

证券研究报告—深度报告

金融工程

数量化投资

数量化投资系列报告之三十六

2011年2月22日

本报告的独到之处:

- 提出基于基本面先行因子的行业收益预测
- 能够对各行业的未来排序进行预测

专题报告

A股市场行业基本面先行因子及其对行业收益的预测作用

本文主要从行业的基本面角度分析先行因子存在的可能性,通过海量的基本面因子筛选,寻找可行的行业先行因子,并通过构建基于先行因子的回归模型来对行业收益进行预测。

全文主要从以下几个方面进行分析:

证券分析师:董艺婷

电话: 021-60933155

E-mail: dongyt@guosen.com.cn证券投资咨询执业资格证书编码:
S0980510120055

联系人: 郑云

电话: 021-60875163

E-mail: zhengyun@guosen.com.cn

联系人: 毛甜

电话: 021-60933154

Email: maotian@guosen.com.cn

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。

1.我们提出了基于一级行业和二级行业的全新的行业分类标准,通过不同级别行业之间的相关性对比,来对原行业进行了重新筛选和分类,为我们后续的研究提供更恰当的行业分类支持。

2.我们运用 Granger 因果检验对行业基本面因子进行筛选,剔除掉那些一方面与行业收益率没有因果关系,或者属于滞后变量的因子,最终筛选出各个行业的先行基本因子。

3.在完成先行因子的筛选后,我们根据这些因子建立单行业的预期收益模型,进行行业收益率的预测,同时为了解决先行因子间存在的多重共线性,我们采用岭回归来建立具体的收益预测模型。

4.模型结果显示,2010年后的样本外区间,各行业预测收益率与真实收益率的相关性非常强,只有房地产行业和食品饮料行业的预测效果不尽如人意。通过对15个行业的预测情况进行分析,我们发现,模型的预测命中率是比较不错的,尤其是对于多头的判断,非常准确。我们预测出的涨跌幅前三的行业命中率曾一度高达66.67%,效果令人满意。且多头组合的实际排名一直处于高位,多头战胜空头的排名命中率赫然是100%!

5.目前阶段,我们对于先行因子的使用及预测模型的运用,主要集中于行业收益率的预测,并初步根据预测收益率对各行业进行排序,以预测热点行业及危险行业。未来,我们仍会在原始基本面因子集的系统化整理和搜集、先行因子的稳定性检验和样本外检验的持续性方面作出改善。

内容目录

行业的分类方法	4
行业分类逻辑	4
行业分类步骤	4
行业分类结果	4
先行因子的筛选	7
原始数据搜集	7
数据的预处理方法—补齐数据	7
先行因子的筛选方法	8
Granger 因果检验原理	9
Granger 因果检验的步骤	9
先行因子的筛选步骤	10
行业先行基本面因子集	11
单个行业预期收益模型	26
多重共线性的存在	26
多重共线性的处理	26
数据及回归方法说明	28
预测收益模型建立	28
单行业到行业配置的可行性分析	33
模型应用及未来优化	35
模型解释	35
模型应用及改进	36

图表目录

图 1: 缺失数据的处理步骤	7
图 2: 缺失数据的处理步骤	8
图 3: 先行因子筛选步骤	10
图 4: 多重共线性简介	27
图 5: 煤炭行业预测收益率对比	29
图 6: 有色金属行业预测收益率对比	29
图 7: 商业零售行业预测收益率对比	29
图 8: 电子元器件行业预测收益率对比	30
图 9: 信息设备行业预测收益率对比	30
图 10: 房地产行业预测收益率对比	30
图 11: 建筑建材行业预测收益率对比	31
图 12: 钢铁行业预测收益率对比	31
图 13: 医药行业预测收益率对比	31
图 14: 食品饮料行业预测收益率对比	32
图 15: 证券行业预测收益率对比	32

图 16: 农林牧渔行业预测收益率对比	32
图 17: 行业配置多空实际排名统计	34
图 18: 预测收益率行业配置收益情况对比	35
表 1: 重新筛选后的行业分类结果	5
表 2: 石油化工行业先行因子筛选结果及先行阶数	11
表 3: 房地产行业先行因子筛选结果及先行阶数	11
表 4: 钢铁行业先行因子筛选结果及先行阶数	12
表 5: 电子元器件行业先行因子筛选结果及先行阶数	13
表 6: 煤炭开采行业先行因子筛选结果及先行阶数	13
表 7: 零售行业先行因子筛选结果及先行阶数	14
表 8: 信息设备先行因子筛选结果及先行阶数	15
表 9: 有色金属先行因子筛选结果及先行阶数	15
表 10: 证券行业先行因子筛选结果及先行阶数	16
表 11: 医药生物行业（剔除医疗服务）先行因子筛选结果及先行阶数	16
表 12: 食品饮料行业先行因子筛选结果及先行阶数	17
表 13: 建筑建材行业先行因子筛选结果及先行阶数	18
表 14: 银行业先行因子筛选结果及先行阶数	19
表 15: 纺织行业先行因子筛选结果及先行阶数	20
表 16: 化工行业先行因子筛选结果及先行阶数	21
表 17: 保险行业先行因子筛选结果及先行阶数	22
表 18: 农林牧渔（剔除林业）行业先行因子筛选结果及先行阶数	23
表 19: 家用电器行业先行因子筛选结果及先行阶数	25
表 20: 多重共线性诊断（部分因子结果）	26
表 21: 煤炭行业预测收益率准确率	29
表 22: 有色金属行业预测收益率准确率	29
表 23: 商业零售行业预测收益率准确率	29
表 24: 电子元器件行业预测收益率准确率	30
表 25: 信息设备行业预测收益率准确率	30
表 26: 房地产行业预测收益率准确率	30
表 27: 建筑建材行业预测收益率准确率	31
表 28: 钢铁行业预测收益率准确率	31
表 29: 医药行业预测收益率准确率	31
表 30: 食品饮料行业预测收益率准确率	32
表 31: 证券行业预测收益率准确率	32
表 32: 农林牧渔行业预测收益率准确率	32
表 33: 行业配置涨跌幅预测命中率统计	33
表 34: 行业配置涨跌速预测命中率统计	34
表 35: 预测收益率行业配置收益情况统计	35

行业的分类方法

如我们所知，市场上现在存在很多行业分类方法，有的分类大而泛，几个行业就代表了市场的全部，有的分类广而细，些许差别就要另称一个行业。虽然这些分类各有特点，有着不同的受众群体，侧重的行业区分角度存在差异，细分程度有所不同，但不可否认，都具有一定的现实性和科学性，能够反映一定需求。

诸如申万一级行业分类和申万二级行业分类，前者包括 23 个子行业，后者包括 81 个子行业，虽然细分程度不同，但均考虑到中国经济发展和证券市场的现实性和发展特点，在行业的财务指标分析、行业市场表现方面都能够体现出较好的业内一致性和业外的差异性。然而，从我们的研究角度出发，这仍然无法满足我们的分析需求，我们还需要进行进一步的划分以实现相应的研究目的。

行业分类逻辑

现有的一级行业分类标准和二级行业分类标准虽然已经非常科学，但实际运用中，前者过简，行业内存在一定的差异性，后者过繁，行业间又出现了一定的相关性，对于我们的研究均不太有利，那么是否存在一种划分方式能够既继承他们的划分优势信息，又能更合理地在差异性和相关性上找到平衡。

对此，我们有以下的分析逻辑：对于一个二级行业，它与自己所属的一级行业一定存在着较高的收益相关性，有着相近的市场表现及运行轨迹，自然没有必要单独进行分析，而与此同时，该一级行业又与其他一级行业存在着一定的相关性，毕竟市场上可能存在着一些因素同时作用于各行业，这就使得一级行业间的相关性也不容忽视；如果该二级行业与所属一级行业之间的相关性不够高，甚至要低于该一级行业与其它行业间的相关性，那么这一二级行业就完全可以脱离于所属的一级行业，不应再当作同一个行业来看待。

这样的话，我们就可以从一级行业中分离出一批二级行业而独立存在，使得新构造的行业分类既保存了行业内的一致性，又兼顾了行业间的差异性。

行业分类步骤

下面是我们根据上述逻辑进行行业划分的简单步骤；

- 1、计算各个二级行业 i 与所属一级行业 k 的相关性 C_{ik} ；
- 2、计算所属一级行业与其他一级行业相关系数的平均值 V_k

$$V_k = \frac{\sum_{j \neq k}^r v_{kj}}{r-1}$$

V_{kj} 为一级行业 k 与一级行业 j 的相关系数， r 为一级行业的总数；

- 3、比较 C_{ik} 与所属一级行业 k 的 V_k 的大小，如果 $V_k < C_{ik}$ 则二级行业 i 与所属一级行业 k 存在高度的相关性，不须单独成为一个行业，而如果 $V_k \geq C_{ik}$ 则二级行业 i 与所属一级行业 k 存在较低的相关性，须要单独成为一个行业；
- 4、最后将原有一级行业及单独成立的二级行业合并，形成新的行业分类。

行业分类结果

如图 1 所示，我们将申万一级行业及二级行业重新进行分类后，得到全新的行业分类。

表 1: 重新筛选后的行业分类结果

一级行业	二级行业	新分类
农林牧渔	林业	林业
	农产品加工	农林牧渔（剔除林业）
	农业综合	
	饲料	
	渔业	
	种植业	
采掘	煤炭开采	煤炭开采
	其他采掘	其他采掘
	石油开采	石油开采
化工	化工新材料	化工新材料
	化学纤维	石油化工
	化学原料	塑料
	化学制品	橡胶
		化工(剔除化工新材料、石油化工、塑料、橡胶)
	石油化工	
	塑料	
黑色金属	钢铁	黑色金属
	有色金属	有色金属
	金属新材料	
	有色金属冶炼	
建筑建材	建筑材料	建筑建材
	建筑装饰	
机械设备	电气设备	仪器仪表
	普通机械	金属制品
	仪器仪表	机械设备(剔除仪表仪器、金属制品)
	专用设备	
	金属制品	
电子元器件	半导体	电子元器件
	其他电子器件	
	元件	
	显示器件	
交运设备	非汽车交运设备	汽车服务
	汽车服务	交运设备(剔除汽车服务)
	汽车零部件	
	汽车整车	
信息设备	计算机设备	信息设备
	通信设备	
家用电器	白色家电	家用电器
	视听器材	
食品饮料	食品加工	食品加工

	食品制造	食品饮料(剔除食品加工)
	饮料制造	
纺织服装	纺织	纺织服装
	服装	
轻工制造	包装印刷	轻工制造
	其他轻工制造	
	造纸	
医药生物	化学制药	医疗服务
	生物制品	医药生物(剔除医疗服务)
	医疗器械	
	医药商业	
	中药	
	医疗服务	
公用事业	电力	环保
	环保	燃气
	燃气	公用事业(剔除环保、燃气)
	水务	
交通运输	港口	交通运输
	公交	
	航空运输	
	机场	
	高速公路	
	航运	
	铁路运输	
	物流	
房地产	房地产开发	房地产
	园区开发	
金融服务	信托	信托
	银行	银行
	证券	证券
	保险	保险
商业贸易	零售	商业贸易
	贸易	
餐饮旅游	餐饮	餐饮
	景点	餐饮旅游(剔除餐饮)
	酒店	
	旅游综合	
信息服务	传媒	通信运营
	计算机应用	信息服务(剔除通讯运营)
	通信运营	
	网络服务	
综合	综合	综合

资料来源：国信证券经济研究所

经过重新筛选和划分，我们将 81 个二级行业缩小为 43 个新行业，一级行业中，一些相关性较高的二级行业得以合并，而相关性较低的二级行业也能够独立出来，单独进行分析。此外，我们还依据申万二级行业的成分股进出记录，重新组成了各个新行业的成分股进出记录，根据已有的成分股历史信息，回溯出各个新行业的行业收益记录，为后续的研究提供支持。

