

专题报告

金融工程

数量化投资

数量化投资技术系列报告之三十一

2010年9月6日

本报告的独到之处

- 区别于传统多因子的因子贡献度模型
- 中证 800 成分股的证监会行业分析
- 行业内的区分度动量选股

专题报告

行业内区分度动量选股探讨

我们之前已经陆续推出三篇关于因子贡献度和区分度动量的报告（《多因子 alpha 选股——将行业轮动落实到 Top 组合》、《多因子选股模型结合股指期货实现 Alpha 收益》和《将区分度动量模型延伸至三维空间》），依次研究了以中证 800 和沪深 300 为股票池的区分度动量选股，并且通过剥离价格，将因子集分类为基本面因子和市场因子。

但是在不同的样本股票池中，因子集的持续表现是否一致？如果不一致，那么是否所有的股票池样本集都能用这种方法，股票池对应的因子集又该如何选择。

最常见的样本子集，就是行业了。按照至上而下的投资逻辑，首先判断宏观经济，根据所处的经济周期位置进行行业配置。那么从量化的结构性角度出发，行业内的因子有什么特性，又该用哪些因子选股呢？这些便是本篇报告希望解决的问题。

在对行业内量化选股面临的难点进行了初步探讨之后，我们决定构建一套因子贡献度的衡量体系，即从**持续性、显著性和偏离性**三个角度去衡量行业内因子集的表现。通过对中证 800 成分股证监会行业特性的分析，最终得出**交运仓储、批发零售、房地产、石化、金属非金属、机械设备和医药生物等行业**可能适用于区分度动量选股。

进一步我们设计了通过对因子贡献度衡量体系加权打分的方式选择行业内因子集的方法，并以石化和机械设备行业为例，给出了行业内区分度动量选股策略的实证分析。

通过对实证结果的分析，发现相对全市场多因子选股而言，由于行业内信息扩散速度更快，投资者挖掘更为深入，使得行业内选股难度更大，区分度动量的持续时间要短于全市场。

在后续的研究工作中，将重点对适合量化选股的行业进行逐一分析，以线性模型的区分度动量选股为基础，针对行业特征进行选股方法和行业参数的进一步优化。

分析师：董艺婷

电话：021-60933155

E-mail: dongyt@guosen.com.cn

SAC 执业证书编号：S09802100030026

联系人：程景佳

电话：021-60933166

E-mail: chengjia@guosen.com.cn

分析师：葛新元

电话：0755-82133332

E-mail: gexy@guosen.com.cn

SAC 执业证书编号：S0980200010107

独立性声明：

作者保证报告所采用的数据均来自合规渠道，分析逻辑基于本人的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

内容目录

区分度动量扩展至行业子集	4
背景介绍	4
问题的提出	5
行业内多因子量化选股失效的可能性探讨	5
基于因子贡献度的行业分析	6
中证 800 的证监会行业子集	6
行业内考察的因子集	8
行业内因子贡献度衡量体系	9
行业内区分度动量选股策略	10
适合多因子选股的行业筛选	10
行业内区分度动量选股示例	15
思考及后续研究	16
对行业内因子分析的思考	16
一个有趣的现象	17

图表目录

图 1: 因子贡献度定义	4
图 2: 行业内的因子贡献度特性	5
图 3: 行业内多因子量化选股失效原因分析	6
表 1: 中证 800 证监会行业统计 (不含制造业)	7
表 2: 中证 800 制造业行业统计	7
图 4: 中证 800 证监会行业自由流通市值占比	8
表 3: 基本面因子集	8
表 4: 估值因子集	9
图 5: 行业内因子贡献度衡量指标	10
表 6: 证监会行业 (不含制造业) 基本面因子集分析	11
表 7: 证监会行业 (不含制造业) 估值因子集与市场因子集分析	12
表 8: 证监会制造业子行业基本面因子集分析	13
表 9: 证监会制造业子行业估值因子集与市场因子集分析	13
表 10: 证监会行业股票数目与因子集持续性对比	15
表 11: 石化行业内区分度动量选股效果	16
表 12: 机械设备行业内区分度动量选股效果	16

区分度动量扩展至行业子集

背景介绍

我们之前已经陆续推出三篇关于因子贡献度和区分度动量的报告（《多因子 alpha 选股——将行业轮动落实到 Top 组合》、《多因子选股模型结合股指期货实现 Alpha 收益》和《将区分度动量模型延伸至三维空间》）。模型的整体思路是用贡献度来衡量因子的区分强度，半衰期来跟踪区分度动量的持续时间。

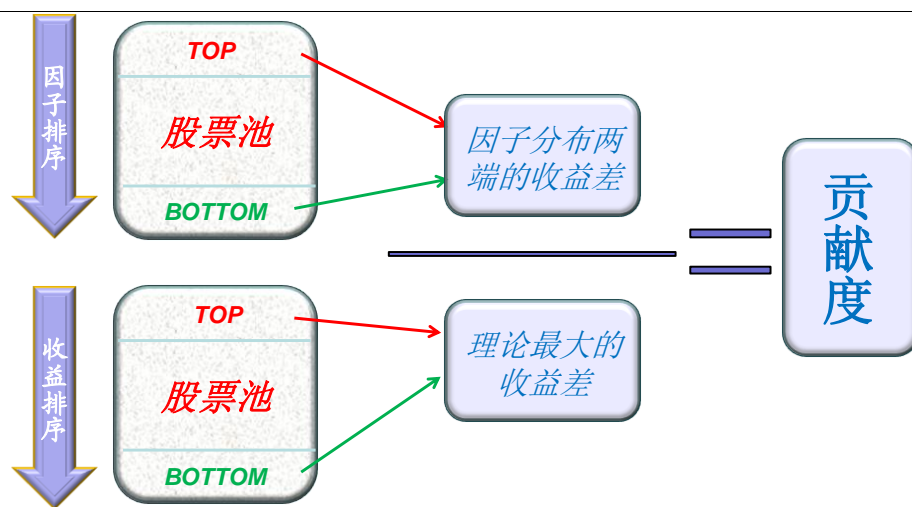
因子贡献度是指股票池按照因子排序前 20% 和后 20% 的两个组合，在观察期内的累计收益差，与理论最大收益差（即观察期内区间收益率最高的 20% 和最低的 20% 两个组合的收益差）的比值。因此因子贡献度是一个标准化的量，没有量纲，且在 [-1, 1] 区间内。

$$\begin{aligned} CON_F &= \frac{\frac{\text{mean}[R_i] - \text{mean}[R_j]}{i \in SET_{low-F} \quad j \in SET_{high-F}}}{\frac{\text{mean}[R_i] - \text{mean}[R_j]}{i \in SET_{high-R} \quad j \in SET_{low-R}}} \\ &= \frac{\sum_{i \in SET_{low-F}} R_i - \sum_{j \in SET_{high-F}} R_j}{\sum_{i \in SET_{high-R}} R_i - \sum_{j \in SET_{low-R}} R_j} \end{aligned}$$

