的是，由于分位数组合法只考虑了多空两个极端组合的收益，而忽略了中间各分位组合的相关信息，因此对因子有效性的刻画上可能存在一定的局限性。

图表 17、分位数组合法



资料来源：兴业证券经济与金融研究院整理

## 3.2、科技动量因子测试

首先我们观察一下因子 IC 的测试结果。原始因子 IC 徘徊在 0 左右，T 值仅为 0.25。在行业中性化后，科技动量因子的表现得到了系统性提升，T 值高达 2.87，有效性较强。

图表 18、科技动量因子 IC 表现

	平均值	标准差	IC_IR	T 值
原始因子	0.001	0.03	0.03	0.25
行业中性化	0.011	0.03	0.34	2.87

注：市值中性化没有改变因子的表现，这里就不赘述



资料来源：德高行，Wind，兴业证券经济与金融研究院整理

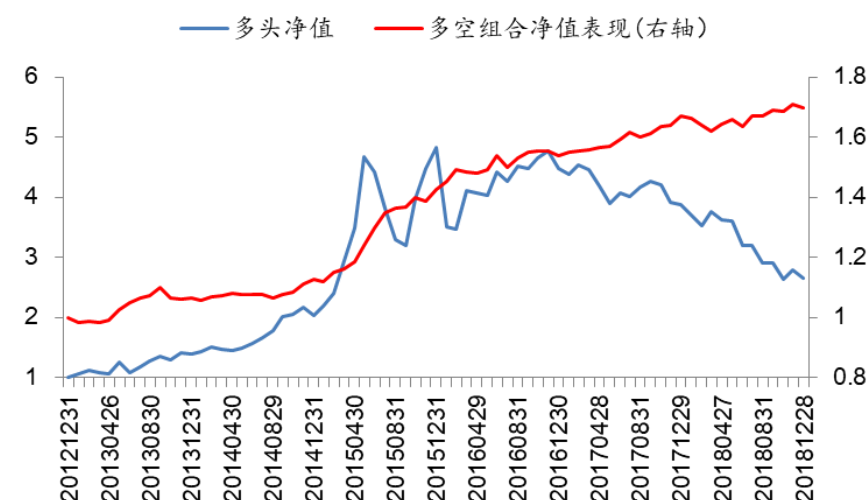
进一步，我们观察该因子的分位数组合测试结果：各组别年化收益率、夏普率严格单调：首组年化收益率高达 17.7%，夏普率达到 0.53，多空年化收益率达到 9.2%，夏普率高达 1.76。

图表 19、科技动量因子分位数表现

	总收益	年化收益	波动率	Sharpe比	平均换手	最大回撤	年化超额	跟踪误差	信息比	胜率
Top	166.1%	17.7%	33.2%	0.53	153.9%	45.7%	5.6%	4.8%	117.1%	58.9%
1	123.8%	14.4%	31.6%	0.45	171.0%	49.3%	2.2%	2.9%	75.8%	54.8%
2	112.6%	13.4%	31.5%	0.43	173.9%	52.7%	1.3%	2.7%	48.4%	58.9%
3	106.1%	12.8%	30.5%	0.42	171.3%	51.2%	0.5%	2.6%	20.2%	49.3%
4	87.9%	11.1%	30.8%	0.36	168.1%	50.6%	-0.9%	3.4%	-27.9%	49.3%
5	86.4%	10.9%	30.5%	0.36	169.6%	51.2%	-1.2%	2.8%	-41.6%	47.9%
6	78.8%	10.2%	30.5%	0.33	173.7%	54.4%	-1.8%	2.8%	-64.0%	47.9%
7	93.1%	11.6%	31.2%	0.37	174.4%	49.8%	-0.4%	2.6%	-14.1%	54.8%
8	75.0%	9.8%	31.5%	0.31	171.7%	54.0%	-1.9%	3.1%	-62.1%	43.8%
Bottom	56.3%	7.7%	32.6%	0.24	154.9%	59.5%	-3.4%	3.4%	-100.2%	35.6%
市场	97.0%	11.5%	31.2%	0.38		51.7%				
L-S	69.8%	9.2%	5.3%	1.76		4.0%				

资料来源：德高行，Wind，兴业证券经济与金融研究院整理

图表 20、科技动量因子多头及多空净值曲线



资料来源：德高行，Wind，兴业证券经济与金融研究院整理

### 3.3、科技动量因子特异性分析

量化选股领域目前面临着选股因子同质化、新因子开发难度加大的窘境。一般情况下，当我们研究一个新的因子，都需要考察该因子与传统因子的共线性、增强等问题，对专利因子的研究同样不例外。为了解决这个问题，我们将首先研究专利因子与传统因子的相关性，并进行 FamaMacBeth 回归测试。同时为了衡量该因子与传统专利因子的差异性，我们将上一篇专利研究报告（）中构建的过往 5 年授权发明专利数因子也纳入了考察范围之内。