先行因子的筛选

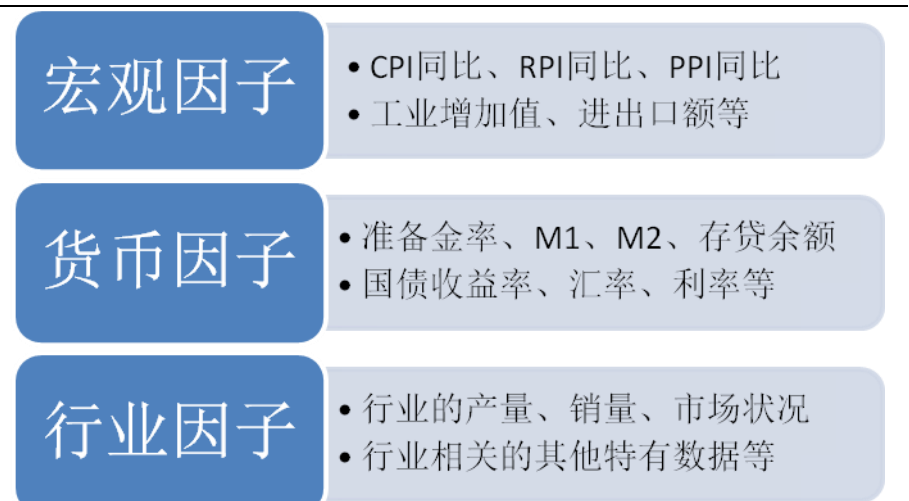
原始数据搜集

为了寻找到行业运行的基本面先行因子，我们从未知的因子世界里搜罗了一系列备选因子以供分析。

因子数量虽然繁多，但归结起来，大致可以分为以下三类：宏观因子、货币因子、行业因子。

其中宏观因子与货币因子为共有因子，行业因子为特有因子。

图 1：缺失数据的处理步骤



资料来源：国信证券经济研究所

数据的预处理方法—补齐数据

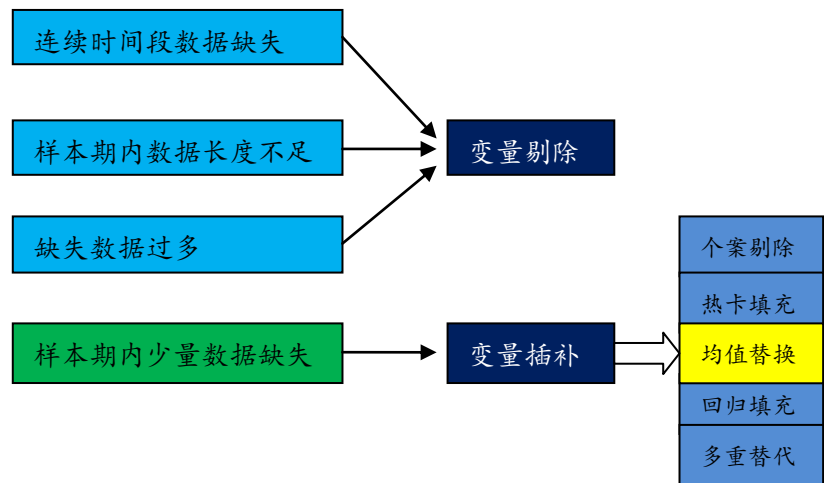
出于各种原因，如数据源出错、数据发布存在一定的缺失规律、人为录入出现纰漏等，数据的不同程度缺失总是不可避免的。

然而，缺失的数据势必会对我们的进一步分析造成相当大的障碍，如进行多元回归时，数据中如果存在缺失项，而不加任何处理就直接进入回归，那么其可信度势必令人怀疑。

一般来说，数据缺失量如果比较大的话，应该毫不犹豫将该变量剔除，进行插补已经没有意义了，而数据缺失只是发生在极个别时间点的话，我们还是可以进行插值填补的。

显然，对于数据质量过差的数据，进行任何处理都只会影响我们的分析结果，故而应直接剔除掉。值得进行插补的只有那些缺失数据少，样本期内信息完整的数据变量。

图 2：缺失数据的处理步骤



资料来源：国信证券经济研究所

下面我们简单对各种填补方式进行一下选择：

1. **个案剔除**：如果任何一个变量含有缺失数据的话，就把相对应的缺失个案从分析中剔除。如果缺失值所占比例比较小的话，这一方法十分有效。但这里，我们的分析涉及到滞后变量，剔除个案，会使得滞后变量无法对应，故不采用；
2. **热卡填充**：对于一个包含缺失值的变量，在数据库中找到一个与它最相似的对象，然后用这个相似对象的值来进行填充。这显然违背了我们进行因子分析的初衷，变量与其他变量相似，使得该因子的分析变得没有必要；
3. **均值替换**：这里的均值替换，并不是指用其他所有对象的取值的平均值来填充该缺失的变量值，而是指用缺省值附近的相邻数据均值来进行替换。这样的话，替换值就能尽可能地接近原有值，减少误差，且使用起来也极为方便，我们后来就是采用了这一方法。
4. **回归填充**：需要选择若干个预测缺失值的自变量，然后建立回归方程估计缺失值，即用缺失数据的条件期望值对缺失值进行替换。仍旧是利用其他变量来估计进行填充，可行性不高。
5. **多重替代**：用一系列可能的值来替换每一个缺失值，以反映被替换的缺失数据的不确定性。然后，用标准的统计分析过程对多次替换后产生的若干个数据集进行分析。最后，把来自于各个数据集的统计结果进行综合，得到总体参数的估计值。虽然这一方法反映出了由于数据缺失而导致的不确定性，但替代过程显然不便于进行批量的变量填补，也不应采用。

最后，我们采用的是邻近值均值插补来进行缺失数据的补齐处理。

除了缺失值的处理之外，还有一系列问题有待我们解决。

行业的基本面指标种类繁多，动辄上百，如此大量的数据，若全部都纳入到回归模型中，显然是不合理的，因为在回归中常常会出现没有因果关系的变量之间常常有很好的回归拟合，变量之间存在多重共线性等等，从经济意义上和统计角度上削弱模型的解释力度和稳定性的问题。我们需要一一解决这些问题。

先行因子的筛选方法

我们这里运用 Granger 因果检验对行业基本面因子进行筛选，筛选的目的是为了

剔除掉那些一方面与行业收益率没有因果关系，另一方面即使有因果关系，但属于滞后变量，无法在实际中应用到行业收益率的预测中的因子，换句话说，是为了筛选出行业的先行基本面因子。

Granger 因果检验原理

Granger 因果检验由 2003 年诺贝尔经济学奖得主克莱夫·格兰杰（Clive W. J. Granger）所开创，用于分析经济变量之间的因果关系。

在时间序列情形下，两个经济变量 X、Y 之间的 Granger 因果关系定义为：若在包含了变量 X、Y 的过去信息的条件下，对变量 Y 的预测效果要优于只单独由 Y 的过去信息对 Y 进行的预测效果，即变量 X 有助于解释变量 Y 的将来变化，则认为变量 X 是引致变量 Y 的 Granger 原因。

Granger 因果检验假设了有关 y 和 x 每一变量的预测的信息全部包含在这些变量的时间序列之中。检验要求估计以下的回归：

$$y_t = \sum_{i=1}^q \alpha_i x_{t-i} + \sum_{j=1}^q \beta_j y_{t-j} + u_{1t}$$

$$x_t = \sum_{i=1}^s \lambda_i x_{t-i} + \sum_{j=1}^s \delta_j y_{t-j} + u_{2t}$$

其中白噪音 u_{1t} 和 u_{2t} 假定为不相关的。

式（1）假定当前 y 与 y 自身以及 x 的过去值有关，而式（2）对 x 也假定了类似的行为。

对式（1）而言，其零假设 $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$ 。

对式（2）而言，其零假设 $H_0 : \delta_1 = \delta_2 = \dots = \delta_s = 0$ 。

因果关系研究的有趣例子是回答“先有鸡还是先有蛋”的问题。1988 年有两位学者 Walter N. Thurman 和 Mark E. Fisher 用美国 1930—1983 年鸡蛋产量 (EGGS) 和鸡的产量 (CHICKENS) 的年度数据，对此问题进行了统计研究。他们运用 Granger 的方法检验鸡和蛋之间的因果关系，结果发现，鸡生蛋的假设被拒绝，而蛋生鸡的假设成立，因此，蛋为因，鸡为果，也就是先有蛋。他们并建议作其他诸如“谁笑在最后谁笑得最好”、“骄傲是失败之母”之类的 Granger 因果检验。

Granger 因果检验的步骤

1. y 对 y 的滞后项 $y_{t-1}, y_{t-2}, \dots, y_{t-q}$ 及其他变量进行回归，这一回归中没有把滞后项 x 包括进来，是一个受约束的回归。回归得到受约束的残差平方和 RSS_R 。

2. 在 1 中的回归式中加进滞后项 x，如下

$$y_t = \sum_{i=1}^q \alpha_i x_{t-i} + \sum_{j=1}^q \beta_j y_{t-j} + u_{1t}$$

这是一个无约束的回归，由此回归得到无约束的残差平方和 RSS_{UR} 。

3. 检验的零假设为：

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$$

为了检验此假设，用 F 检验，即：

$$F = \frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(n-k)}$$

其中，n 是样本容量，q 为滞后项 x 的个数，即有约束回归方程中待估参数的个数，k 是无约束回归中待估参数的个数。

4. 如果在选定的显著性水平 α 上计算的 F 值超过临界值（或 P 值小于 0.05），则拒绝零假设，表明 x 是 y 的 Granger 原因。
5. 同样，为了检验 y 是否是 x 的原因，可将变量 y 与 x 相互替换，重复步骤（1）~（4）。

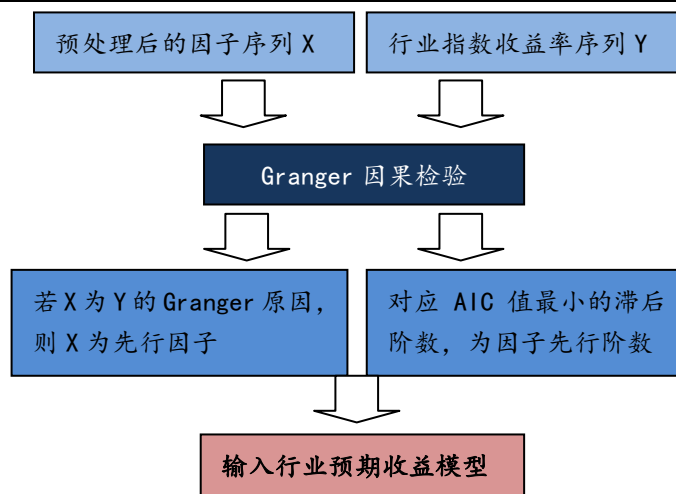
Granger 因果检验对于滞后期长度的选择有时很敏感。不同的滞后期可能会得到完全不同的检验结果。因此，一般而言，常进行不同滞后期长度的检验，根据赤池信息准则 AIC 来选取滞后期。

Granger 因果检验的结论只是统计意义上的因果性，而不一定是真正的因果关系。虽然可以作为真正的因果关系的一种支持，但不能作为肯定或否定因果关系的最终根据。当然，即使格兰杰因果关系不等于实际因果关系，也并不妨碍其参考价值。因为统计意义上的因果关系也是有意义的，对于经济预测等仍然能起很大的作用。

先行因子的筛选步骤

对每个行业，我们将进行过预处理的行业基本因子时间序列逐个与该行业指数的月收益率序列进行 Granger 因果检验，根据 P 值大小确定显著性因子，根据 AIC 准则确定最优滞后阶数，保留为指数收益率 Granger 原因的因子，得到该行业的先行因子集及因子的滞后阶数，用于下一步行业预期收益模型的输入变量。

图 3：先行因子筛选步骤



资料来源：国信证券经济研究所

行业先行基本面因子集

根据前文中的因子筛选过程，我们得到市场占比大于 1% 的诸多行业的先行基本面因子。先行因子均采用月度数据，阶数以月为单位。

表 2：石油化工行业先行因子筛选结果及先行阶数

宏观指标		行业指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
宏观经济景气指数:先行指数	1	国际原油价格(美元/桶)-Brent	3	加权平均价格-石脑油	2
CPI:当月同比	1	新加坡成品油价格(美元/每桶)-Jet Kerosene	3	成品油实现的原油价格(美元/桶)-加权平均价格	2
CPI:累计同比	1	90#汽油-含税月均价	2	上月迪拜原油价格-成品油实现的原油价格(美元/桶)	6
PPI:全部工业品:当月同比	2	0#柴油-90#汽油	2	国际原油价格(美元/桶)-Dubai	3
PPI:全部工业品:累计同比	1	航煤-0#柴油	6	固定资产投资完成额:石油加工、炼焦及核燃料加工业:累计同比	1
RPI:当月同比	1	加权平均价格-石脑油	4	美国成品油价格(美分/加仑)-柴油	5
进口金额:当月同比	1	90#汽油-不含税月均价	2	固定资产投资完成额:橡胶制品业:累计同比	1
社会消费品零售总额:当月同比	1	航煤-0#柴油	6	美国成品油价格(美分/加仑)-航空煤油	3

资料来源：国信证券经济研究所

表 3：房地产行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	1	准货币 - 定期存款	4	房地产开发资金来源累计合计累计同比(%)	2
RPI:当月同比	1	CHIBOR 加权平均	2	房地产开发资金来源:利用外资累计同比(%)	1
PPI:全部工业品:当月同比	2	CHIBOR 隔夜	2	房地产开发资金来源:自筹资金累计同比(%)	2
PPI:全部工业品:环比	3	CHIBOR7 天	6	全国土地购置面积同比	1
PPI:全部工业品:累计同比	1	法定准备金利率	4	实际完成投资额同比增长	4
进出口金额:当月同比	1	银行间回购利率加权平均	6	美国住宅开工数同比	4
进口金额:当月同比	1	银行间回购利率 7 天	6	美国住宅获建筑许可数同比	1
社会消费品零售总额:累计同比	6	人民币贷款基准利率 6 个月	4	美国单户住宅销售套数同比	3
产销率:累计值	1	人民币存款基准利率活期	4		
CPI:当月同比	4	人民币存款基准利率 3 个月	4		
		人民币存款基准利率 6 个月	4		
		汇率欧元	5		
		M1	6		
		活期存款	1		

资料来源：国信证券经济研究所

表 4: 钢铁行业先行因子筛选结果及先行阶数

宏观及货币指标		行业指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
宏观经济景气指数:先行指数	1	线材:6.5 高线:北京	5	产量:生铁:当月值	4
CPI:当月同比	3	IISI Crude Steel Production Data/Libya	2	线材:6.5 高线:上海	5
CPI:累计同比	1	IISI Crude Steel Production Data/South Korea	5	产量:铁矿石原矿量:当月同比	3
PPI:全部工业品:当月同比	2	IISI Crude Steel Production Data/Taiwan	6	热轧:3.0 热轧板卷:北京	4
PPI:全部工业品:累计同比	1	进口数量:铁道用材:当月值	1	铁道用钢材:产量:当月同比	4
RPI:当月同比	1	房屋施工面积:累计同比	1	全球粗钢总产量	4
银行间回购利率加权平均	6	商品房销售面积:累计同比	2	北美	4
		线材:独联体(FOB):最低价	5	南美	4
		线材:美国(CIF):最低价	5	热轧:3.0 热轧板卷:上海	5
		产量:工业锅炉:累计同比	1	亚洲	4
		产量:民用钢质船舶:累计同比	1	冷轧:1.0mm:北京	5
		产量:工业锅炉:累计同比	1	IISI Crude Steel Production Data/Slovakia	6
		螺纹钢:独联体(FOB):最低价	5	IISI Crude Steel Production Data/Slovenia	2
		产量:民用钢质船舶:累计同比	1	冷轧:1.0mm:广州	4
		螺纹钢:美国(CIF):最低价	1	IISI Crude Steel Production Data/Serbia and Montenegro	2
		型材(300 ~ 600):美国(CIF):最低价	1	IISI Crude Steel Production Data/Moldavia	6
		H 型钢(600mm 以上):美国(CIF):最低价	1	冷轧:1.0mm:上海	5
		社会消费品零售总额:当月同比	1	IISI Crude Steel Production Data/Saudi Arabia	4
		社会消费品零售总额:累计同比	1	IISI Crude Steel Production Data/El Salvador	2
		宏观经济景气指数:预警指数	1	型材:25#工字钢:上海	5
		线材:6.5 高线:广州	5	IISI Crude Steel Production Data/United States	4
		中板 3 ~ 10:美国(CIF):最低价	1	IISI Crude Steel Production Data/Argentina	5
		PPI:全部工业品:当月同比	5	IISI Crude Steel Production Data/Brazil	4
		厚板 10 以上:美国(CIF):最低价	1	热轧板卷:美国(CIF):最低价	1
		工业增加值:重工业:当月同比	5	冷轧板卷:美国(CIF):最低价	1

资料来源: 国信证券经济研究所

表 5: 电子元器件行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	1	M1、M2 增速差	1	出口数量:收音设备(包括收录音组合机及整套散件):当月同比	4
RPI:当月同比	1	CHIBOR 加权平均	5	出口数量:黑白电视机(包括整套散件):当月同比	3
PPI:全部工业品:当月同比	2	CHIBOR 隔夜	6	出口数量:印刷电路:当月同比	4
PPI:全部工业品:环比	1	CHIBOR7 天	6	出口数量:照相机:当月同比	5
PPI:全部工业品:累计同比	1	法定准备金利率	4	进口数量:自动数据处理设备的零件:当月同比	3
进出口金额:当月同比	1	银行间回购利率加权平均	5		
出口金额:当月同比	2	银行间回购利率隔夜	6		
进口金额:当月同比	1	银行间回购利率 7 天	5		
产销率:累计值	1	人民币贷款基准利率 6 个月	4		
CPI:当月同比	1	人民币贷款基准利率 1 年	4		
		人民币贷款基准利率 1-3 年	4		
		人民币贷款基准利率 3-5 年	4		
		人民币贷款基准利率 5 年以上	4		
		人民币存款基准利率活期	4		
		人民币存款基准利率 3 个月	4		
		人民币存款基准利率 6 个月	4		
		人民币存款基准利率 1 年	4		
		人民币存款基准利率 2 年	4		
		M1	1		
		活期存款	3		
		准货币 - 储蓄存款	3		

资料来源: 国信证券经济研究所

表 6: 煤炭开采行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	1	CHIBOR 加权平均	5	无烟煤进口量增长率	2
PPI:全部工业品:当月同比	2	CHIBOR 隔夜	6		
PPI:全部工业品:环比	3	CHIBOR7 天	6		
PPI:全部工业品:累计同比	1	法定准备金利率	4		
进口金额:当月同比	1	银行间回购利率加权平均	6		
CPI:当月同比	6	银行间回购利率隔夜	5		
		银行间回购利率 7 天	6		
		M1	6		
		活期存款	1		

资料来源: 国信证券经济研究所

表 7：零售行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	1	CHIBOR 加权平均	6	家具类零售总额增速	1
PPI:全部工业品:当月同比	2	CHIBOR 隔夜	6	建筑及装潢材料类零售总额增速	1
PPI:全部工业品:环比	1	CHIBOR7 天	6	名义消费品零售总额增速	1
PPI:全部工业品:累计同比	1	银行间回购利率加权平均	6	CPI 食品同比	1
进出口金额:当月同比	1	银行间回购利率隔夜	6	RPI 食品同比	1
进口金额:当月同比	1	银行间回购利率 7 天	6	日用品类零售总额增速	1
产销率:累计值	1	准货币 - 储蓄存款	3	粮油类零售总额增速	2
CPI:当月同比	1			肉禽蛋类零售总额增速	4
RPI 环比	1				
CPI 同比	1				
RPI 同比	2				

资料来源：国信证券经济研究所

表 8：信息设备先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	1	M1、M2 增速差	1	包裹:当月同比	1
RPI:当月同比	1	CHIBOR 隔夜	5	订销报纸:当月同比	4
PPI:全部工业品:当月同比	2	CHIBOR7 天	5	汇票:当月同比	6
PPI:全部工业品:环比	1	法定准备金利率	4		
PPI:全部工业品:累计同比	1	银行间回购利率加权平均	5		
进出口金额:当月同比	1	银行间回购利率隔夜	5		
出口金额:当月同比	2	银行间回购利率 7 天	5		
进口金额:当月同比	1	人民币存款基准利率活期	4		
工业增加值:累计同比	1	人民币存款基准利率 3 个月	4		
工业增加值:重工业:累计同比	1	人民币存款基准利率 6 个月	2		
产销率:累计值	1	汇率港元	3		
CPI:当月同比	1	M1	1		
		活期存款	1		
		准货币 - 储蓄存款	5		

资料来源：国信证券经济研究所

表 9：有色金属先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	3	CHIBOR 加权平均	2	精铝产量同比	6
RPI:当月同比	1	CHIBOR 隔夜	6	锑选矿产品同比	1
PPI:全部工业品:当月同比	2	CHIBOR7 天	6	锑选矿产品环比	5
PPI:全部工业品:环比	2	法定准备金利率	5	进口数量:铜材:净进口	6
PPI:全部工业品:累计同比	2	银行间回购利率加权平均	6	进口数量:氧化铝:净进口	5
进口金额:当月同比	1	银行间回购利率 7 天	6	铜材产量同比	5
社会消费品零售总额:累计同比	2	人民币贷款基准利率 6 个月	4	氧化铝产量环比	6
产销率:累计值	1	人民币贷款基准利率 1 年	4		
CPI:当月同比	3	人民币贷款基准利率 1-3 年	4		
		人民币贷款基准利率 3-5 年	4		
		人民币贷款基准利率 5 年以上	4		
		人民币存款基准利率活期	4		
		人民币存款基准利率 3 个月	4		
		人民币存款基准利率 6 个月	4		
		人民币存款基准利率 1 年	4		
		M1	1		

资料来源：国信证券经济研究所

表 10：证券行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:当月同比	3	人民币贷款基准利率 1 年	2	深成指涨跌幅	2
CPI:环比	2	人民币贷款基准利率 1-3 年	2	深 100 涨跌幅	2
CPI:累计同比	1	人民币贷款基准利率 3-5 年	2		
RPI:当月同比	3	汇率欧元	1		
工业增加值:轻工业:当月同比	2				
社会消费品零售总额:当月同比	1				
社会消费品零售总额:累计同比	1				

资料来源：国信证券经济研究所

表 11：医药生物行业（剔除医疗服务）先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:当月同比	1	准货币 - 储蓄存款	6	出口金额:医用敷料:累计同比	6
CPI:累计同比	1	M1、M2 增速差	1	进口金额:抗菌素(制剂除外):累计同比	6
PPI:全部工业品:当月同比	2	CHIBOR 加权平均	5	出口平均单价:药材:当月值	4
PPI:全部工业品:环比	1	CHIBOR 隔夜	6	出口数量:中式成药:累计同比	6
PPI:全部工业品:累计同比	1	CHIBOR7 天	5	出口数量:医用敷料:累计值	4
RPI:当月同比	1	银行间回购利率加权平均	6	出口金额:医药品:累计值	4
出口金额:当月同比	2	银行间回购利率隔夜	6	出口金额:抗菌素(制剂除外):累计值	5
工业增加值:轻工业:当月同比	5	银行间回购利率 7 天	5	出口金额:医用敷料:累计值	4
进口金额:当月同比	1	银行间国债收益率 10 年期	5	出口金额:医疗仪器及器械:累计值	1
出口金额:中式成药:累计同比	6	人民币存款基准利率 3 个月	5	出口数量:医用敷料:累计同比	6
		人民币存款基准利率 6 个月	5	进口金额:抗菌素制剂:累计值	4
		人民币存款基准利率 1 年	5	产量:化学药品原药:累计同比	6
		人民币存款基准利率 2 年	5		
		人民币存款基准利率 3 年	5		
		人民币存款基准利率 5 年	5		

资料来源：国信证券经济研究所

表 12: 食品饮料行业先行因子筛选结果及先行阶数

宏观及货币指标		行业指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
宏观经济景气指数:先行指数	1	城市 CPI:食品:当月同比	1	产量:饮料酒:当月值	4
CPI:当月同比	3	城市 CPI:食品:肉禽及其制品:当月同比	6	出口数量:蘑菇罐头:当月值	1
CPI:累计同比	1	城市 CPI:食品:水产品:当月同比	1	出口数量:烘焙花生:当月值	1
PPI:全部工业品:当月同比	2	城市 CPI:食品:累计同比	1	出口金额:蘑菇罐头:当月值	1
PPI:全部工业品:环比	3	城市 CPI:食品:肉禽及其制品:累计同比	1	出口金额:烘焙花生:当月值	1
PPI:全部工业品:累计同比	1	城市 CPI:食品:水产品:累计同比	6	产量:啤酒:当月值	5
RPI:当月同比	1	农村 CPI:食品:当月同比	6	出口数量:猪肉罐头:累计同比	1
进口金额:当月同比	1	农村 CPI:食品:肉禽及其制品:当月同比	6	出口数量:烟草:累计同比	2
社会消费品零售总额:当月同比	1	农村 CPI:食品:水产品:当月同比	1	产量:果汁及果汁饮料:当月值	6
社会消费品零售总额:累计同比	1	农村 CPI:烟酒及用品:当月同比	1	出口金额:猪肉罐头:累计同比	3
M1、M2 增速差	2	农村 CPI:食品:累计同比	1	产量:液体乳:当月值	1
银行间国债收益率 20 年期	4	农村 CPI:食品:肉禽及其制品:累计同比	1	进口数量:纸烟:当月值	4
		农村 CPI:食品:水产品:累计同比	6	进口金额:纸烟:当月值	6
		农村 CPI:烟酒及用品:累计同比	1	进口数量:纸烟:累计值	4
		产量:饮料酒:当月同比	1	进口金额:纸烟:累计值	4
		产量:啤酒:当月同比	1	进口数量:食糖:累计同比	2
		产量:葡萄酒:当月同比	6	进口金额:食糖:累计同比	2
		产量:精制茶:当月同比	2	CPI:食品:当月同比	2
		产量:方便面:当月同比	1	CPI:食品:油脂:当月同比	6
		产量:酱油:当月值	2	CPI:食品:肉禽及其制品:当月同比	6
		产量:乳制品:当月值	2	CPI:食品:水产品:当月同比	1
		产量:食用植物油:累计同比	2	CPI:烟酒及用品:当月同比	1
		产量:葡萄酒:累计同比	4	CPI:烟酒及用品:酒类:当月同比	1
		产量:发酵酒精:当月值	1	CPI:食品:累计同比	1
		产量:精制茶:累计同比	1	CPI:食品:肉禽及其制品:累计同比	1

资料来源: 国信证券经济研究所

表 13: 建筑建材行业先行因子筛选结果及先行阶数

宏观及货币指标		行业指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	1	水泥价格指数:累计同比	5	房地产开发投资完成额:安装工程:累计同比	1
PPI:全部工业品:当月同比	1	木材价格指数:当月环比	6	产量:砖:当月同比	1
PPI:全部工业品:环比	1	木材价格指数:当月同比	3	产量:炆质砖:当月同比	6
M1	1	木材价格指数:当月比年初	6	产量:刨花板:当月同比	1
人民币贷款余额	5	木材价格指数:累计同比	1	产量:水泥:累计同比	1
CHIBOR 加权平均	5	木材价格指数:累计比年初	5	产量:水泥熟料:累计同比	1
CHIBOR 隔夜	6	原木价格指数:当月同比	3	产量:水泥混凝土排水管:累计同比	1
银行间国债收益率 10 年期	5	原木价格指数:当月比年初	6	产量:水泥混凝土电杆:累计同比	1
银行间国债收益率 20 年期	5	原木价格指数:累计同比	1	产量:砖:累计同比	1
		原木价格指数:累计比年初	5	产量:天然大理石建筑板材:累计同比	4
		胶合板价格指数:当月比年初	5	产量:刨花板:累计同比	1
		胶合板价格指数:累计同比	1	建材类价格指数:当月同比	1
		产量:水泥混凝土排水管:当月值	5	建材类价格指数:当月比年初	1
		出口金额:粘土及其他耐火矿物:当月值	3	建材类价格指数:累计同比	1
		出口金额:床垫、寝具及类似品:累计值	1	建材类价格指数:累计比年初	1
		出口数量:花岗岩石材及制品:累计同比	4	产量:窑外分解窑水泥熟料:当月值	5
		出口金额:花岗岩石材及制品:累计同比	4	产量:炆质砖:当月值	6
		房地产开发资金来源:利用外资累计同比 (%)	1	产量:水泥混凝土排水管:累计值	4

资料来源：国信证券经济研究所

表 14：银行行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
宏观经济景气指数:先行指数	1	1 年期-银行间国债收益率--即期 (%)	1	短期贷款-贷款余额 (十亿元人民币)	3
宏观经济景气指数:一致指数	6	3 年期-银行间国债收益率--即期 (%)	2		
CPI:当月同比	3	5 年期-银行间国债收益率--即期 (%)	2		
CPI:累计同比	1	7 年期-银行间国债收益率--即期 (%)	1		
PPI:全部工业品:当月同比	2	10 年期-银行间国债收益率--即期 (%)	1		
PPI:全部工业品:环比	3	20 年期-银行间国债收益率--即期 (%)	1		
PPI:全部工业品:累计同比	1	1 年期-银行间政策性金融债收益率--即期 (%)	1		
RPI:当月同比	3	3 年期-银行间政策性金融债收益率--即期 (%)	2		
进口金额:当月同比	1	5 年期-银行间政策性金融债收益率--即期 (%)	1		
社会消费品零售总额:当月同比	1	7 年期-银行间政策性金融债收益率--即期 (%)	1		
社会消费品零售总额:累计同比	1	10 年期-银行间政策性金融债收益率--即期 (%)	1		
		20 年期-银行间政策性金融债收益率--即期 (%)	1		
		人民币对外币汇率 - 月末-欧元	1		

资料来源：国信证券经济研究所

表 15: 纺织行业先行因子筛选结果及先行阶数

宏观及货币指标		行业指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:当月同比	1	产量:绒线(毛线):累计同比	2	出口金额:棉纱线:当月值:百分比	6
CPI:累计同比	1	产量:天然皮革制手提包(袋)、背包:累计同比	2	出口金额:丝织物:当月值:百分比	1
PPI:全部工业品:当月同比	1	出口金额:毛纺织物:累计同比	1	出口金额:玩具:当月值:百分比	4
PPI:全部工业品:累计同比	1	出口金额:丝织物:累计同比	1	出口数量:外底及鞋面均以橡胶或塑料制的鞋:当月值:百分比	4
RPI:当月同比	1	产量:印染布:当月同比	4	出口数量:塑料编织袋(周转袋除外):当月值:百分比	4
进口金额:当月同比	1	产量:绒线(毛线):当月同比	2	进口数量:针织或钩编织物:当月值:百分比	5
M1	1	出口数量:外底及鞋面均以橡胶或塑料制的鞋:累计同比	3	进口数量:合成短纤与棉混纺织物:当月值:百分比	6
人民币贷款余额	5	出口金额:玩具:累计同比	4	进口金额:棉织物:当月值:百分比	6
CHIBOR 加权平均	5	出口数量:棉纱线:当月值:百分比	6	进口金额:丝织物:当月值:百分比	1
CHIBOR 隔夜	6	产量:天然皮革制手提包(袋)、背包:当月同比	4	进口数量:二醋酸纤维丝束:累计同比	1
CHIBOR7 天	5	进口金额:合成短纤与棉混纺织物:累计同比	5	进口金额:二醋酸纤维丝束:累计同比	2
银行间回购利率加权平均	5			进口数量:毛纱线:累计同比	5
银行间回购利率隔夜	6			进口数量:针织或钩编织物:累计同比	4
银行间回购利率 7 天	5			进口金额:毛纱线:累计同比	5
CPI:当月同比	1	产量:绒线(毛线):累计同比	2	出口金额:棉纱线:当月值:百分比	6
CPI:累计同比	1	产量:天然皮革制手提包(袋)、背包:累计同比	2	出口金额:丝织物:当月值:百分比	1
PPI:全部工业品:当月同比	1	出口金额:毛纺织物:累计同比	1	出口金额:玩具:当月值:百分比	4

资料来源: 国信证券经济研究所

表 16: 化工行业先行因子筛选结果及先行阶数

宏观及货币指标		行业指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:当月同比	3	国际硝酸波罗的海 FOB	1	国际尿素	6
CPI:累计同比	1	国际硝酸黑海 FOB	1	合成氨	6
PPI:全部工业品:当月同比	2	甘氨酸	5	国内尿素出厂价	6
PPI:全部工业品:累计同比	1	纯 MDI 市场价	3	国内尿素批发价	6
进口金额:当月同比	1	甲苯	5	国内尿素零售价	6
宏观经济景气指数:一致指数	5	TDI	4	氯化铵	6
准货币 - 其他存款	4	己二酸	5	国内硫磺出厂价	6
CHIBOR 加权平均	6	电石	4	国际磷矿石/印度 CFR	6
CHIBOR 隔夜	6	乙炔法 PVC	6	宏福磷矿石	6
CHIBOR7 天	6	乙烯法 PVC	6	黄磷	6
银行间回购利率加权平均	6	乙炔法 PVC - 1.5*电石	6	磷酸	3
银行间回购利率隔夜	6	原盐/华东	6	三聚磷酸钠 (工业级)	2
银行间回购利率 7 天	6	甲醇	6	宏福二铵	6
银行间国债收益率 3 年期	2	DMF	5	金河一铵	6
银行间国债收益率 5 年期	2	煤焦油	6	氯化钾/温哥华 FOB	1
银行间国债收益率 7 年期	2	炭黑	5	氯化钾/东南亚 CFR	1
银行间国债收益率 10 年期	1	炭黑 - 2*煤焦油	5	丁苯橡胶	6
人民币存款基准利率 3 个月	2	纯苯	6		

资料来源: 国信证券经济研究所

表 17：保险行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
进口金额:当月同比	1	M2	5		
社会消费品零售总额:当月同比	1	准货币 - 定期存款	6		
社会消费品零售总额:累计同比	2	CHIBOR 隔夜	1		
宏观经济景气指数:先行指数	2	银行间回购利率隔夜	1		
CPI:当月同比	3	银行间国债收益率 30 年期	6		
CPI:环比	3	汇率日元	6		
CPI:累计同比	1				
PPI:全部工业品:当月同比	2				
PPI:全部工业品:环比	6				
RPI:当月同比	3				
工业增加值:当月同比	2				
工业增加值:轻工业:当月同比	2				
工业增加值:重工业:当月同比	2				

资料来源：国信证券经济研究所

表 18: 农林牧渔（剔除林业）行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
CPI:累计同比	1	出口金额:鸡肉制品:累计值:	6	出口数量:食用油籽:当月值:	1
PPI:全部工业品:当月同比	1	出口金额:冻鸡:累计值:	4	出口数量:大豆:当月值:	2
PPI:全部工业品:环比	2	出口金额:猪鬃:累计值:	4	出口数量:冻虾仁:当月值:	1
准货币 - 储蓄存款	6	出口金额:水海产品:累计值:	5	出口数量:天然蜂蜜:当月值:	6
M1、M2 增速差	2	出口金额:活鱼:累计值:	4	出口金额:大豆:当月值:	1
CHIBOR 加权平均	5	出口金额:墨鱼及鱿鱼:累计值:	5	出口金额:猪肉:当月值:	4
CHIBOR 隔夜	6	出口金额:虾:累计值:	5	农产品集贸市场价格指数:粳稻:环比	1
CHIBOR7 天	6	出口金额:鲜、冻对虾:累计值:	5	出口金额:天然蜂蜜:当月值:	6
银行间回购利率加权平均	5	出口金额:冻虾仁:累计值:	5	农产品集贸市场价格指数:花生油:环比	2
银行间回购利率隔夜	6	出口金额:冻鱼、冻鱼片:累计值:	5	出口数量:冻虾仁:当月值:	4
		出口金额:鲜蛋:累计值:	4	出口数量:天然蜂蜜:当月值:	5
		出口金额:植物油:累计值:	4	出口数量:棉花:当月值:	3
		出口金额:肠衣:累计值:	4	农产品集贸市场价格指数:豆油:环比	1
		出口金额:蜂蜜:累计值:	4	出口金额:活鱼:当月值:	1
		出口金额:天然蜂蜜:累计值:	4	农产品集贸市场价格指数:菜籽油:环比	1
		出口金额:棉花:累计值:	1	出口金额:天然蜂蜜:当月值:	5
		出口金额:辣椒干:累计值:	5	出口金额:棉花:当月值:	3
		出口金额:茶叶:累计值:	5	出口数量:稻谷和大米:累计值:	6
		农产品集贸市场价格指数:牛肉:当月同比	3	出口数量:鲜、干水果及坚果:累计值:	5
		出口数量:肠衣:累计同比	1	出口数量:橘、橙:累计值:	6
		出口数量:棉花:累计同比	3	出口数量:苹果:累计值:	6
		农产品集贸市场价格指数:烤烟:当月同比	1	出口数量:花生仁果:累计值:	6
		出口金额:活鱼:累计同比	1	出口数量:花生、花生仁:累计值:	6
		出口金额:棉花:累计同比	3	出口数量:干的食用菌类:累计值:	6
		进口金额:谷物及谷物粉:当月值	6	出口数量:大蒜:累计值:	5
		进口金额:谷物及谷物粉:累计值	6	出口数量:松子仁:累计值:	5
		进口金额:稻谷和大米:累计值:	6	出口数量:食用油籽:累计值:	6
		农产品集贸市场价格指数:豆油:当月同比	4	出口数量:活家禽:累计值:	4
		进口金额:小麦粉:累计值:	5	出口数量:猪肉:累计值:	5
		进口金额:鲜、干水果及坚果:累计值:	5	出口数量:鸡肉:累计值:	5
		进口金额:香蕉(包括芭蕉):累计值:	4	出口数量:冻鸡:累计值:	4
		进口金额:食用植物油:累计值:	6	出口数量:猪鬃:累计值:	4
		进口金额:菜籽油和芥子油:累计值:	4	出口数量:水海产品:累计值:	5
		进口金额:豆油:累计值:	6	出口数量:活鱼:累计值:	4

进口金额:冻鱼:累计值:	6	出口数量:墨鱼及鱿鱼:累计值:	5
农产品集贸市场价格指数:大豆:当月同比	6	出口数量:虾:累计值:	5
进口数量:食用植物油:累计值:	6	农产品集贸市场价格指数:仔猪:当月同比	3
进口数量:豆油:累计值:	6	出口数量:鲜、冻对虾:累计值:	5
进口数量:稻谷和大米:累计值:	6	出口数量:冻虾仁:累计值:	5
进口数量:鲜、干水果及坚果:累计值:	6	出口数量:冻鱼、冻鱼片:累计值:	5
进口数量:香蕉(包括芭蕉):累计值:	5	出口数量:鲜蛋:累计值:	4
进口数量:冻鱼:累计值:	6	出口数量:食用植物油(包括棕榈油):累计值:	5
农产品批发价格总指数:同比	3	出口数量:豆油:累计值:	1
进口数量:菜籽油和芥子油:累计同比	4	出口数量:肠衣:累计值:	4
进口金额:谷物及谷物粉:累计同比	1	出口数量:蜂蜜:累计值:	4
进口金额:大麦:累计同比	1	出口数量:天然蜂蜜:累计值:	4
进口金额:菜籽油和芥子油:累计同比	4	农产品集贸市场价格指数:猪肉:当月同比	6
农产品集贸市场价格指数:仔猪:累计同比	1	出口数量:辣椒干:累计值:	6
进口金额:大麦:当月值:	6	出口数量:茶叶:累计值:	5
农产品集贸市场价格指数:猪肉:累计同比	6	出口金额:谷物及谷物粉:累计值:	5
进口金额:大豆:当月值:	1	出口金额:稻谷和大米:累计值:	4
农产品集贸市场价格指数:油菜籽:累计同比	6	出口金额:鲜、干水果及坚果:累计值:	5
农产品集贸市场价格指数:大豆:累计同比	1	出口金额:苹果:累计值:	6
平均价:二级豆油:	4	出口金额:花生仁果:累计值:	5
平均价:二级菜籽油:	1	出口金额:花生、花生仁:累计值:	5
平均价:3等粳稻谷:	4	出口金额:干的食用菌类:累计值:	6
平均价:2等粳稻谷:	5	出口金额:蔬菜:累计值:	5
平均价:2等白小麦(普通):	6	出口金额:鲜或冷藏蔬菜:累计值:	5
越南大米报价:5%破碎率:	5	出口金额:大蒜:累计值:	5
泰国大米报价:5%蒸谷米:	4	出口金额:松子仁:累计值:	6
泰国大米报价:5%破碎率:	6	出口金额:活家禽:累计值:	4
泰国大米报价:100% B级:	6	出口金额:牛肉:累计值:	4
美国大米报价:南方碾磨中粒米:	1	出口金额:猪肉:累计值:	5
农产品集贸市场价格指数:牛肉:累计同比	1	出口金额:鸡肉:累计值:	5

资料来源: 国信证券经济研究所

表 19: 家用电器行业先行因子筛选结果及先行阶数

宏观指标		货币指标		行业指标	
先行指标	阶数	先行指标	阶数	先行指标	阶数
产销率:当月值	2	法定准备金率	4	家用燃气灶具:当月值	6
产销率:累计值	3	准货币-定期存款	6	家用电热烘烤器具:累计值	6
宏观经济景气指数:先行指数	1	准货币-其他存款	1	美国:密歇根大学消费者信心指数	6
CPI:当月同比	6	贷款余额/存款余额	4	美国:世界大型企业联合会:消费者信心指数	6
CPI:环比	4	电气空间加热器:累计量	4	日本:消费者信心指数(排除一人户家庭)	6
CPI:累计同比	1	汇率美元	6	英国:NationalWide:消费者信心指数:定基数(2004=100)	6
PPI:全部工业品:当月同比	2	汇率港元	3	房屋施工面积:累计同比:Divide(constant)	1
PPI:全部工业品:累计同比	1	汇率日元	6	商品房销售额:现房:累计值:ADD:PERCENTCHG:Divide(constant)	2
RPI:当月同比	5			商品房销售面积:累计值:DIFFERENCE	5
工业增加值:累计同比	1			房屋施工面积:累计值:DIFFERENCE	5
工业增加值:轻工业:当月同比	6			商品房销售额:现房:累计值:ADD:DIFFERENCE	5
工业增加值:重工业:累计同比	1			完成开发土地面积:累计值:DIFFERENCE	5
进出口金额:当月同比	1			美国:耐用品:未完成订单	5
进口金额:当月同比	1			电热水器:累计量	5
社会消费品零售总额:当月同比	1			瓦斯煎烤器:当月量	6
社会消费品零售总额:累计同比	2			洗碗机:当月量	5
				零售价格:36个城市:月平均:抽油烟机:当地主销:直吸式	5
				零售价格:36个城市:月平均:燃气灶具:当地主销	4
				零售价格:36个城市:月平均:电热水器:当地主销:50L 钢化搪瓷内胆	6
				零售价格:36个城市:月平均:燃气热水器:当地主销:8L	5
				零售价格:36个城市:月平均:彩电:康佳:29寸:纯平	6
				零售价格:36个城市:月平均:彩电:当地主销:29寸:纯平	2
				家用洗衣机:当月值	6
				房间空调器:当月值	6
				家用吸排油烟机:当月值	6
				家用吸排油烟机:累计值	2

资料来源: 国信证券经济研究所

单个行业预期收益模型

在进行完先行因子的筛选后，下一步便是根据这些因子建立单个行业的预期收益模型，进行行业收益率的预测，为我们进行行业配置做基础准备。

在行业预期收益模型的建立方法上，我们采用多元线性回归的方法，因为回归模型是一种比较成熟的预测模型，也是在预测过程中使用较多的模型。这里，我们遇到了需要解决的第二个问题——多重共线性。

多重共线性的存在

多重共线性指的是多元回归模型的自变量间存在近似的线性关系，它的存在使得估计的精确性大幅降低，估计值稳定性变差，甚至在回归方程整体高度显著时，一些回归系数通不过显著性检验，正负号倒置，使得无法从回归方程得到合理的经济解释，降低回归方程的应用价值。

这里，我们以有色行业为例，诊断一下该多元回归模型是否存在多重共线性。

由于变量过多，通过方差比例（Variance Proportion）表和条件数（Condition Index）不易观察多重共线性，我们可以通过方差膨胀因子（VIF）简单观察一下。下表是部分因子的回归结果，并不完整。

表 20：多重共线性诊断（部分因子结果）

因子	参数估计值	标准误	t 值	Pr> t	VIF
Intercept	-0.8438	6.2925	-0.13	0.8941	0
pp1	2.1576	4.1863	0.52	0.6096	97.8303
pp2	7.3752	5.3500	1.38	0.1770	173.5304
pp3	-4.1415	1.5746	-2.63	0.0127	43.0399
pp4	-5.0239	2.8103	-1.79	0.0827	4.0502
pp5	0.6685	1.7167	0.39	0.6994	44.4384
pp6	0.6152	0.4207	1.46	0.1529	47.4191
pp7	0.5519	0.2334	2.36	0.0239	17.1726
pp8	-0.3353	0.2474	-1.36	0.1843	19.6371
pp9	12.6517	13.4350	0.94	0.3530	13.4701
pp10	-8.1410	6.8477	-1.19	0.2427	277.1130
pp11	0.0091	0.0036	2.54	0.0158	1.5035

资料来源：国信证券经济研究所

显然，除pp4和pp11外，其他变量的VIF值均远远超过10，说明存在严重的多重共线性。再来看p值，通过显著性检验的变量也是少之又少，很多变量无法准确地对模型做出解释。而同时，这一回归的R方则高达0.7390，再次说明了多重共线性对整个回归方程的影响。

多重共线性的处理

由于解释变量间完全不相关的情形是非常少见的，大多数变量都在某种程度上存在

着一定的共线性，多重共线性是进行多元线性回归时普遍存在和需要解决的问题。而在处理内在关联性非常强的行业基本面因子时，这一问题就更为显著。

图 4 简要概括了多重共线性的概念、后果、判断方法及主要处理方法，本报告就不再一一进行详细阐述，只重点介绍下多重共线性的几种主要处理方法。

图 4：多重共线性简介

多重共线性的概念	多重共线性的后果	多重共线性的判断方法	多重共线性的主要处理方法
<ul style="list-style-type: none"> 线性回归模型中某一个解释变量与其他解释变量间存在线性关系，就称线性回归模型中存在多重共线性 	<ul style="list-style-type: none"> 增大最小二乘估计量的方差； 参数估计值不稳定，对样本变化敏感 检验可靠性降低，产生弃真错误。 	<ul style="list-style-type: none"> 直观判断方法 方差扩大因子法 (VIF)。VIF 超过 10，常常表示存在多重共线性 	<ul style="list-style-type: none"> 逐步回归法 岭回归法 主成分分析法

资料来源：国信证券经济研究所

逐步回归法。逐步回归法是指利用被解释变量 Y 对每一个解释变量 X_i 作一个回归方程，构造统计量，进行统计检验，并根据相应的经济理论进行解释，从中选取最优的回归方程；然后逐步引入其他的解释变量，再做相应的回归方程，扩大模型的规模，同时对所有解释变量的回归系数进行检验。

主成分分析法。主成分分析法是通过线性变换，将原来的多个指标组合成相互独立的少数几个能充分反映总体信息的指标，从而在不丢掉重要信息的前提下避开变量间共线性问题，便于进一步分析。在主成分分析中提取出的每个主成分都是原来多个指标的线性组合。比如有两个原始变量 X_1 和 X_2 ，则一共可提取出两个主成分：

$$\begin{cases} Z_1 = b_{11}X_1 + b_{21}X_2 \\ Z_2 = b_{12}X_1 + b_{22}X_2 \end{cases}$$

原则上如果有 n 个变量，则最多可以提取出 n 个主成分，但如果将它们全部提取出来就失去了该方法简化数据的意义。一般情况下提取出 2-3 个主成分（已包含了 90% 以上的信息），其他的可以忽略不计。

主成分分析的基本原理是将解释变量转换成若干个主成分，这些主成分从不同侧面反映解释变量的综合影响，并且互不相关。因此，可以将被解释变量关于这些主成分进行回归，再根据主成分与解释变量之间的对应关系，求得原回归模型的估计方程。

岭回归估计法。Horel (1970) 提出了岭回归估计方法，该方法放弃最小二乘的无偏性，损失部分信息，以放弃部分精确度为代价来寻求效果稍差但更符合实际的回归方程。

岭回归分析的基本思想是当自变量间存在共线性时，解释变量的相关矩阵行列式近似为零， $X'X$ 是奇异的，也就是说它的行列式的值也接近于零，此时 OLS 估计将失效。此时可采用岭回归估计。岭回归就是用 $(X'X + kI)$ 代替正规方程中的 $X'X$ ，人为地把最小特征根由 $\min \lambda_i$ 提高到 $\min (\lambda_i + k)$ ，希望这样有助于降低均方误差。这时 β 的岭估计定义为：

$$b(k) = (X'X + kI)^{-1} X'Y$$

上式作为 β 的岭估计，当 $k=0$ 时即为通常所说的最小二乘估计（OLS 估计）。

当 $k \rightarrow \infty$ 时， $b(k) \rightarrow 0$ ，一般情况下， k 取 0 到 1 之间的数值。

$(X'X + kI)$ 表示在 $X'X$ 的矩阵对角线上每个元素都加上一个正数，就象形成一个山脊一样，用式 $b(k)$ 来估计 β ，这种估计参数的方法，称为岭回归估计法， k 为岭回归系数。

对于上述三种方法，逐步回归法会对解释变量进行删除，在一些经济模型中，要求一些很重要变量必须包含在里面，这时如果贸然地删除就不符合现实的经济意义；主成分分析法只是从解释变量中寻找主成分，与被解释变量无关，因而主成分与被解释量在算法上关系不密切，从而导致最后主成分在实际应用中无法更好的进一步拟合被解释变量。在下一步的预期收益模型建立中，我们采用的是岭回归方法。

数据及回归方法说明

1. 源数据：新分类下的行业指数收益率序列，行业先行基本面因子
2. 数据频度：月
3. 样本区间：2005 年 1 月至 2010 年 10 月，起始观察点为 2005 年 1 月，最短观察期长度为 36 个月，递月增加观察期长度，预测区间为一个月
4. 回归方法：岭回归
5. K 值选择：为了既满足方差膨胀因子（VIF）小于 10，又符合 k 值越小越好的原则，我们将对比各 k 值下的的各变量 VIF 值，选取满足所有变量 VIF 值均小于 10 的最小的 k 值。

预测收益模型建立

在对行业收益数据及各先行因子滞后先行阶数后的滞后数据进行岭回归后，我们便得到了各解释变量的系数，而这些系数就为我们下步的预测搭好了框架。原来的回归模型中，各变量的系数都已经悉数得到，只要代入下一月的各变量数值，即可得到下期的预测收益值。假设一共有 m 个先行因子，则下式是岭回归时的方程：

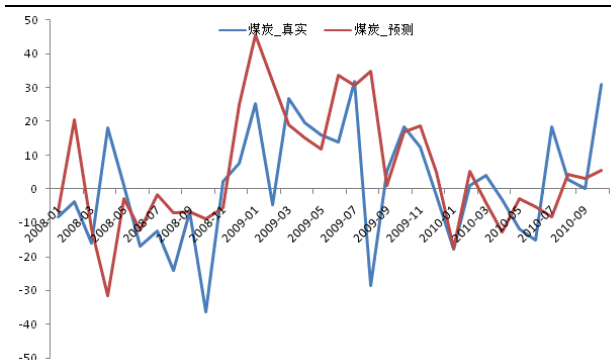
$$y_t = intercept + b_1 * pp_{1t} + b_2 * pp_{2t} + \dots + b_m * pp_{mt} + \varepsilon_t$$

得到系数后，我们只要根据 $t+1$ 时刻（即下月）的变量数值，即可预测出下月的收益 y 。

这里需要解释的是，为什么我们能够根据模型预测收益。这主要取决于我们所选取的因子特性，我们的因子均为先行因子，而先行阶数至少为 1，也就是说我们本月进行回归时所用到的数据，最晚也是上一个月的数据。仅仅根据历史的数据我们就可以得到最新的回归模型参数，而预测下个月的收益所需要的数据我们本月也一定是可以获得的，这就使得我们能够顺利地进行收益预测，而无须担心会用到未来数据的原因了。

让我们来看一下部分行业预测的结果如何。图表分别为真实收益率与预测收益率的对比，真实收益率与预测收益率的相关性。

图 5: 煤炭行业预测收益率对比



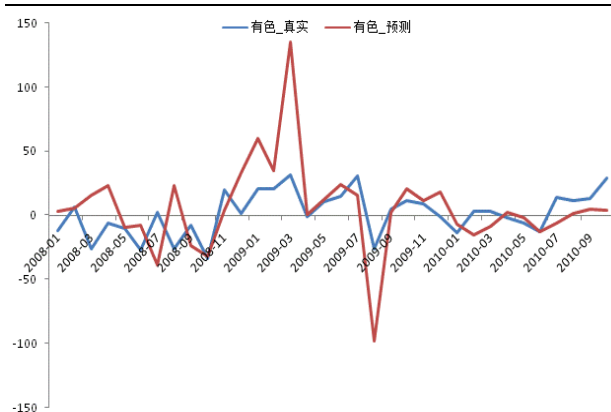
资料来源: 国信证券经济研究所

表 21: 煤炭行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.4833	0.1745

资料来源: 国信证券经济研究所

图 6: 有色金属行业预测收益率对比



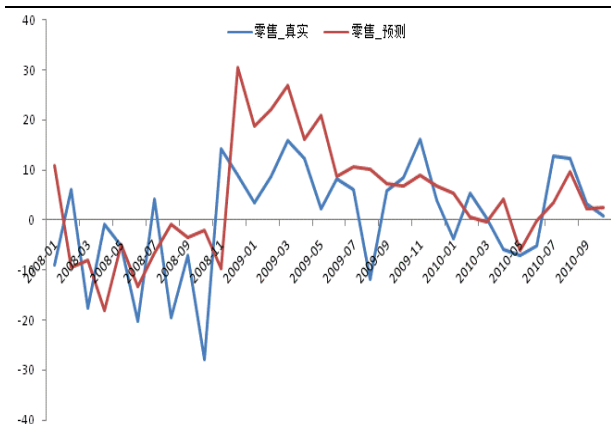
资料来源: 国信证券经济研究所

表 22: 有色金属行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.5365	0.8112

资料来源: 国信证券经济研究所

图 7: 商业零售行业预测收益率对比



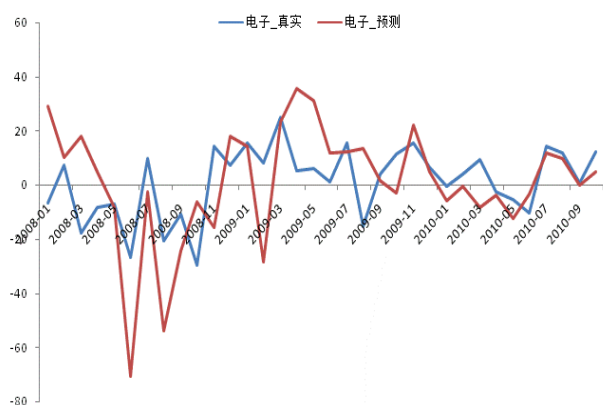
资料来源: 国信证券经济研究所

表 23: 商业零售行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.5470	0.5943

资料来源: 国信证券经济研究所

图 8: 电子元器件行业预测收益率对比



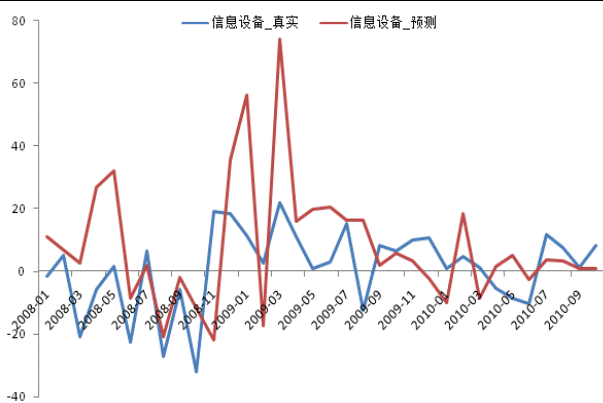
资料来源: 国信证券经济研究所

表 24: 电子元器件行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.6938	0.0081

资料来源: 国信证券经济研究所

图 9: 信息设备行业预测收益率对比



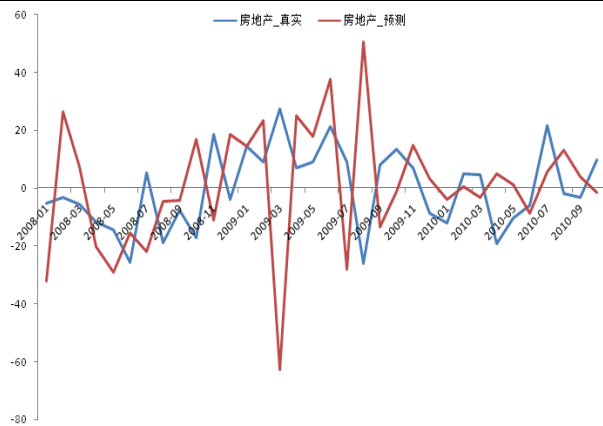
资料来源: 国信证券经济研究所

表 25: 信息设备行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.1929	-0.0663

资料来源: 国信证券经济研究所

图 10: 房地产行业预测收益率对比



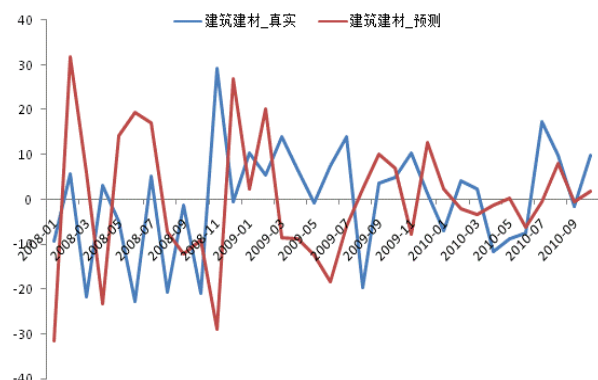
资料来源: 国信证券经济研究所

表 26: 房地产行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.0889	0.6078

资料来源: 国信证券经济研究所

图 11: 建筑建材行业预测收益率对比



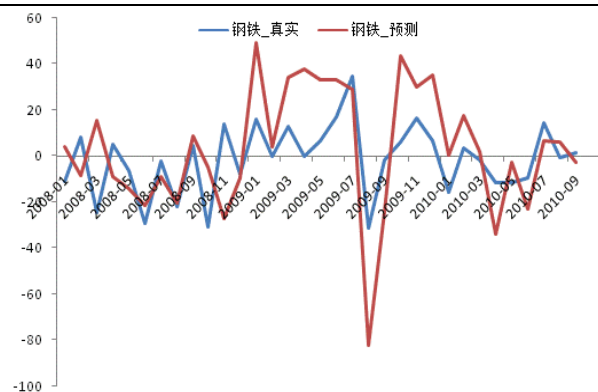
资料来源: 国信证券经济研究所

表 27: 建筑建材行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.3607	0.6071

资料来源: 国信证券经济研究所

图 12: 钢铁行业预测收益率对比



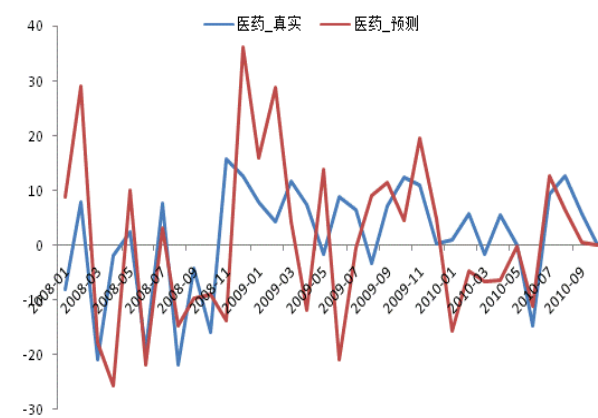
资料来源: 国信证券经济研究所

表 28: 钢铁行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.5535	-0.0022

资料来源: 国信证券经济研究所

图 13: 医药行业预测收益率对比



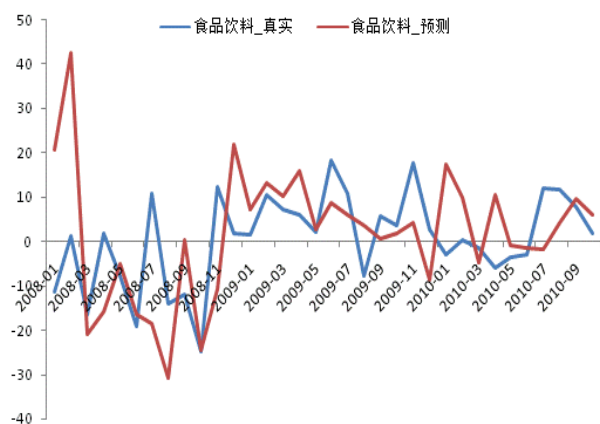
资料来源: 国信证券经济研究所

表 29: 医药行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.6116	-0.2270

资料来源: 国信证券经济研究所

图 14: 食品饮料行业预测收益率对比



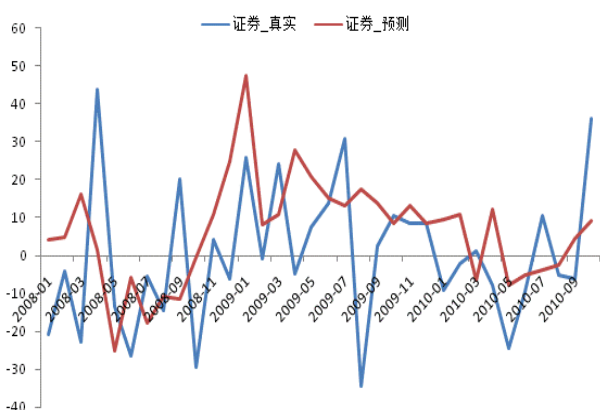
资料来源: 国信证券经济研究所

表 30: 食品饮料行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	-0.1550	0.3879

资料来源: 国信证券经济研究所

图 15: 证券行业预测收益率对比



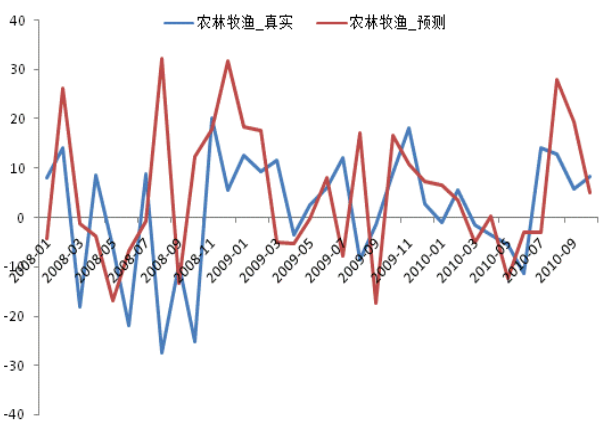
资料来源: 国信证券经济研究所

表 31: 证券行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.2866	0.1907

资料来源: 国信证券经济研究所

图 16: 农林牧渔行业预测收益率对比



资料来源: 国信证券经济研究所

表 32: 农林牧渔行业预测收益率准确率

样本外区间	真实-预测相关性	真实-滞后一期预测相关性
2010.01-2010.10	0.5604	0.7908

资料来源: 国信证券经济研究所

从上述图表，我们不难发现，2010年后的样本外区间，各行业预测收益率与真实收益率的相关性非常强，只有房地产行业 and 食品饮料行业的预测效果不尽如人意。不过大部分行业的预测准确性都相当高，说明该模型对于单行业的收益预测具有一定的效果。

再来对比一下当期预测收益率与真实收益率的相关性和滞后一期预测收益率与真实收益率的相关性。两种相关性的对比，实际上是为了证明预测收益率并不是通过所求得参数的变化而变量值不变造成的，而应该是我们所求的参数稳定情况下，通过先行因子值的变化而准确估计到的。最终是为了证明我们的预测还是比较准确的。通过各行业具体的比较，我们欣慰地看到，还是有不少行业表现出了足够的预测准确性。尤其是煤炭、电子元器件、钢铁、医药这些行业，较高的当期预测收益率与真实收益率的相关性以及较低的滞后一期预测收益率与真实收益率的相关性的反差，无不表明该模型预测的有效性。

单行业到行业配置的可行性分析

上面的分析可以看出，单行业的收益预测的确具有一定的准确性和应用意义，那么如果就诸多行业联系起来而进行进一步的行业配置，是否可行呢？下面我们尝试通过行业间的预测收益排名进行真实的行业排名预测，来讨论实施行业配置的可行性。这里我们对一共15个行业的预测情况进行了分析，并结合15个行业的相关结果进行了行业配置的测试。

具体测试方式为：

每月首先根据各个行业的预测收益率进行排序，选出收益排名前三的为多头，收益排名后三的为空头，然后同期对比根据各个行业的真实收益率进行排序后的多头与空头，观察我们的模型预测成功的比例。

表 33：行业配置涨跌幅预测命中率统计

月度	多头命中率	空头命中率	多空命中率
2010-01	0.00%	66.67%	33.33%
2010-02	0.00%	0.00%	0.00%
2010-03	0.00%	33.33%	16.67%
2010-04	0.00%	33.33%	16.67%
2010-05	33.33%	0.00%	16.67%
2010-06	66.67%	66.67%	66.67%
2010-07	0.00%	33.33%	16.67%
2010-08	33.33%	33.33%	33.33%
2010-09	33.33%	0.00%	16.67%
2010-10	33.33%	33.33%	33.33%

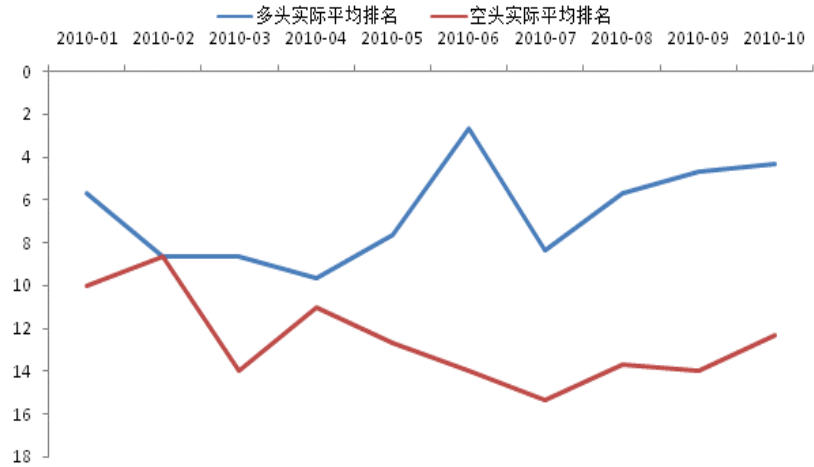
资料来源：国信证券经济研究所

可以看到，2010年5月份以后，模型的预测命中率是比较不错的，尤其是对于多头的判断，非常准确。我们预测出的涨跌幅前三的行业命中率曾一度高达66.67%，效果令人满意。

由于这里计算的命中率过于严格，可能会忽略很多信息，如预测排名前三的多头，其真实排名虽然不是前三，但很可能是4至6名，也是预测比较准确的或效果比较好的，但命中率却是0，容易使人误解。为此，我们还对多头组合及空头组合的实际

排名进行了测试。

图 17: 行业配置多空实际排名统计



资料来源: 国信证券经济研究所

事实证明, 在2010年, 多头组合的实际排名一直处于高位, 虽然从多头命中率上来看, 并没有异常突出的表现, 但多头实际排名的靠前位置还是足以说明所选行业的有效性。尤其是对于空头组合, 测试的效果更为理想, 排名一直稳定在低位, 且从未超过过多头组合。多头战胜空头的排名命中率赫然是100%!

此外, 我们为了进一步检测关于模型行业配置的命中率, 又设计了一个实验, 具体步骤为:

每月首先根据各个行业的预测收益率涨跌速进行排序(涨跌速为本月收益率相较上月收益率的变化速度), 选出涨跌速排名前三的为多头, 涨跌速排名后三的为空头, 然后同期对比根据各个行业的真实收益率涨跌速进行排序后的多头与空头, 观察我们的模型预测成功的比例。

表 34: 行业配置涨跌速预测命中率统计

月度	多头命中率	空头命中率	多空命中率
2010-01	0.00%	0.00%	0.00%
2010-02	0.00%	0.00%	0.00%
2010-03	33.33%	0.00%	16.67%
2010-04	0.00%	66.67%	33.33%
2010-05	0.00%	0.00%	0.00%
2010-06	33.33%	33.33%	33.33%
2010-07	0.00%	33.33%	16.67%
2010-08	33.33%	33.33%	33.33%
2010-09	33.33%	66.67%	50.00%
2010-10	0.00%	33.33%	16.67%

资料来源: 国信证券经济研究所

结果显示, 空头组合的命中率显然要高于多头组合, 跌速最快(或涨速最慢)的行业更容易被成功预测。

如果我们采用根据预测收益率排名的多头组合或空头组合来进行行业配置的话, 下面我们来看一下每月等权重配置多头及空头行业的收益情况。

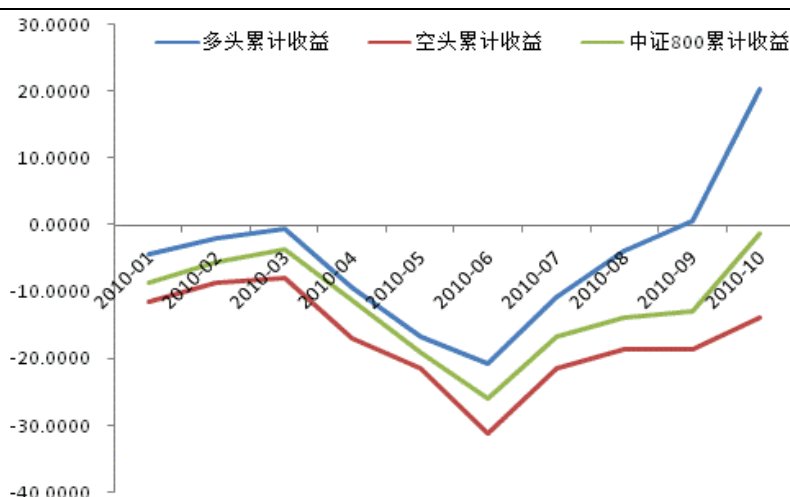
表 35: 预测收益率行业配置收益情况统计

月度	多头组合收益	多头累计收益	空头组合收益	空头累计收益
2010-01	-4.3875	-4.3875	-11.6078	-11.6078
2010-02	2.5418	-1.9572	3.4266	-8.5789
2010-03	1.5354	-0.4519	0.7252	-7.9159
2010-04	-8.9608	-9.3722	-9.7289	-16.8747
2010-05	-8.0317	-16.6511	-5.6979	-21.6111
2010-06	-4.9954	-20.8147	-12.3539	-31.2952
2010-07	12.7181	-10.7439	14.0933	-21.6124
2010-08	7.6249	-3.9382	3.6482	-18.7527
2010-09	4.7411	0.6161	0.0701	-18.6957
2010-10	19.6616	20.3988	6.0040	-13.8142

资料来源: 国信证券经济研究所

很显然, 多头组合表现远远要好于空头组合, 且表现比较稳定。在诸多测试中, 都表明我们所选出的多头组合的确具有一定的预测准确性, 可以应用于实际配置之中。

图 18: 预测收益率行业配置收益情况对比



资料来源: 国信证券经济研究所

模型应用及未来优化

模型解释

在我们前面的分析中, 可能很多先行因子被挖掘出来, 并没有得到足够的重视, 甚至只是被看做一种回归分析中的偶然现象。其实不然, 我们挖掘出的很多的先行因子, 虽然其中一些只是单纯的Granger因果关系, 但仔细观察, 不难发现, 很多因子不仅因子本身具有一定的先行特征, 所得到的先行阶数也都具有一定的经济学含义。

下面列举对模型结果的一些经济学解释:

1. **CPI: 当月同比。**CPI指标的重要性不言而喻,很多行业都与之有直接的联系。在我们的研究中,该指标也表现得非常突出,几乎所有行业的先行指标中都有它的出现。唯一的区别就是个别行业在先行阶数的上可能有所不同,但大部分行业的先行阶数均为1,说明CPI指标对于行业的反应常常是立竿见影。CPI反映了市场通货膨胀的程度,而有关通胀的预期常常会使市场进入提前反应、提前透支的阶段,通过各行业指数的波动影响市场。
2. **货币因子。**货币因子中大部分都体现了货币政策的影响,且我们的指标均经过处理,体现的是货币政策的月度变动。而先行因子集的结果表明,大部分货币政策的先行性比较强,先行阶数较高,从另一个角度反映的是货币政策以月为观察单位,并没有立即对各行业造成显著的影响。但这并不是说没有影响,影响的确存在,只是作用的时间比较缓慢。这与一些研究中,货币政策对证券市场短期影响有限,而长期影响比较稳定的结论,是比较符合的。
3. **行业预测效果差异:**我们看到煤炭、钢铁、有色、医药等行业的预测效果非常可观,周期性比较强,能够从因子的变化中比较准确地推测出行业的走势;而诸如房地产、食品饮料行业的效果就差强人意。这其实是有道理的。像煤炭、钢铁、有色这样的行业,行业因子里面或者包括了上下游的情况,或者包括了进出口的情况,或者包括了产量销量等信息,都是能反映行业运行的关键因子,且这种因子的影响是长期存在的,而不是偶然或阶段现象。而房地产、食品饮料就比较特殊,如房地产,各种因素对其产生的影响效果总是变化的,在不同阶段可能影响程度都不尽相同,这就使得我们的模型效果大受影响,预测准确性有所欠缺。

模型应用及改进

在目前阶段,我们对于先行因子的使用及预测模型的运用,主要集中于行业收益率的预测,并初步根据预测收益率对各行业进行排序,以预测热点行业及危险行业。关于先行因子,我们在进行因果Granger检验时是将因子与对应的行业相匹配的所有历史数据进行回归,所以并不只是最近几年的偶然规律,而是长期历史的规律体现,故当我们得到先行因子后,并不会频繁地改变,只会在未来恰当时机做出调整。先行阶数也是同样道理。

关于预测模型,根据2010年后的检验效果来看,我们完全可以将之作为市场两端行业的判断依据,根据模型预测的收益率排序来进行行业的初步筛选和甄别,聚焦热点行业、规避危险行业。

不过,根据目前的进展状况,我们尚不会将之完全应用于行业配置的策略之中,还将通过更长时间更多数据的样本外检验,以证明我们的模型设计及因子结果已足够稳定和成熟。

未来一段时间,我们仍将从以下几点着力于该模型的完善:

1. **原始基本面因子集的系统化整理和搜集:**任何量化策略的制定必须依靠数据的强力支持。虽然在众多同仁的鼎力支持下,我们历经数月完成了各行业基本面因子的海量搜集和处理,并提出了较为完整的行业预测收益模型,但为了使我们的模型更具说服力并经得起市场的考验,我们必须对数据提出更高的要求,包括数据源、数据格式、数据发布时间、数据时效等。
2. **先行因子的稳定性检验:**通过一定样本内的反复检验,我们有理由相信先行因子的稳定性并不会因为检验区间的变化而有所动摇。但在原始基本面因子集尚

未完全固定下来之时，我们也不能贸然认为我们已找到了最优的先行因子，所以对于先行因子的研究仍会成为我们工作的重点。

3. 样本外检验的持续性：由于我们研究的对象毕竟是基本面因子，受限于因子数据的复杂和数据质量的有待考验，我们不会过分停留于过往的研究和结论，而会始终坚持严格的样本外检验，以观察和完善我们的模型，坚持研究思路和逻辑的贯彻。

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±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±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

风险提示

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。

证券投资咨询业务的说明

证券投资咨询业务是指取得监管部门颁发的相关资格的机构及其咨询人员为证券投资者或客户提供证券投资的相关信息、分析、预测或建议，并直接或间接收取服务费用的活动。

证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。

国信证券经济研究所团队成员

宏观		固定收益		策略	
周炳林	0755-82130638	李怀定	021-60933152	黄学军	021-60933142
林松立	010-66026312	侯慧梯	021-60875161	闫 莉	010-88005316
崔 嵘	021-60933159	张 旭	010-66026340	吴土金	0755-82130833-1332
交通运输		银行		房地产	
郑 武	0755- 82130422	邱志承	021- 60875167	方 焱	0755-82130648
陈建生	0755- 82133766	黄 飙	0755-82133476	区瑞明	0755-82130678
岳 鑫	0755- 82130432	谈 焯	010- 66025229	黄道立	0755- 82133397
高 健	0755-82130678	汽车及零配件		钢铁	
机械		李 君	021-60933156	郑 东	010- 66025270
郑 武	0755- 82130422	左 涛	021-60933164	秦 波	010-66026317
余爱斌	0755-82133400	基础化工		医药	
陈 玲	0755-82130646	张栋梁	0755-82130532	贺平鸽	0755-82133396
杨 森	0755-82133343	陈爱华	0755-82133397	丁 丹	0755- 82139908
商业贸易		邱 斌	0755-82130532	杜佐远	0755-82130473
孙菲菲	0755-82130722	罗 洋	0755-82150633	传媒	
祝 彬	0755-82131528	吴琳琳	0755-82130833-1867	陈财茂	021-60933163
田惠蓝	0755-82133263	梁 丹	0755- 82134323	非银行金融	
石油与石化		电力设备与新能源		邵子钦	0755- 82130468
严蓓娜	021-60933165	杨敬梅	021-60933160	田 良	0755-82130513
有色金属		电力与公用事业		童成敦	0755-82130513
彭 波	0755-82133909	徐颖真	021-60875162	家电	
谢鸿鹤	0755-82130646	谢达成	021-60933161	王念春	0755-82130407
通信		造纸		黄海培	021-60933150
严 平	021-60875165	李世新	0755-82130565	纺织服装	
程 锋	021-60933167	邵 达	0755-82130706	方军平	021-60933158
计算机		电子元器件		食品饮料	
段迎晟	0755- 82130761	段迎晟	0755- 82130761	黄 茂	0755-82138922
农业		黄 磊	0755-82151833	建筑	
张 如	021-60933151	高耀华	0755-82130771	邱 波	0755-82133390
建材		旅游		刘 萍	0755-82130678
新兴产业		陈财茂	021-60933163	指数与产品设计	
陈 健	010-66022025	煤炭		焦 健	0755-82133928
李筱筠	010-66026326	研究支持		彭甘霖	0755-82133259
投资基金		沈 瑞	0755-82132998	阳 瑾	0755-82133538
杨 涛	0755-82133339	雷 达	0755-82132098	周 琦	0755-82133568
刘舒宇	0755-82133568	袁 剑	0755-82139918	赵学昂	0755-66025232
康 亢	010-66026337	余 辉	0755-82130741	邓 岳	0755- 82150533
刘 洋	0755-82150566	王越明	0755-82130478	交易策略	
李 腾	0755-82130833-6223	数量化投资		戴 军	0755-82133129
		葛新元	0755-82133332	秦国文	0755-82133528
		董艺婷	021-60933155	徐左乾	0755-82133090
		林晓明	0755-25472656	黄志文	0755-82133928
		赵斯尘	021-60875174		
		程景佳	010-88005326		
		郑 云	021-60875163		
		毛 甜	021-60933154		

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		魏 宁	0755-82133492 13823515980 weining@guosen.com.cn	
王晓建	010-66026342 13701099132 wangxj@guosen.com.cn		马小丹	021-60875172 13801832154 maxd@guosen.com.cn		邵燕芳	0755-82133148 13480668226 shaoyf@guosen.com.cn	
焦 戡	010-66026343 13601094018 jiaojian@guosen.com.cn		郑 毅	021-60875171 13795229060 zhengyi@guosen.com.cn		林 莉	0755-82133197 13824397011 linli2@guosen.com.cn	
徐文琪	010-66026341 13811271758 xuwwq@guosen.com.cn		黄胜蓝	021-60875166 13761873797 huangsl@guosen.com.cn		王昊文	0755-82130818 18925287888 wanghaow@guosen.com.cn	
李文英	010-88005334 13910793700 liwying@guosen.com.cn		刘 塑	021-60875177 13817906789 liusu@guosen.com.cn		甘 墨	0755-82133456 15013851021 ganmo@guosen.com.cn	
赵海英	010-66025249 13810917275 zhaohy@guosen.com.cn		叶琳菲	021-60875178 13817758288 yelf@guosen.com.cn		段莉娟	0755-82130509 18675575010 duanlj@guosen.com.cn	
原 玮	010-88005332 15910551936 yuanyi@guosen.com.cn		孔华强	021-60875170 13681669123 konghq@guosen.com.cn		徐 冉	0755-82130655 13632580795 xuran1@guosen.com.cn	
						严小燕	13590436977 yanxy@guosen.com.cn	
						赵晓曦	15999667170 zhaoxxi@guosen.com.cn	
						温 馨	13612901872 wenxin@guosen.com.cn	