在模型的应用部分，我们构建了一个 FACTOR-SCORE 的打分模型，将贡献度向量和排名向量的内积作为单只股票的截面分数，并选出分数最小的 20% 股票作为 TOP 组合，利用多因子半衰期作为持仓期限，即**多因子半衰期选股策略**。以中证 800 成分股为全样本的历史回溯检验中，该策略有较为稳定的超额收益。

进一步，我们对因子集进行了分类，分离成基本面因子与市场因子两个维度，将整个模型框架扩展至时间-基本面-市场的三维结构，从而解决了因子集的共线性问题。

图 1：因子贡献度定义



资料来源：国信证券经济研究所

问题的提出

可以看到因子贡献度纯粹是一个度量的工具，它反映的市场上已经真实发生的信息。在之前的报告中，我们的逻辑是假设这种因子的区分度或者说市场对因子的关注度是可以持续的，然后利用这种持续性去进行选股。

之前的工作都是以中证 800 或者沪深 300 成分股为样本股票池。但是在不同的样本股票池中，因子集的持续表现是否一致？如果不一致，那么是否所有的股票池样本集都能用这种方法，股票池对应的因子集又该如何选择，这些便是本篇报告希望解决的问题。

最常见的样本子集，就是行业了。按照至上而下的投资逻辑，首先判断宏观经济，根据所处的经济周期位置进行行业配置。那么从量化的结构性角度出发，行业内的因子有什么特性，又该用哪些因子选股呢？由于公司的行业性差异，可能在一个行业非常重要的财务指标，在另一个行业则不受关注。行业的结构性差异，还会影响到因子的持续性，有一些的行业动量效应明显，另一些行业的股票可能呈现的是反转效应，还有一些行业可能根本无法用量化指标来进行选股。

图 2：行业内的因子贡献度特性



资料来源：国信证券经济研究所

因此我们觉得决定利用现有的因子贡献度体系，对行业进行一次梳理，初步筛选出适合用多因子方法选股的行业，以及相应有效的因子集。这样的行业梳理工作完成后，不仅可以使行业内半衰期选股策略，也可以为其他基于因子的行业内选股策略提供参考。

行业内多因子量化选股失效的可能性探讨

在之前的报告中我们以中证 800 成分股为股票池，对区分度动量模型进行了历史回溯检验，各观察期都能稳定战胜中证 800 指数。同时在每周的《正 alpha 策略周报》中也会给出根据策略选出的 160 只股票组合，样本外区间都取得了不错的超额收益。

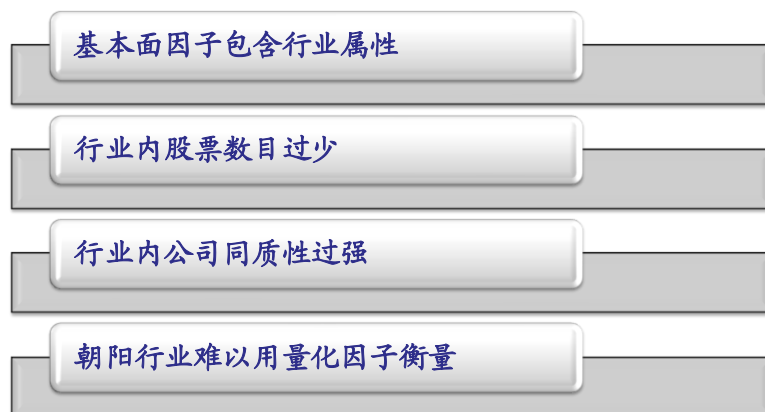
但是当尝试还是用相同的方法和因子集去作用到行业子集上时，结果却不如意，行业间产生了较大的差异，制造业中石油化工、机械设备等行业仍然能取得超

额收益，而像金融服务行业则完全失效，累计收益明显低于基准。

造成这种结果的原因，我们分析可能源于以下几点：

- (1) 以中证 800 成分股作为股票池，虽然基本面因子本身不是用来选行业，但其本身可能就包含了行业信息。比如有些行业在财务指标上具有很强的共性，此时基本面因子就具有一定的行业属性；但是在行业内部，这些基本面因子就失去了指导意义。
- (2) 上市公司的行业分布不均匀，导致有些行业股票数量过少。当一个行业股票数量不超过 10 只时，20%就只有 2 只，此时个股的异常波动会极大的影响整个策略的效果，而量化策略更多是基于概率。
- (3) 有个行业内公司的同质性强，股价有平稳回归的特性，此时反应出来的往往是反转效应而非动量效应，这一类行业更适合配对交易这样的策略，而不太适合用多因子模型进行选股。
- (4) 对于朝阳行业而言，如新能源，主动投资者关注的是上市公司的业务发展方向、产业布局、扩展潜力等等，这些很难从财务报表中体现出来，此时用基本面因子集选股难以获得理想的效果。

图 3：行业内多因子量化选股失效原因分析



资料来源：国信证券经济研究所

基于因子贡献度的行业分析

中证 800 的证监会行业子集

目前市场上存在多种行业分类方法，常见的有证监会行业分类、GICS 行业分类、券商行业分类（如国信行业分类、申万行业分类等）和一些资讯公司提供的行业分类（如 wind 行业分类）。考虑到规范性和通用性，本篇报告皆采用证监会的行业分类方法。

证监会在 2001 年 4 月 3 日公布了《上市公司行业分类指引》，正式规定了上市公司分类的原则、编码方法、框架及其运行与维护制度。当公司某类业务的营业收入比重大于或等于 50%，则将其划入该业务相对应的类别。当公司没有一类业务的营业收入比重大于或等于 50% 时，如果某类业务营业收入比重比其他业务收入比重均高出 30%，则将该公司划入此类业务相对应的行业类别；否则，将其划为综合类。

按照证监会行业分类，共分为 13 个一级行业，其中制造业又下分为 10 个二级子行业。最终我们使用 22 个行业分类进行划分。下面我们列出了这 22 个行业在中证 800 成分股中的数目以及自由流通市值权重占比。由于有些股票已经退市找不到相应的行业分类，导致在早期行业数目之和不足 800，但影响不大。

表 1: 中证 800 证监会行业统计 (不含制造业)