传统的选股因子包括：价值、成长、分析师情绪、动量反转、质量、另类因子 6 大类，鉴于质量因子选股作用较弱，另类因子更多的聚焦于价量变动，我们

请务必阅读正文之后的信息披露和重要声明

这里仅研究科技动量因子与其他四类因子以及授权发明专利数因子的相关性。

测试结果表明，科技动量因子与各类因子的相关性非常低，均在 0 左右。而 FamaMacBeth 回归测试结果进一步说明，在加入价值、成长、分析师情绪、动量反转以及专利因子后，科技动量因子的选股有效性依然非常稳定，能够提供显著的增量信息。

**图表 21、科技动量因子与传统因子的相关性**

	专利	成长	价值	情绪	动量
科技动量因子	0.022	0.008	-0.004	0.003	-0.024

资料来源：德高行，Wind，兴业证券经济与金融研究院整理

**图表 22、基于各大类因子的 Fama-MacBeth 测试 T 值**

	专利因子	科技动量因子	成长	价值	情绪	动量反转
综合回归	1.43	3.71	7.09	1.68	7.16	4.14
单独回归	5.26	3.72	6.09	4.62	7.78	4.43

资料来源：德高行，Wind，兴业证券经济与金融研究院整理

至此，我们通过多个维度的测试验证了科技动量因子的选股有效性，后续我们将基于该因子构建选股策略。

## 4、基于科技动量因子的选股策略研究

单独使用科技动量因子的表现在上一章节中已经有了比较充分的探讨，本节我们将主要聚焦于将科技动量与反转因子相结合。从逻辑上讲，科技动量因子捕捉的是个股对与其具有潜在科技关联的公司涨跌反应不充分，股价表现相对滞后所带来的机会，而反转则表现了市场对公司信息的过度反应。两者的结合，首先帮我们筛选出了股价可能被错杀的个股（前期大幅下跌，反转效应强），而后再从这其中挑选科技动量强劲的标的。具体的策略设定如下：

### ➤ 股票池

全市场所有股票，剔除 ST、中信一级行业中的大金融板块（房地产、银行、非银）。同时调仓时保证非涨跌停。

### ➤ 时间窗

2013 年 1 月-2018 年 12 月底

### ➤ 加权方式

等权/流通市值加权

### ➤ 选股策略

月度调仓，同时分层筛选：1、每期分行业选择跌幅最大的 50%的股票；2、进一步分行业选择科技动量因子最大的 20%的股票。

作为基准策略，我们观察该策略与单独的反转策略进行对比。从测试结果来看，无论是等权亦或是流通市值加权，基于科技动量因子和反转因子结合的策略

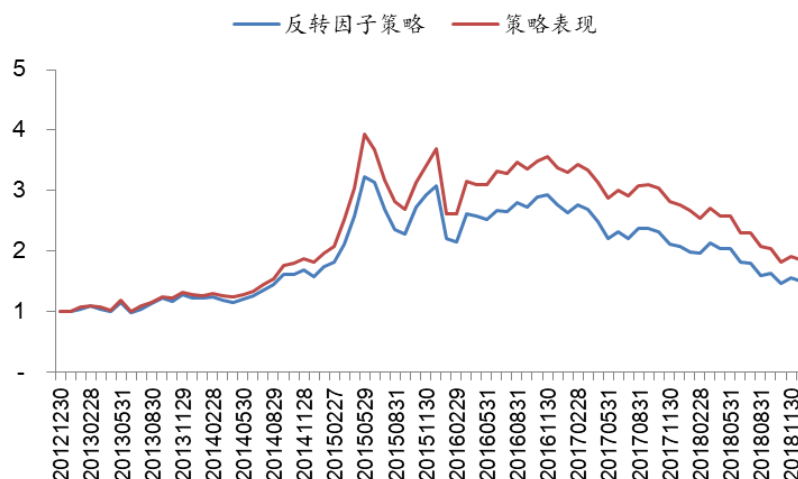
表现远优于单独反转因子的表现。

图表 23、流通市值加权下基于科技动量因子构建的选股策略表现

	总收益率	年化收益	年化波动	Sharpe 比	胜率	最大回撤
反转因子策略	50.2%	7.0%	31.3%	0.22	48.6%	54.4%
合成策略表现	85.7%	10.9%	31.2%	0.35	50.0%	53.6%

资料来源：Wind，德高行，兴业证券经济与金融研究院整理

图表 24、流通市值加权下基于科技动量因子构建的选股策略净值曲线



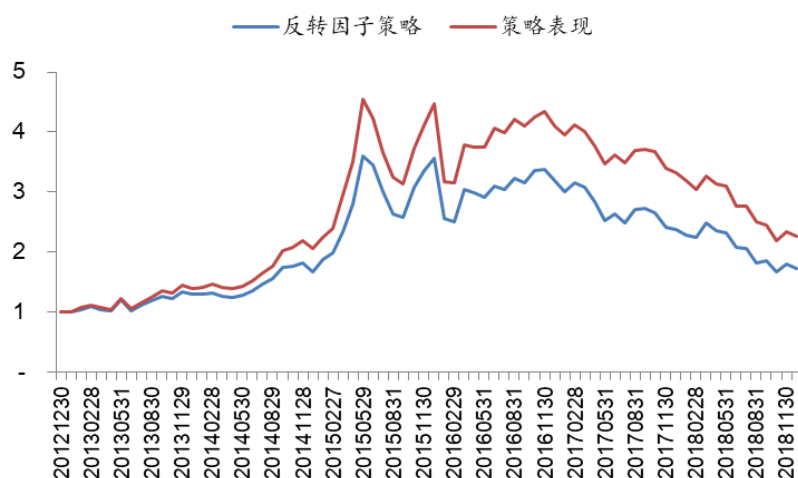
资料来源：Wind，德高行，兴业证券经济与金融研究院整理

图表 25、等权情况下基于科技动量因子构建的选股策略表现

	总收益率	年化收益	年化波动	Sharpe 比	胜率	最大回撤
反转因子策略	71.9%	9.5%	32.2%	0.29	50.0%	53.4%
合成策略表现	126.2%	14.6%	32.3%	0.45	50.0%	51.7%

资料来源：Wind，德高行，兴业证券经济与金融研究院整理

图表 26、等权情况下基于科技动量因子构建的选股策略净值曲线



资料来源：Wind，德高行，兴业证券经济与金融研究院整理

## 5、总结

在专利数据专题研究的第二篇报告中,我们从专利 IPC 分类数据的介绍出发,通过引入科技关联度这一重要概念,构建了科技动量因子。科技动量在与传统因子和之前研究过的普通专利因子保持着很低的相关性的前提下,依旧具有稳定的股票超额收益的预测能力。

专利是一类比较新颖而复杂的另类数据类型。国外在专利的申请、审核、专利的质量把控方面有着较为完善的体制和丰富的研究成果,而我国在这方面起步相对较晚,目前对相关数据的利用还不是太成熟。未来,我们将继续在专利数据领域精耕细作,努力为大家揭开那些隐藏在深处的规律、特征,以便更好地服务于大家的投资与研究工作。

## 6、参考文献

[1].Charles M.C. Lee, Stephen Teng Sun,Rongfei Wang, and Ran Zhang (2018).Technological Links and Predictable Returns. *Journal of Financial Economics*, Mar 13 2018

[2].徐寅,郑兆磊 (2018). 从中国心到中国芯--贸易战引发的专利选股有效性研究, 兴业证券经济与金融研究院, Dec 20 2018



## 7、附录

图表 27、专利因子定义

	概念	定义
存量	TP101	统计截止日前所有的有效发明专利的专利数（存量数据）
	TP010201	统计截止日前三年内新生成的有效发明专利的专利数
所有增量维度的专利因子	TP010202	统计截止日前三年内新生成的有效发明专利的说明书总字数
	TP010203	统计截止日前三年内新生成的有效发明专利的权利要求总项数
	TP010204	统计截止日前三年内新生成的有效发明专利的独权总项数
	TP010205	统计截止日前三年内新生成的有效发明专利的说明书的附图总数
	TP010206	统计截止日前三年内新生成的有效发明专利的摘要总字数
	TP010207	统计截止日前三年内新生成的有效发明专利的 IPC 分类号总数
	TP010208	统计截止日前三年内新生成的有效发明专利的寿命加总
	TP010401	统计截止日前三年内新生成的有效发明专利授权的专利数
	TP010402	统计截止日前三年内新生成的有效发明专利授权的说明书总字数
	TP010403	统计截止日前三年内新生成的有效发明专利授权的权利要求总项数
	TP010404	统计截止日前三年内新生成的有效发明专利授权的独权总项数
	TP010405	统计截止日前三年内新生成的有效发明专利授权的说明书的附图总数
	TP010406	统计截止日前三年内新生成的有效发明专利授权的摘要总字数
	TP010407	统计截止日前三年内新生成的有效发明专利授权的 IPC 分类号总数
	TP010408	统计截止日前三年内新生成的有效发明专利授权的寿命加总
	TP010409	统计截止日前三年内新生成的有效发明专利授权的专利审查期加总
	TP020401	统计截止日前三年内新生成的有效实用新型的专利数
	TP020402	统计截止日前三年内新生成的有效实用新型的说明书总字数
	TP020403	统计截止日前三年内新生成的有效实用新型的权利要求总项数
	TP020404	统计截止日前三年内新生成的有效实用新型的独权总项数
	TP020405	统计截止日前三年内新生成的有效实用新型的说明书的附图总数
	TP020406	统计截止日前三年内新生成的有效实用新型的摘要总字数
	TP020407	统计截止日前三年内新生成的有效实用新型的 IPC 分类号总数
	TP020408	统计截止日前三年内新生成的有效实用新型的寿命加总
	TP030401	统计截止日前三年内新生成的有效外观设计的专利数
	TP030408	统计截止日前三年内新生成的有效外观设计的寿命加总
	TP040401	统计截止日前三年内新生成的有效授权专利总数
	TP040410	统计截止日前三年内新生成的有效授权专利的前引总数
	TP040411	统计截止日前三年内新生成的有效授权专利的后引总数
	TP040412	统计截止日前三年内新生成的有效授权专利的非专利类前引总数
	TP040413	统计截止日前三年内新生成的有效授权专利的海外同族专利总数
	TP010411	统计截止日前三年内新生成的有效发明专利授权专利的后引总数