代码	A	B	D	E	F	G	H	I	J	K	L	M
对应行业	农林牧渔	采掘	电力煤气及水	建筑	交通运输仓储	信息技术	批发和零售贸易	金融保险	房地产	社会服务	传播文化产	综合类
股票数目 (最新)	10	37	38	22	48	34	61	32	59	14	10	12
股票数目 (平均)	12	29	43	17	44	38	69	19	55	16	9	15
行业权重 (最新, %)	0.8	8.13	3.51	3.06	4.43	3.74	5.86	22.34	6.18	1.23	0.73	0.99
行业权重 (平均, %)	0.72	6.84	5.92	1.42	6.79	4.77	5.95	16.19	6.75	1.51	1.05	1.10

资料来源: wind 资讯

上表为除制造业外的 11 个证监会行业，在中证 800 成分股内的统计状况。统计区间由 2005 年 1 月 5 日至 2010 年 8 月 13 日。可以看到农林牧渔、文化传播和综合类在中证 800 里都只有 10 只左右。金融保险相对特殊一下，在 07 年之前只有 11 只，而 07 年至今随着金融股的大量上市已扩大至 32 只。

再来看一下制造业的统计状况，从数目上来看，中证 800 中木材家具和其他制造业的数量都只有个位数，而机械设备、石油化工和金融非金属为股票数量最多的三个二级行业，甚至超过了其他的一级行业，这也侧面印证了制造业需要考虑子行业的合理性。

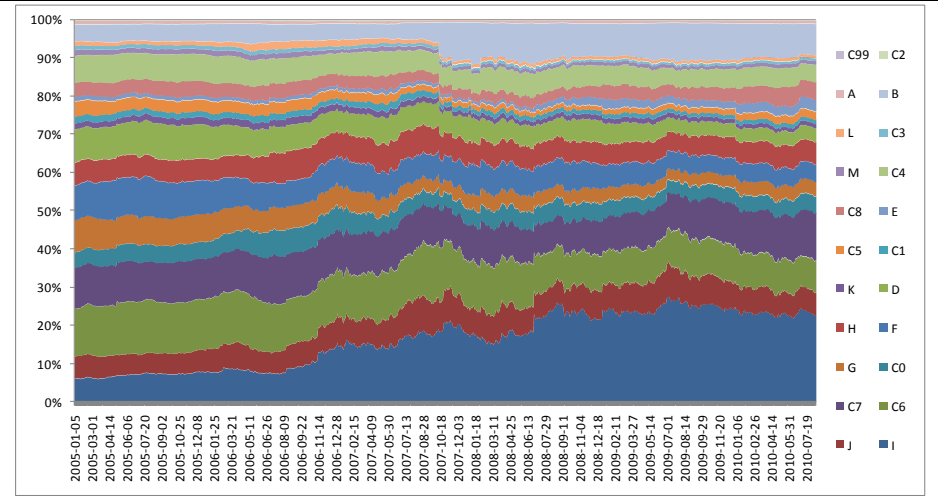
表 2: 中证 800 制造业行业统计

代码	C0	C1	C2	C3	C4	C5	C6	C7	C8	C99
对应行业	食品、饮料	纺织、服装、皮毛	木材家具	造纸印刷	石油化学塑胶塑料	电子	金属非金属	机械设备仪表	医药生物制品	其他制造业
股票数目 (最新)	32	18	3	13	75	33	78	117	51	2
股票数目 (平均)	32	20	3	11	85	32	82	112	45	1
行业权重 (最新, %)	4.53	1.11	0.15	0.68	4.54	2.12	8.23	12.41	4.71	0.08
行业权重 (平均, %)	4.54	1.42	0.16	0.77	5.76	2.06	11.23	10.18	3.45	0.04

资料来源: wind 资讯

最后关注这 22 个行业在中证 800 中的自由流通市值占比。目前占比最高的是金融保险业，占到了 22.34%，较 05 年提升了数倍。07 年中石油进入中证 800，对采掘业的权重占比产生了直接影响。实际上这种一只股票影响整个行业的情形，也会对行业内的量化选股造成干扰。

图 4: 中证 800 证监会行业自由流通市值占比



资料来源: wind 资讯

行业内考察的因子集

从上面的统计可以看到, 中证 800 的行业股票数量分布很不均衡, 此时因子贡献度如果依然用两端的 20% 来计算, 会导致度量的组合只有 2、3 只股票, 放大了个股的不确定性, 导致因子度量出现偏差。

因此我们在计算行业内的因子贡献度时, 对其定义进行了调整, 由之前的两端 20% 变为 50%, 即将行业内的股票按因子排名分成两半, 看这两部分的收益差与理论最大收益差之比。

结合之前的三篇报告, 我们考虑了基本面因子、估值因子和市场因子三类指标, 共 19 个因子, 如下表所列:

表 3: 基本面因子集

因子名称	说明	因子意义
流通市值	最新季报日流通市值	市值规模
净资产	最新季报股东权益合计	资产规模
净资产增长率	过去两年净资产季报数据回归	经营能力
净资产收益率	最新年报 ROE	盈利能力
每元净资产经营现金流	过去十二个月经营现金流/净资产	现金流水平
每元净资产营业收入	过去十二个月营业收入/净资产	收入水平
净利润增长率	过去两年净利润季报数据回归	成长能力
股息率	最新年报股息率	投资收益
资产负债率	最新年报资产负债率	偿债能力

资料来源: 国信证券经济研究所

表 4：估值因子集

因子名称	说明	因子意义
市盈率	市值/过去十二个月净利润	估值水平
市净率	市值/最新季报股东权益合计	估值水平
市现率	市值/过去十二个月经营现金流	估值水平
市销率	市值/过去十二个月营业收入	估值水平

资料来源：国信证券经济研究所

表 5：市场因子集

因子名称	说明	因子意义
价格动量	过去 60 个交易日累计收益率	市值规模
换手率	过去 60 个交易日平均换手率	资产规模
波动率	过去 60 个交易日收益率标准差	经营能力
总市值	最新总市值	市值规模
股价	最新股票价格	单位价格
机构持股	机构持股/流通股本	机构参与

资料来源：国信证券经济研究所

行业内因子贡献度衡量体系

确定了这 19 个因子后，我们可以计算每个行业内各个因子的历史因子贡献度表现。接下来的问题是，如何根据得到的贡献度去分析行业和筛选因子。回顾区分度动量选股策略的逻辑，是希望观察到因子区分度的动量，并利用半衰期跟踪动量的衰竭程度，因此选择的因子集应该具备以下两点：

（1） 因子区分度应当显著

因子贡献度实际是反映在观察期内，投资者在关注什么因子，从而导致了股票收益率的差异。因子贡献度的绝对值越大，说明该因子的区分效果越少，从而根据因子选出的股票组合可能获得较高的超额收益。

（2） 区分度的动量效应有较好的持续性

因为我们采用的动量策略，半衰期跟踪动量的持续时间，动量持续性越好，说明因子作用越稳定；而持续性差的话，会导致换仓频率的增加，甚至可能面临观察期刚观测到因子的区分度，动量便衰竭的问题。

在报告《将区分度动量模型延伸至三维空间》中将国信资金强弱 GSMS 值也作为一个市场因子，发现它较为稳定的偏向一侧，这种贡献度偏向一侧的因子隐含的规律性对量化策略也有很大的意义。因此除了以上两点，还应该把因子贡献度的偏离状况也作为衡量体系中的一个指标。

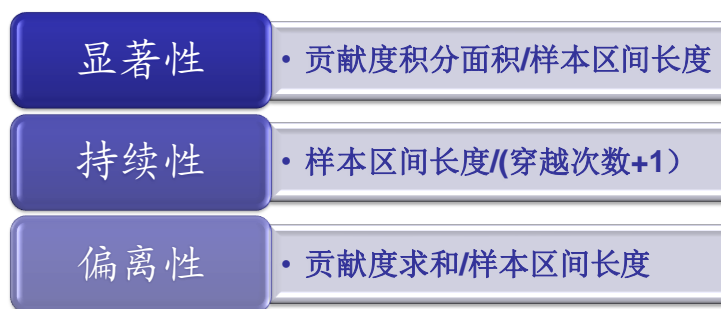
最终，我们定义了因子贡献度的**显著性、持续性和偏离性**做为衡量体系。

（1） **显著性**：用因子贡献度绝对值的平均值来衡量，也即样本区间内的因子贡献度积分面积除以区间长度。由于考察的因子取相同的样本区间，因

此实际是比较积分面积。

- (2) **持续性**: 用因子贡献度在 x 轴一侧持续的平均时间来衡量, 计算方法为样本区间长度除以贡献度曲线与 x 轴交叉次数加 1。由于考察的因子取相同的样本空间, 因此实际是比较贡献度改变符号的次数 (即因子方向的变化频率, 因子贡献度在坐标轴一侧持续的时间越长, 则由此选出的股票组合获得的超额收益越稳定)。
- (3) **偏离性**: 用因子贡献度的平均值来衡量。如果某个因子长期能稳定的偏向一侧, 则该因子的选股效果将更加可靠。但需要注意的是, 这种偏离可能只是一种统计规律, 如果不能找到产生偏离的本质逻辑, 随着市场结构的变化, 在未来使用相同的因子和方向是非常危险的做法。

图 5: 行业内因子贡献度衡量指标



资料来源: 国信证券经济研究所

行业内区分度动量选股策略

适合多因子选股的行业筛选

第二部分中, 我们构建了因子贡献度的衡量体系, 从显著性、持续性和偏离性三个方面在行业内来衡量因子。

由于股改的原因, 06 年大量股票停牌, 有些行业本身股票数量就少, 会出现同一时间超过一半股票停牌的极端情形, 此时因子贡献度的计算也会出现异常。因此在行业内统计时, 我们选择的样本区间为 2007 年 1 月 4 日至 2010 年 8 月 13 日。

由于木材家具和其他制造业在中证 800 中的股票数量过少, 因此最终只考虑了剩下的 20 个行业 (11 个一级行业和 8 个制造业二级行业)。各个行业的基本因子集、估值因子集和市场因子集的分析数据如下面 4 张表所示:

表 6: 证监会行业（不含制造业）基本面因子集分析

基本面因子集

证监会行业	指标	<div> <div>每元净资产经</div> <div>每元净资产营</div> </div>								
		流通市值	净资产	净资产增长率	净资产收益率	管现金流	业收入	净利润增长率	股息率	资产负债率
A	均值	0	0.04	-0.03	0.13	-0.02	0.07	0.08	0.12	0.11
	绝对值均值	0.28	0.36	0.36	0.39	0.32	0.37	0.35	0.36	0.34
	平均持续时间	9.8	13.78	21	18.77	14.46	22.62	14.46	14.7	15.47
B	均值	0.07	0.1	0.02	-0.01	-0.07	-0.01	0.05	-0.14	0.01
	绝对值均值	0.25	0.31	0.19	0.25	0.22	0.21	0.3	0.31	0.2
	平均持续时间	15.75	17.64	11.03	16.96	10.38	11.45	25.94	27.56	18.38
D	均值	0.14	0.14	0.04	0.06	-0.05	0.04	0.09	0.02	0.03
	绝对值均值	0.27	0.32	0.17	0.24	0.21	0.16	0.18	0.17	0.18
	平均持续时间	19.17	18.38	10.89	12.08	16.33	9.48	11.31	15.47	12.25
E	均值	0.1	0.08	0.14	0.02	0.11	0.06	0.08	-0.04	0.05
	绝对值均值	0.33	0.34	0.28	0.34	0.26	0.3	0.32	0.33	0.35
	平均持续时间	20.05	11.31	12.08	25.94	14.23	14.7	26.73	21	22.62
F	均值	0.13	0.15	0	0.06	-0.05	-0.03	-0.04	0.04	0.05
	绝对值均值	0.26	0.31	0.17	0.23	0.2	0.2	0.21	0.18	0.16
	平均持续时间	18.38	25.94	13.36	19.17	13.57	13.16	21	15.75	15.47
G	均值	0.01	-0.04	0.06	-0.02	-0.04	-0.01	-0.02	-0.04	-0.06
	绝对值均值	0.21	0.18	0.27	0.21	0.17	0.17	0.16	0.21	0.17
	平均持续时间	11.76	12.78	20.05	14.46	11.76	12.42	9.91	12.6	12.97
H	均值	0.09	0.05	-0.01	-0.03	-0.07	-0.03	-0.03	-0.01	-0.04
	绝对值均值	0.2	0.14	0.14	0.25	0.28	0.19	0.18	0.15	0.18
	平均持续时间	23.84	14.95	10.38	20.51	30.41	14.95	19.6	17.29	17.64
I	均值	0.03	0.04	-0.05	-0.05	0.03	-0.03	-0.07	0.01	0.01
	绝对值均值	0.28	0.29	0.29	0.27	0.27	0.29	0.32	0.22	0.43
	平均持续时间	11.31	10.5	16.33	18	11.31	12.78	11.31	12.97	12.08
J	均值	0.07	0.06	0.03	0	0	0	0.04	0.02	0.01
	绝对值均值	0.27	0.25	0.22	0.22	0.14	0.17	0.16	0.17	0.17
	平均持续时间	33.92	16.96	12.25	17.29	12.6	12.97	14.95	25.94	16.64
K	均值	0.06	0.04	-0.05	0.04	0.01	-0.05	-0.03	0.05	-0.02
	绝对值均值	0.35	0.35	0.2	0.32	0.23	0.31	0.32	0.23	0.25
	平均持续时间	14.7	13.36	12.25	21.51	11.45	18	16.64	13.57	14.46
L	均值	0.07	0.13	-0.04	-0.01	0.03	0.02	0.12	0	-0.08
	绝对值均值	0.34	0.37	0.45	0.44	0.27	0.29	0.31	0.38	0.44
	平均持续时间	22.05	14.46	17.29	17.29	9.59	9.48	12.6	18.38	18.38
M	均值	0.01	-0.1	-0.01	-0.03	-0.03	0.05	0	-0.16	0.21
	绝对值均值	0.27	0.27	0.3	0.36	0.23	0.23	0.27	0.31	0.31
	平均持续时间	11.16	10.26	14.23	18.38	10.26	9.59	11.03	15.75	15.75

资料来源: tinysoft, 国信证券经济研究所

请务必阅读正文之后的免责条款部分

全球视野 本土智慧

表 7: 证监会行业（不含制造业）估值因子集与市场因子集分析

证监会行业	指标	估值因子集				市场因子集					
		市盈率	市净率	市销率	市现率	换手率	价格动量	波动率	股价	机构持股	总市值
A	均值	0.24	0.11	0.01	0.02	0.05	-0.11	0	0.13	0.08	0.06
	绝对值均值	0.35	0.33	0.35	0.36	0.37	0.43	0.35	0.36	0.37	0.34
	平均持续时间	14.95	11.61	18.38	23.21	10.76	21.51	7.95	10.02	15.47	12.25
B	均值	0.1	0.14	0.13	0.15	-0.06	-0.09	0.02	0.21	-0.04	0.15
	绝对值均值	0.26	0.31	0.22	0.24	0.22	0.18	0.26	0.3	0.29	0.28
	平均持续时间	14.23	18.77	16.96	11.45	11.92	12.08	11.45	19.6	16.64	15.75
D	均值	0.05	0.09	0.01	0.03	-0.01	0.04	0.02	0.18	0.09	0.17
	绝对值均值	0.19	0.2	0.19	0.18	0.26	0.2	0.27	0.24	0.26	0.29
	平均持续时间	13.78	11.92	12.25	9.38	15.21	20.05	12.25	13.36	17.64	15.21
E	均值	0.16	0.2	0.06	0.03	0.11	-0.1	0.01	0.31	0.12	0.25
	绝对值均值	0.28	0.32	0.31	0.23	0.32	0.31	0.34	0.41	0.29	0.4
	平均持续时间	12.08	13.16	12.42	13.57	9.48	19.6	12.6	17.64	16.04	16.96
F	均值	0.1	0.17	0.13	0.12	0.07	-0.02	0.04	0.22	0.1	0.19
	绝对值均值	0.18	0.22	0.21	0.22	0.2	0.15	0.25	0.28	0.25	0.31
	平均持续时间	10.02	12.78	16.04	16.04	11.92	14.46	11.03	21	16.64	18.38
G	均值	0.08	0.18	0.08	0.07	0.09	0.01	0.1	0.12	0.09	0.05
	绝对值均值	0.19	0.24	0.2	0.17	0.24	0.18	0.22	0.26	0.24	0.24
	平均持续时间	11.03	10.89	11.92	12.97	14.46	12.97	10.02	12.25	14.7	12.08
H	均值	0.16	0.14	0.11	0.02	0.1	-0.04	0.07	0.17	0.02	0.1
	绝对值均值	0.22	0.21	0.18	0.12	0.24	0.13	0.25	0.26	0.26	0.24
	平均持续时间	20.51	12.78	32.67	11.45	13.57	11.16	15.21	23.21	22.62	22.62
I	均值	0.11	0.13	0.1	0.03	0.04	-0.03	0.01	0.14	0.08	0.13
	绝对值均值	0.29	0.35	0.3	0.29	0.38	0.24	0.32	0.36	0.32	0.34
	平均持续时间	9.09	11.31	9.69	15.75	13.36	12.08	10.38	17.64	13.57	11.61
J	均值	0.1	0.19	0.11	-0.03	0.02	-0.01	0.07	0.24	0.06	0.14
	绝对值均值	0.17	0.23	0.2	0.13	0.18	0.14	0.19	0.27	0.29	0.26
	平均持续时间	10.14	16.64	16.64	11.76	12.6	12.6	9.38	13.57	18.77	16.04
K	均值	0.17	0.27	0.2	0.16	0.01	0.03	0.11	0.15	0	0.12
	绝对值均值	0.27	0.33	0.33	0.3	0.33	0.28	0.36	0.37	0.31	0.38
	平均持续时间	11.16	11.45	12.78	13.57	13.78	14	14	18.38	15.47	20.05
L	均值	0.16	0.17	0.03	0.12	-0.06	0.03	-0.06	0.17	-0.02	0.21
	绝对值均值	0.39	0.38	0.32	0.32	0.42	0.44	0.43	0.33	0.45	0.36
	平均持续时间	11.16	13.57	9.38	12.25	14.46	20.05	12.25	9.91	11.92	12.97
M	均值	-0.04	0.14	-0.01	-0.03	0.23	-0.04	0.14	0.2	0	0.05
	绝对值均值	0.28	0.26	0.23	0.33	0.32	0.3	0.33	0.37	0.28	0.26
	平均持续时间	10.02	10.5	10.14	15.47	10.26	14.95	12.6	15.75	9.28	8.82

资料来源: tinysoft, 国信证券经济研究所

表 8：证监会制造业子行业基本面因子集分析
基本面因子集

证监会行业	指标	<div> <div>每元净资产经</div> <div>每元净资产营</div> </div>								
		流通市值	净资产	净资产增长率	净资产收益率	管现金流	业收入	净利润增长率	股息率	资产负债率
C0	均值	0.16	0.06	0.03	-0.01	-0.01	0.03	0.07	-0.01	-0.01
	绝对值均值	0.25	0.23	0.18	0.28	0.23	0.21	0.24	0.23	0.27
	平均持续时间	11.61	16.33	15.21	24.5	12.78	15.47	17.29	15.47	18.77
C1	均值	0.07	-0.03	-0.02	0.02	-0.01	-0.1	0.04	-0.1	0.08
	绝对值均值	0.31	0.26	0.27	0.26	0.23	0.24	0.24	0.22	0.18
	平均持续时间	14.46	13.16	13.57	12.78	11.03	10.5	16.64	14	10.26
C3	均值	0.16	0.06	0.01	0.05	0.03	-0.09	0.18	-0.03	0.1
	绝对值均值	0.4	0.33	0.38	0.39	0.41	0.42	0.44	0.32	0.4
	平均持续时间	14	14.46	15.21	14.95	13.16	20.05	17.29	15.47	16.64
C4	均值	0.11	0.02	0.01	0.02	-0.02	0	0.05	0.02	0
	绝对值均值	0.17	0.16	0.19	0.2	0.16	0.13	0.14	0.16	0.12
	平均持续时间	15.47	19.17	16.64	19.6	16.64	14.95	11.76	16.04	19.6
C5	均值	0.06	0.01	0.11	0.04	-0.01	-0.03	-0.02	0.03	-0.04
	绝对值均值	0.17	0.24	0.24	0.26	0.17	0.19	0.23	0.21	0.19
	平均持续时间	12.42	14.46	12.97	15.75	11.61	13.78	14.23	16.96	12.97
C6	均值	0.09	0.1	0.04	0	-0.02	0.11	0.03	-0.02	0
	绝对值均值	0.25	0.29	0.19	0.26	0.15	0.26	0.22	0.13	0.14
	平均持续时间	17.64	25.94	16.04	18.77	16.04	24.5	25.94	17.64	15.47
C7	均值	0.06	0.07	0.09	0.03	-0.02	-0.01	0.01	-0.02	0.01
	绝对值均值	0.2	0.17	0.18	0.17	0.15	0.14	0.15	0.09	0.12
	平均持续时间	23.21	13.78	23.21	13.16	20.51	15.75	14.7	8.24	19.17
C8	均值	0.09	0.03	0.03	0.05	-0.03	-0.06	0.01	-0.05	-0.03
	绝对值均值	0.19	0.16	0.21	0.22	0.2	0.19	0.21	0.19	0.2
	平均持续时间	17.29	16.33	18.77	14.95	15.47	22.62	13.57	22.62	14.95

资料来源: tinysoft, 国信证券经济研究所

表 9：证监会制造业子行业估值因子集与市场因子集分析
估值因子集
市场因子集

证监会行业	指标	<div> <div>市盈率</div> <div>市净率</div> <div>市销率</div> <div>市现率</div> </div>				<div> <div>换手率</div> <div>价格动量</div> <div>波动率</div> <div>股价</div> <div>机构持股</div> <div>总市值</div> </div>					
		市盈率	市净率	市销率	市现率	换手率	价格动量	波动率	股价	机构持股	总市值
C0	均值	0.1	0.08	0.02	0.01	0.07	0	0.09	0.08	-0.04	0.11
	绝对值均值	0.2	0.25	0.2	0.2	0.27	0.16	0.26	0.31	0.27	0.23
	平均持续时间	13.16	12.78	13.16	12.08	10.76	14.46	14.95	13.16	16.04	10.26
C1	均值	0.02	0.13	0.15	-0.04	0.13	-0.01	0.03	0.14	-0.01	0.06
	绝对值均值	0.26	0.37	0.34	0.19	0.31	0.26	0.3	0.31	0.23	0.35

	平均持续时间	11.03	22.62	20.05	8.17	11.45	14.7	10.63	13.16	14.23	18.77
C3	均值	-0.01	0.07	0.13	0.01	0.05	-0.13	0.08	0.13	0.2	0.08
	绝对值均值	0.38	0.36	0.38	0.32	0.31	0.38	0.29	0.37	0.37	0.37
	平均持续时间	13.16	11.31	11.76	12.08	8.65	15.21	6.35	10.76	13.78	13.36
C4	均值	0	0.12	0.07	0.05	0.1	0.03	0.03	0.2	0.07	0.12
	绝对值均值	0.17	0.19	0.17	0.11	0.18	0.1	0.14	0.27	0.23	0.22
	平均持续时间	22.05	13.57	14.46	9.69	16.96	12.6	9.59	23.21	17.64	14.46
C5	均值	0.02	0.14	0.11	0.05	0.08	0.02	0.09	0.2	0.07	0.15
	绝对值均值	0.22	0.24	0.2	0.21	0.24	0.22	0.27	0.35	0.27	0.19
	平均持续时间	11.45	14.7	10.38	11.31	10.38	20.51	11.92	18.77	15.75	10.5
C6	均值	0	0.08	0.01	0.07	0.02	0.04	0.02	0.14	0.13	0.14
	绝对值均值	0.22	0.29	0.23	0.14	0.17	0.09	0.25	0.29	0.21	0.28
	平均持续时间	15.75	22.05	22.05	11.61	11.92	9.59	10.38	21	14.46	16.33
C7	均值	0.05	0.12	0.08	0	0.08	0	0.1	0.17	0.07	0.1
	绝对值均值	0.18	0.21	0.19	0.13	0.17	0.12	0.18	0.22	0.21	0.24
	平均持续时间	22.62	19.6	28.45	16.04	12.78	19.17	14	14.95	15.47	30.41
C8	均值	0.07	0.09	0.15	0.09	0.06	-0.01	0.01	0.12	-0.02	0.04
	绝对值均值	0.19	0.26	0.21	0.2	0.26	0.16	0.25	0.29	0.28	0.23
	平均持续时间	12.08	12.6	14.7	13.78	21.51	20.51	13.36	24.5	19.6	16.64

资料来源: tinysoft, 国信证券经济研究所

上面列出了 20 个行业因子贡献度的三个衡量特征。为什么看因子贡献度而不是看因子本身呢? 有些量化选股策略通过统计规律分市场状况选择相应的因子指标, 但这种选股策略往往包含了两个隐含错误:

第一, 确实在特定的市场状态下, 可以通过因子的力度和持续程度总结出超额收益股票的一些指标特征。但问题在于, 这种市场状况的归类, 本身就是不可测的。站在任何一个时点上, 我们都可以知道过去是什么市场状况; 有经验的投资者, 也能大致判断出当前所处的市场阶段; 但是谁能预知到未来几个月, 或者说未来你的量化选股组合持有期, 是在牛市、熊市还是震荡市?

第二, 即使对市场状况的分类预测对, 这种统计规律也只是必要非充分条件, 不能确保在未来能发生相同的故事。其本质在于这些因子特征的来源, 究竟是市场状况导致的, 还是其他因素导致的, 比如宏观政策、结构变化等等。有可能市场状况和统计规律都只是现象, 两者在过去是由同一种内因驱动, 但未来即使市场状况相同, 因子的表现也未必相同。

我们反复强调, 因子贡献度也是客观历史的真实反映, 但它不涉及统计规律。而选股起作用的是区分度动量, 它不依赖于历史的规律。选择的因子, 如果未来不表现, 则在加权时将自动赋予低权重, 不需要主观判断。

那么什么样的行业适合利用这一类多因子量化选股呢?

(1) 首先从数量上来看, 由于量化策略是建立在大概率上的超额收益, 因此行业内的股票数量不应过少, 这样才能消除个别股票的异常表现对整个组合冲击。

(2) 从股票池而言，池中的股票应该具有差异性。在之前的报告中已经提过，区分度动量选股在沪深 300 股票池中的效果就远不如中证 800，其原因就在于沪深 300 股票由于是大多数机构的标配，走势同质性强。同样的，适合用多因子量化选股的行业也应该即具有差异性，又具有可比性。

(3) 在实证中我们发现因子贡献度的绝对值大小同股票池的数量有一定关系，当股票池数量较小时，贡献度水平会高于正常水平。造成这种结果的原因可能是由于数量少导致个股在贡献度计算时的影响程度被放大，比如极端情形一个行业只有 2 只股票，那么因子贡献度肯定是 1 或-1。所以显著性更适用于行业选择因子，而不适合作为判断行业是否有效的统计量

(4) 持续性与贡献度的绝对值水平无关，因此它不仅可以用在行业内选择因子，也可以用于行业间的比较。

(5) 偏离性是分析因子贡献度时对显著性和持续性的补充。直观来看，偏离性好的因子应该持续性也较好，但反过来不成立。

综合上述 5 点的分析，可以得出结论，行业内股票数目多且因子集整体持续性好，上市公司兼具差异性和可比性的行业适合采用多因子量化策略选股。其中前两点可以量化，而差异性和可比性则依赖于主观判断。

表 10：证监会行业股票数目与因子集持续性对比

代码	A	B	D	E	F	G	H	I	J	K	L	M
对应行业	农林 牧渔	采掘	水电 煤气	建筑	交运 仓储	信息 技术	批发 零售	金融、 保险	房地 产业	社会 服务	传播 文化	综合 类
因子持续性（平均）	15.32	16.00	14.02	16.43	16.01	12.74	18.70	12.69	15.88	14.77	14.08	12.33
股票数目（平均）	12	29	43	17	44	38	69	19	55	16	9	15
代码	C0	C1	C3	C4	C5	C6	C7	C8				
对应行业	食品 饮料	纺织 服装	造纸 印刷	石化 塑胶	电子	金属 非金	机械 设备	医药 生物				
因子持续性（平均）	14.64	13.75	13.56	16.01	13.73	17.53	18.17	17.15				
股票数目（平均）	32	20	11	85	32	82	112	45				

资料来源：wind 资讯，国信证券经济研究所

从上表可以看到，在不考虑差异性和对比性以及是否朝阳行业的前提下，**交运仓储、批发零售、房地产、石化、金属非金属、机械设备和医药生物**可能适用于区分度动量选股。

行业内区分度动量选股示例

在这一节，我们将以石油化工塑胶塑料业 C4 为示范，进行行业内区分度动量选股。那么首先需要解决的，是该行业应该用哪些因子。

前面定义了因子贡献度的持续性、显著性和偏离性，但对于区分度动量策略而言，这三个方面的重要程度并不一致，持续性>显著性>偏离性。在选因此

择因子时，按照以下步骤进行

(1) 将因子集的三个衡量指标分别进行排名，对于每个因子来说，都得到一个三维数组。

(2) 将持续性、显著性和偏离性排名分别以 3: 2: 1 的权重进行加权，得到每个因子的加权排名

(3) 选取 19 个因子中加权排名考前的 8 个因子作为区分度动量策略的因子集。

根据上述三个步骤，最终选出适用石化行业的 8 个因子分别为股价、机构持股、净资产收益率、换手率、市盈率、总市值、净资产和净资产增长率。以这 8 个因子作为因子集，用区分度动量和半衰期在行业内选股，实证的结果如下：

表 11: 石化行业内区分度动量选股效果

	行业收益	T=20	T=30	T=40	T=50	T=60
多头组合	106.26%	168.73%	118.47%	144.77%	81.35%	98.48%
空头组合	106.26%	19.32%	59.27%	70.27%	131.95%	115.47%

资料来源：国信证券经济研究所

可以看出当观察期为 20-40 天时，多头组合有明显的超额收益。而当观察期为 60 时，同中证 800 作为股票池不同，此时空头组合反而占优。在我们的测试中，不仅仅是石化行业出现这个现象，很多其他的行业也有类似的结果。

这并不是说行业内选股失败了。恰恰相反，这证明了行业里因子轮动与全市场的轮动特性有很大不同。长观察期空头组合取得了超额收益，说明在行业内，因子区分度首先表现出的是动量效应，随后逐步转为反转效应。下表是机械设备行业的实证结果。可以看到多头组合虽然全部跑赢了行业基准，但也随着观察期变长效果逐渐减弱。而空头组合在 50-60 同样表现出反转效应。

表 12: 机械设备行业内区分度动量选股效果

	行业收益	T=20	T=30	T=40	T=50	T=60
多头组合	133.91%	161.17%	178.77%	176.88%	154.36%	150.86%
空头组合	133.91%	91.80%	83.87%	106.11%	159.12%	157.18%

资料来源：国信证券经济研究所

思考及后续研究

对行业内因子分析的思考

纵观我们这个方向的 4 篇系列报告，其核心都是因子贡献度和区分度动量。虽然经常把两者放在一起，但它们实质上是两个独立的框架。

因子贡献度本身是一个纯粹的度量工具，它是对市场已经发生的过程的客观反应，既不存在规律的统计，也不涉及对未来的预测。区分度动量模型则不同，它的假设前提是动量存在，核心在于怎么跟踪动量和观察动量衰竭，而因子贡献度是它跟踪的通道。

回顾区分度动量策略的逻辑，它本质还是一个线性模型：首先，它需要假定因子的影响具有线性性或者说单调性；其次，在打分的时候利用贡献度加权求和，实际是因子排名的线性变换。在贡献度的基础上，如果度量了因子，选出了因子集，是不是一定要用线性模型来选股？我们觉得答案是否定的，因子贡献度就像砖头，是建宫殿还是盖园林，完全是每个投资者自己的选择。

回到行业内选股，我们发现同样的因子集，在全市场和样业内表现出不同的特征。全市场中还能体现动量效应的观察期，在行业内表现出的却是反转。事实上，在行业内股票的信息反应较全市场更有迅速。目前来看尽管用因子贡献度选出了行业内的因子集，但如何选股，方法上还有很大优化的空间，这也是我们后续研究方向的重点。

一个有趣的现象

在观察因子贡献度的偏离性时，我们发现一个有趣的现象，就是所有的行业内，股价因子都是偏向同一侧的。换言之，在历史上所有行业内低价股存在超额收益。

为什么会出现这种有趣的现象呢？我们认为这和市场参与者的心理有关，如果用金融行为学便可以得到解释。首先，单价低会让人有便宜的错觉，比如1元的股票让人感觉还有很大上涨空间，而100元的股票涨到200元则从感觉上难度较大；其次，价格较低的股票能有较好的流动性，尤其是针对中小投资者，100元的股票，成交最少单位也需要1万元的资金，这便无形中产生了投资门槛。

这就容易解释事件套利的研究中，为什么统计结果表明在中国股市上市公司拆股前后股票往往会有超额收益产生。

国信证券投资评级

类别	级别	定义
股票 投资评级	推荐	预计 6 个月内，股价表现优于市场指数 20%以上
	谨慎推荐	预计 6 个月内，股价表现优于市场指数 10%-20%之间
	中性	预计 6 个月内，股价表现介于市场指数 $\pm 10\%$ 之间
	回避	预计 6 个月内，股价表现弱于市场指数 10%以上
行业 投资评级	推荐	预计 6 个月内，行业指数表现优于市场指数 10%以上
	谨慎推荐	预计 6 个月内，行业指数表现优于市场指数 5%-10% 之间
	中性	预计 6 个月内，行业指数表现介于市场指数 $\pm 5\%$ 之间
	回避	预计 6 个月内，行业指数表现弱于市场指数 5%以上

免责条款

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员对使用本报告及其内容所引发的任何直接或间接损失概不负责。我公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。本报告版权归国信证券所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登。

国信证券经济研究所团队成员

宏观		策略		交通运输	
周炳林	0755-82130638	黄学军	021-60933142	郑 武	0755- 82130422
林松立	010-66026312	崔 嵘	021-60933159	陈建生	0755- 82133766
		廖 喆	021-60933162	岳 鑫	0755- 82130432
				高 健	0755-82130678
银行		房地产		机械	
邱志承	021- 60875167	方 焱	0755-82130648	余爱斌	0755-82133400
黄 飙	0755-82133476	区瑞明	0755-82130678	黄海培	021-60933150
谈 煜	010- 66025229	黄道立	0755- 82133397	陈 玲	0755-82130646
				杨 森	0755-82133343
				李筱筠	010-66026326
汽车及零配件		钢铁		商业贸易	
李 君	021-60933156	郑 东	010- 66026308	孙菲菲	0755-82130722
左 涛	021-60933164	秦 波	010-66026317	吴美玉	010-66026319
				祝 彬	0755-82131528
基础化工		医药		石油与石化	
张栋梁	0755-82130532	贺平鸽	0755-82133396	李 晨	021-60875160
陈爱华	0755-82133397	丁 丹	0755- 82139908	严蓓娜	021-60933165
邱 斌	0755-82130532	陈 栋	021-60933147		
电力设备与新能源		传媒		有色金属	
皮家银	021-60933160	陈财茂	021-60933163	彭 波	0755-82133909
				谢鸿鹤	0755-82130646
电力与公用事业		非银行金融		通信	
徐颖真	021-60875162	邵子钦	0755- 82130468	严 平	021-60875165
谢达成	021-60933161	田 良	0755-82130513	程 峰	021-60933167
		童成敦	0755-82130513		
造纸		家电		计算机	
李世新	0755-82130565	王念春	0755-82130407	段迎晟	0755- 82130761
邵 达	0755-82130706				
电子元器件		纺织服装		农业	
段迎晟	0755- 82130761	方军平	021-60933158	张 如	021-60933151
高耀华	0755-82130771				
旅游		食品饮料		建材	
廖绪发	021-60875168	黄 茂	0755-82138922	杨 昕	021-60933168
刘智景	021-60933148				
煤炭		建筑		固定收益	
李 然	010-66026322	邱 波	0755-82133390	李怀定	021-60933152
陈 健	010-66215566	李遵庆	0755-82133055	高 宇	0755- 82133538
苏绍许	021-60933144			侯慧娣	021-60875161
				张 旭	010-66026340
				蔺晓熠	021-60933146
				刘子宁	021-60933145
指数与产品设计		投资基金		量化投资	
焦 健	0755-82133928	杨 涛	0755-82133339	葛新元	0755-82133332
王军清	0755-82133297	彭怡萍	0755-82133528	董艺婷	021-60933155
彭甘霖	0755-82133259	刘舒宇	0755-82133568	林晓明	0755-25472656
阳 瑾	0755-82133538	康 亢	010-66026337	赵斯尘	021-60875174
周 琦	0755-82133568	刘 洋		程景佳	021-60933166
赵学昂	0755-66025232			郑 云	021-60875163
				毛 甜	021-60933154
交易策略					
戴 军	0755-82133129				
秦国文	0755-82133528				
徐左乾	0755-82133090				
黄志文	0755-82133928				

国信证券机构销售团队

华北区（机构销售一部）			华东区（机构销售二部）			华南区（机构销售三部）		
王立法	010-66026352 13910524551 wanglf@guosen.com.cn		盛建平	021-60875169 15821778133 shengjp@guosen.com.cn		万成水	0755-82133147 13923406013 wancs@guosen.com.cn	
王晓建	010-66026342 13701099132 wangxj@guosen.com.cn		马小丹	021-60875172 13801832154 maxd@guosen.com.cn		魏 宁	0755-82133492 13823515980 weining@guosen.com.cn	
焦 骥	010-66026343 13601094018 jiaojian@guosen.com.cn		郑 毅	021-60875171 13795229060 zhengyi@guosen.com.cn		邵燕芳	0755-82133148 13480668226 shaoyf@guosen.com.cn	
李 锐	010-66025249 13691229417 lirui2@guosen.com.cn		黄胜蓝	021-60875166 13761873797 huangsl@guosen.com.cn		林 莉	0755-82133197 13824397011 linli2@guosen.com.cn	
徐文琪	010-66026341 13811271758 xuwq@guosen.com.cn		刘 塑	021-60875177 13817906789 liusu@guosen.com.cn		王昊文	0755-82130818 18925287888 wanghaow@guosen.com.cn	
			叶琳菲	021-60875178 13817758288 yelf@guosen.com.cn		甘 墨	0755-82133456 15013851021 ganmo@guosen.com.cn	
			孔华强	021-60875170 13681669123 konghq@guosen.com.cn		段莉娟	0755-82130509 18675575010 duanlj@guosen.com.cn	
						黎 敏	0755-82130681 13902482885 limin1@guosen.com.cn	
						徐 冉	13632580795 xuran1@guosen.com.cn	
						颜小燕	13590436977 yanxy@guosen.com.cn	