资料来源：德高行，兴业证券经济与金融研究院整理

图表 28、专利重要概念汇总

概念	定义
发明公开	发明专利公布指的是形式审查通过之后，发了受理通知书之后的一个流程，发明公布一般是从申请日开始四到五个月时间，公布之后进入实质审查阶段，进入实质审查阶段以后就排队等着审查员审查
权利要求	权利要求书是申请发明专利的和申请实用新型专利的必须提交的申请文件。它是发明或者实用新型专利要求保护的内容，具有直接的法律效力，是申请专利的核心，也是确定专利保护范围的重要法律文件
独权要求	是指无需用其他权利要求来确定其范围和含义的完整权利要求。独立权利要求应从整体上反映发明或实用新型的主要技术内容
IPC 分类号	是目前惟一国际通用的专利文献分类和检索工具，为世界各国所必备；PCT 申请情况能够反应专利法律地域保护范围，说明专利在世界范围内的影响力
专利寿命	如果专利取得授权，那么自申请日起至保护期结束的时间
前引	一个专利引用在申请时，引用的其他专利

请务必阅读正文之后的信息披露和重要声明

后引	该专利被其他专利引用的次数
非专利前引引用	该专利引用的其他文献数
非专利后引	该专利被其他文献引用的次数
海外同族专利	基于同一优先权文件，在不同国家或地区，以及地区间专利组织多次申请、多次公布或批准的内容相同或基本相同的一组专利文献

资料来源：国家专利局，兴业证券经济与金融研究院整理

风险提示：本报告模型及结论全部基于对历史数据的分析，当市场环境变化时，存在模型失效风险。

## 分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，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。

## 投资评级说明

投资建议的评级标准	类别	评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级(另有说明的除外)。评级标准为报告发布日后的12个月内公司股价(或行业指数)相对同期相关证券市场代表性指数的涨跌幅。其中：A股市场以上证综指或深圳成指为基准，香港市场以恒生指数为基准；美国市场以标普500或纳斯达克综合指数为基准。	股票评级	买入	相对同期相关证券市场代表性指数涨幅大于15%
		审慎增持	相对同期相关证券市场代表性指数涨幅在5%~15%之间
		中性	相对同期相关证券市场代表性指数涨幅在-5%~5%之间
		减持	相对同期相关证券市场代表性指数涨幅小于-5%
		无评级	由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级
	行业评级	推荐	相对表现优于同期相关证券市场代表性指数
		中性	相对表现与同期相关证券市场代表性指数持平
		回避	相对表现弱于同期相关证券市场代表性指数

## 信息披露

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。

## 使用本研究报告的风险提示及法律声明

兴业证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告仅供兴业证券股份有限公司(以下简称“本公司”)的客户使用，本公司不会因接收人收到本报告而视其为客户。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本报告所载资料的来源被认为是可靠的，但本公司不保证其准确性或完整性，也不保证所包含的信息和建议不会发生任何变更。本公司并不对使用本报告所包含的材料产生的任何直接或间接损失或与此相关的其他任何损失承担任何责任。

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的表现依据；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。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现。过往的业绩表现亦不应作为日后回报的预示。我们不承诺也不保证，任何所预示的回报会得以实现。分析中所做的回报预测可能是基于相应的假设。任何假设的变化可能会显著地影响所预测的回报。

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门、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兴业证券股份有限公司及其关联子公司等违反当地的法律或法规或可致使兴业证券股份有限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。

## 特别声明

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。

## 兴业证券研究

上 海	北 京	深 圳
地址：上海浦东新区长柳路36号兴业证券大厦15层	地址：北京西城区锦什坊街35号北楼601-605	地址：深圳市福田区皇岗路5001号深业上城T2座52楼
邮编：200135	邮编：100033	邮编：518035
邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn