多因子量化选股系列专题研究

因子定价逻辑与多因子跟踪体系的构 建

2018年2月8日

投资要点

- ❖ 投资聚焦:多因子模型是量化投资领域中最为经典、应用最为广泛的方法体系之一。本文使用改进后的截面回归法,在因子定价逻辑基础上对技术类、预期类、财务类因子的历史收益进行了考察,构建了多因子跟踪体系。
- ❖ 使用逐步回归法提取各类因子的纯收益。(1)截面回归在使用中面临的最大问题是因子之间的共线性问题,本文采用逐步回归法进行规避。(2)逐步回归法的含义是,每一步回归中只对单一变量进行回归,即仅估计单一因子的收益率,而每一步回归时将前一步回归的残差项作为这一步回归的因变量。(3)逐步回归法在实践中100%的规避了共线性问题,而且从逻辑上说,相当于在每步回归时将残差中和单因子有关的信息全部剔除,而后续因子的收益率可认为是已剔除了前序因子影响的纯收益。(4)我们以主观逻辑认识设定因子的考察顺序,并认为不需要对因子的顺序进行过多的筛选优化。当因子的顺序调换后,因子的含义已经发生了本质性改变,即原始考察中按照哪一个顺序来进行,那么在后续的研究中也遵从该顺序即可。
- ❖ 使用等权加权的复合方式构建因子可以提升因子的总体解释度。(1) 对于概念相似、数量较多、且暴露度相关性较高的部分因子来说,我们建议对因子进行复合,以复合后的综合因子进行最终的组合构建。(2) 如果暂不考虑权重优化问题的话,选择等权加权即可。(3) 一般来说,复合因子对总波动的解释度要高于使用单一因子。
- ❖ 对单一因子进行逐步回归、单步回归的对比考察。(1)在逐步、单步考察中,我们通过对比同一因子两种不同方式得到的收益率序列的差异,以及对总波动解释度的差异来判断其与前序因子共线性的高低。(2)全部因子中任意两因子间均具有一定的相关性,而逐步回归方法相当于将属于前一个因子的全部可解释收益归于该因子,后一个因子则用于解释剔除了前一个因子的剩余部分。但如果某一因子的剩余部分占该因子原本总解释度较低的话,则说明该因子与前序因子的共线性较高,应予剔除。

已保留各因子在不同空间的日均解释度(R2)

	沪深 300 空间	中证 500 空间	中证 800 空间
市场因子	25.36%	25.73%	24.93%
行业因子	20.96%	9.91%	11.46%
规模	0.40%	0.36%	0.47%
流动性	0.40%	0.53%	0.40%
流动性动量	0.31%	0.24%	0.23%
动量	0.56%	0.42%	0.50%
一致预期 E2P	0.17%	0.19%	0.14%
一致预期 PEG	0.18%	0.12%	0.09%
E2P	0.07%	0.11%	0.06%
ROE	0.20%	0.10%	0.11%
销售毛利率	0.13%	0.11%	0.07%
净利润增长率	0.14%	0.11%	0.07%
总收入增长率	0.15%	0.11%	0.08%

资料来源:中信证券数量化投资分析系统



中信证券研究部

王兆宇

电话: 021-20262110

邮件: wangzhaoyu@citics.com 执业证书编号: S1010514080008

赵文荣

电话: 010-60836759

邮件: zhaowenrong@citics.com 执业证书编号: S1010512070002

李祖苑

电话: 010-60838026 邮件: lizuyuan@citics.com 执业证书编号: S1010514070002

张依文

电话: 021-20262149 邮件: yiwenzhang@citics.com 执业证书编号: S101051708000

相关研究

- 1. 多因子量化选股系列专题研究一多因 子模型构建中的组合优化与风险预算 研究.......2017-11-17
- 2. 另类交易策略研究—通道突破股票的 市场表现与交易策略......2017-04-19
- 3. 量化选股系列专题一基于价值轮动的 上证 50 增强策略....... 2016-09-05



目录

投资聚焦	1
多因子模型基本理论框架	1
多因子的基本假设与表现形式	1
基于波动性的组合优化	2
基于因子类别的多因子模型分类	2
使用逐步回归法提取各类因子的纯收益	3
样本内考察的其他设定	3
市场与行业因子的处理	4
市场因子的收益表现	4
行业因子的收益表现	5
市场与行业因子对价格波动的解释度	6
技术类因子的考察	7
使用等权加权的复合方式构建因子可以提升因子的总体解释度	7
对单一因子进行逐步回归、单步回归的对比考察	8
一致预期类因子的考察	10
财务类因子的考察	13
多因子跟踪框架体系的最终构成	16



插图目录

图 1:	股价叠加各类信息的主要过程示意	3
图 2:	逐步回归的主要顺序示意	3
图 3:	市场因子收益率一沪深 300 空间	5
图 4:	市场因子收益率一中证 500 空间	5
图 5:	市场因子收益率一中证 800 空间	5
图 6:	市场因子收益率一中证 1000 空间	5
	不同空间的行业因子表现——钢铁行业	
图 8:	不同空间的行业因子表现——计算机行业	6
图 9:	行业因子对行业内波动率解释度(按中证 800 空间降虚排序)	7
图 10:	规模复合因子历史累计收益	7
图 11:	自由流通市值因子累计收益	7
图 12:	流通市值因子累计收益	8
	流动性因子的逐步回归累计收益	
图 14:	流动性因子的单步回归累计收益	9
	流动性动量因子的逐步回归累计收益	
图 16:	流动性动量因子的单步回归累计收益	9
	动量因子的逐步回归累计收益	
	动量因子的单步回归累计收益	
图 19:	波动率因子的逐步回归累计收益1	0
图 20:	波动率因子的单步回归累计收益1	0
图 21:	一致预期 E2P 因子的逐步回归累计收益1	1
	一致预期 E2P 因子的单步回归累计收益1	
	一致预期 PEG 因子的逐步回归累计收益1	
图 24:	一致预期 PEG 因子的单步回归累计收益1	1
	一致预期 ROE 因子的逐步回归累计收益1	
	一致预期 ROE 因子的单步回归累计收益1	
	一致预期 CF2P 因子的逐步回归累计收益1	
图 28:	一致预期 CF2P 因子的单步回归累计收益1	2
	一致预期 B2P 因子的逐步回归累计收益1	
	一致预期 B2P 因子的单步回归累计收益1	
图 31:	一致预期 D2P 因子的逐步回归累计收益 1	2
	一致预期 D2P 因子的单步回归累计收益 1	
图 33:	E2P 因子的逐步回归累计收益1	3
	E2P 因子的单步回归累计收益1	
	B2P 因子的逐步回归累计收益1	
	B2P 因子的单步回归累计收益1	
	S2P 因子的逐步回归累计收益1	
	S2P 因子的单步回归累计收益1	
图 39:	ROE 因子的逐步回归累计收益1	4



图 40:	ROE 因子的单步回归累计收益	14
图 41:	ROA 因子的逐步回归累计收益	15
图 42:	ROA 因子的单步回归累计收益	15
图 43:	销售净利率因子的逐步回归累计收益	15
图 44:	销售净利率因子的单步回归累计收益	15
图 45:	销售毛利率因子的逐步回归累计收益	15
图 46:	销售毛利率因子的单步回归累计收益	15
图 47:	净利润增长率因子的逐步回归累计收益	16
图 48:	净利润增长率因子的单步回归累计收益	16
图 49:	总收入增长率因子的逐步回归累计收益	16
图 50:	总收入增长率因子的单步回归累计收益	16
图 51:	单因子基本架构	17
	計目录 基于因子类别的多因子模型分类	2
	举」囚」关刑的多囚」模型分关 市场与行业因子日均对总波动的解释度	
	规模、自由流通市值、流通市值因子的日均解释度	
	技术类复合因子的构建方式	
	技术类因子的逐步回归日均解释度	
	技术类因子的单步回归日均解释度	
	备选一致预期因子的定义	
	一致预期类因子的逐步回归日均解释度	
	一致预期类因子的单步回归日均解释度	
	备选静态财务类因子的定义	
•	财务类因子的逐步回归日均解释度	
	财务类因子的单步回归日均解释度	
表 13:	已保留各因子在不同空间的日均解释度	



投资聚焦

多因子模型是量化投资领域中最为经典、应用最为广泛的方法体系之一。**我们在本文中使用改进后的截面回归法,在因子定价逻辑基础上对常规的技术类、预期类、财务类因子的历史收益进行了测算考察,构建了多因子跟踪体系。**

本文对截面回归法的主要改进在于使用逐步回归法来对每个因子的收益率进行估计。逐步回归法的含义是,每一步回归中只对单一变量进行回归,即仅估计单一因子的收益率,而每一步回归时将前一步回归的残差项作为这一步回归的因变量。这样的方法在实践中 100%的规避了共线性问题,而且从逻辑上说,相当于在每步回归时将残差中和单因子有关的信息全部剔除,而后续因子的收益率可认为是已剔除了前序因子影响的纯收益。

在逐步回归处理中,除了分析各因子历史收益的表现外,对共线性的处理也是关键一步,本文通过逐步回归、单步回归的对比分析,剔除在因子考察中与已确认因子具有高共线性的因子。全部因子中任意两因子间均具有一定的相关性,而逐步回归方法相当于将属于前一个因子的全部可解释收益归于该因子,后一个因子则用于解释剔除了前一个因子的剩余部分。但如果某一因子的剩余部分占该因子原本总解释度较低的话,则说明该因子与前序因子的共线性较高,应予剔除。与前序因子具有高共线性因子的特征是,在逐步回归下的解释度要比单步回归低很多,同时逐步回归与单步回归的历史收益率序列相差较大。

最终,在使用改进后的截面回归法对常规的技术类、预期类、财务类的历史收益进行测算的基础上,我们完成构建多因子体系跟踪体系的构建。。

多因子模型基本理论框架

多因子的基本假设与表现形式

多因子模型的主要形式可由公式(1)来表示:

$$r_i = x_{i,0} + x_{i,1}f_1 + \dots + x_{i,K}f_K + \epsilon_i$$
 (1)

其中, $\{x_{ik}\}_{k=0}^K$ 为股票 i 的因子暴露(Factor Exposure); $\{f_k\}_{k=1}^K$ 为因子收益率(Factor Return); ϵ_i 为资产 i 的特异风险(Specific Risk),且通常假定其在组合构建中可被充分分散而相互抵消。

公式(1)本质上只是一个示意性的表达法,基于不同的数理处理方法可构成多种不同的多因子模型,本文我们以截面回归(Cross-Sectional Regression)体系构建多因子模型。那么为了实现回归,对公式(1)中的个股收益率还需假定满足正态分布。设收益率向量 $r=(r_1,r_2,...,r_N)'$,且 $r\sim N(\mu,\Sigma)$,个股权重向量 $w=(w_1,w_2,...,w_N)'$,由该权重构成的组合 $r_n=w'r$,则有:

$$r_p \sim N(w'\mu, w'\Sigma w)$$
 (2)

基于公式(1)还可对股票组合的风险进行分析,即构建组合的风险模型。一般来说,计算组合风险时首先必须要得到个股的协方差矩阵,而由于组合中的个股数量可能很多,因此仅以个股来分析的话,要求协方差矩阵可能达到上百阶,这将给数据处理造成极大的压力。通过公式(1),则可以将对个股协方差矩阵的求解转化为对因子收益率序列协方差矩阵的求



解,大大简化处理过程。设风险暴露矩阵 $X=\begin{pmatrix}x_{1,1}&\cdots&x_{1,K}\\ \vdots&\ddots&\vdots\\ x_{N,1}&\cdots&x_{N,K}\end{pmatrix}$,因子收益协方差矩阵 $\Sigma_f=\begin{pmatrix}\sigma_{1,1}&\cdots&\sigma_{1,K}\\ \vdots&\ddots&\vdots\\ \sigma_{K,1}&\cdots&\sigma_{K,K}\end{pmatrix}$,特异风险协方差对角矩 $S=diag(\theta_1^2,\theta_2^2,\ldots,\theta_N^2)$,则股票协方差矩阵 Σ 可写为。

$$\Sigma = X\Sigma_f X' + S (3)$$

基于波动性的组合优化

组合优化是求解满足自身投资目标的最优组合的过程。在这个过程中,首先要设定投资目标,特别是确定需要暴露哪些风险,对冲哪些风险;然后将上述问题量化为具体的数学表达,即确定目标函数和边界条件,最后进行最优效用函数的求解。

在确定投资目标时,首先需要确定如何定义风险。通常来说会以方差(Variance)或者在险价值(Value at Risk)定义风险。如以方差来定义,则组合的总风险 σ^2 可写为:

$$\sigma^2 = w'\Sigma w = (w'X)\Sigma_f(X'w) + w'Sw \quad (4)$$

常见的优化目标函数包含以下几种:

- Minimum Variance: Min $\frac{1}{2}w'\Sigma w$
- Mean-Variance Optimization: Max w'r $-\frac{1}{2}\lambda(w'\Sigma w)$
- Active Mean-Variance Optimization: $\max a'r \frac{1}{2}\lambda(a'\Sigma a)$, a = w b 常见的优化边界条件包括:
- ▶ 仓位约束:例如w'1=0
- **)** 边际风险贡献(Marginal Contribution to Risk): 个股MCR = $\frac{\partial \sigma}{\partial w} = \frac{x \Sigma_f X' w + S w}{\sigma}$; 组合 MCR_t = t' MCR, t = $(t_1, t_2, ..., t_N)'$
- ightharpoons 风险贡献(Contribution to Risk): $CR = w \otimes \frac{\partial \sigma}{\partial w} = \frac{w \otimes (x \sum_f X' w + S w)}{\sigma}$

基于因子类别的多因子模型分类

不同的多因子模型除了所使用的方法体系有所区别外,在使用的因子上可能也有所不同。 基于因子类别来分类,多因子模型可分为宏观多因子模型、基本面多因子模型、统计多因子 模型。表 1 展示了上述三种模型的主要异同。从学术界和业界的主要接受度来看,基本面多 因子模型较受推崇,本文也采用了基本面多因子体系。

表 1: 基于因子类别的多因子模型分类

K 11 E 1 L 1 (
	宏观因子模型	基本面因子模型	统计因子模型			
估计方法	将宏观指标作为因子值,通过时间 序列回归确定个股对每个指标的 敏感度(因子暴露)。	基于基本面数据计算个股的因子 暴露,然后通过界面回归获得对应 期限的因子收益率。	将各因子值和敏感度均视作未知, 对股票历史协方差矩阵进行统计 分析(例如主成分分析),估算因 子值和敏感度。			
常见因子举例	市场收益率、短期利率 、行业产能(及变动)、通胀、长 短期利差、信用利差等	行业、规模、B2P,分红率、动量、成长、价值、波动、杠杆等	所的因子未必会有明确的解释			
稳健性 (样本外的预测性)	较好	较好	较差(导致伪相关问题)			
股票波动的解释度	较差	较好	较好			
动态性(对风险暴露变化进行调整	较慢	较快(当某一因子暴露变化时,对	较慢			



资料来源: Aamir Sheikh (1996),中信证券研究部

使用逐步回归法提取各类因子的纯收益

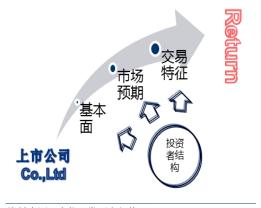
截面回归在使用中面临的最大问题是因子之间的共线性问题,且备选因子越多共线性可能越严重,这会导致回归的结果(因子收益率)很不稳定,对后续处理产生极大障碍。本文采用逐步回归法来应对共线性问题。

逐步回归法的含义是,每一步回归中只对单一变量进行回归,即仅估计单一因子的收益率,而每一步回归时将前一步回归的残差项作为这一步回归的因变量。这样的方法在实践中100%的规避了共线性问题,而且从逻辑上说,相当于在每步回归时将残差中和单因子有关的信息全部剔除,而后续因子的收益率可认为是已剔除了前序因子影响的纯收益。

在备选因子的组成上,首先我们遵从主流的做法,将因子分为市场因子、行业因子、风格因子三大类。其次,在构成最为复杂的风格因子中,我们按照技术类因子→预期类因子→静态财务类因子的顺序进行处理。这是因为我们认为股价对信息的叠加按照基本面信息→市场预期信息→交易特征信息这样的顺序(如图 1)来进行,那么对信息的剥离就需要以相反的顺序来进行。最终各类因子的处理过程如图 2 所示。

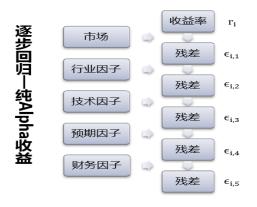
我们对图 2 所示处理顺序的设定完全基于主观逻辑认识,我们认为而不需要对因子的前后顺序进行过多的筛选优化。本质上来说,当因子的前后顺序调换后,因子的含义已经发生了本质性改变,不同回归顺序下的因子收益率序列之间不具备可比性。因此我们认为在原始考察中按照哪一个顺序来进行,那么在后续的研究中遵从该顺序即可。

图 1: 股价叠加各类信息的主要过程示意



资料来源:中信证券研究部整理

图 2: 逐步回归的主要顺序示意



资料来源:中信证券研究部整理

样本内考察的其他设定

考察期: 2009年1月1日至2017年3月31日。

回归频率:每日。

回归方法: 简单最小二乘法(OLS)。

行业分类法:中信一级行业(以实际调整日为准)。



因子标准化过程:

- ▶ 步骤 1:对全部股票按照因子值降序排序,使用 5%分位点至 95%分位点的股票, 计算加权均值和等权标准差,用该均值和标准差对全部股票标准化。
 - ▶ 步骤 2: 对步骤 1 的因子值进行缩尾, ±3 以外的因子值全部设为+/-3。
- ▶ 步骤 3: 对步骤 2 的结果再以全部样本计算加权均值和等权标准差,并以此均值和标准差进行标准化。

全样本空间(全 A 空间)的定义:

- ▶ 全部上市超过 90 个交易日的 A 股;
- ▶ 剔除 ST 股票 (以实际调整日为准);
- ▶ 剔除净利润(TTM)为负的股票。

对样本空间进一步分组:

- ▶ 全A空间:即全市场;
- 沪深 300 空间:沪深 300 成分股;
- ▶ 中证 500 空间: 中证 500 成分股;
- ▶ 中证 800 空间: 中证 800 成分股;
- ▶ 中证 1000 空间: 中证 1000 成分股

市场与行业因子的处理

市场因子的收益表现

根据上文的框架,首先需要处理市场因子,所使用的回归公式如(5)所示:

$$\mathbf{r}_{\mathrm{i},\mathrm{t}} = m_t + \epsilon_{i,t} \ (5)$$

其中, m_t 为第 t 日针对某一样本空间的市场因子收益率, $\epsilon_{i,t}$ 为股票 i 的残差项。市场因子即是回归方程的常数项,对于公式(5)这样一个只有常数项的回归方程来说, m_t 等同于考察日全部样本的算数平均收益率。

图 3 至图 6 分别展示了针对沪深 300、中证 500、中证 800、中证 1000 四个空间市场因子累计收益率的走势,可见由于近年来小盘股效应普遍较强,因此四个空间市场因子表现远远好于同一空间以自由流通市值加权的规模指数的表现。需要注意的是,这里的市场因子是一个每日进行再平衡同时不计交易成本的理论组合,仅用于表征而不具可投资性。



图 3: 市场因子收益率一沪深 300 空间



图 5: 市场因子收益率—中证 800 空间



资料来源:中信证券数量化投资分析系统

图 4: 市场因子收益率一中证 500 空间



资料来源:中信证券数量化投资分析系统

图 6: 市场因子收益率—中证 1000 空间



资料来源:中信证券数量化投资分析系统

行业因子的收益表现

经过公式(5)的回归残差项可认为是在市场中性下的超额收益率,我们将其记为 $\epsilon_{i,t}$,并作为处理行业因子的回归方程的被解释变量,如公式(6)所示:

$$\hat{\epsilon}_{i,t} = \sum_{j=1}^{29} I_{i,j,t} Industry_{j,t} + \epsilon_{i,t}$$
 (6)

其中, $I_{i,j,t}$ 为股票 i 在第 t 日是否属于行业 j 的指示变量,本文采用了中信一级行业分类,行业总数为 29 且行业之间互斥; $Industry_{j,t}$ 为行业 j 在第 t 日的因子收益率,即为市场中性下该行业的超额收益率; $\epsilon_{i,t}$ 为回归残差项,记为股票 i 在市场中性、行业中性下的超额收益。

不同行业在不同空间的表现有所差异,图 7 和图 8 分别列举了钢铁和计算机两行业在三个空间中的累计收益率。



图 7: 不同空间的行业因子表现——钢铁行业



资料来源:中信证券数量化投资分析系统

图 8: 不同空间的行业因子表现——计算机行业



资料来源:中信证券数量化投资分析系统

市场与行业因子对价格波动的解释度

回归模型的R²统计量用于描述解释变量对被解释变量总波动的解释度,一般的计算公式为:

$$R^2 = \frac{\overline{W}}{\overline{W}} = \frac{\overline{W}}{\overline{W}} = \overline{W} = \overline{W}$$
 \overline{W} \overline{W}

对于本文所采用的的逐步回归,可得第1步回归的解释度为:

其中,对于第一步进行的市场因子回归来说,(I-1步残差平方和)=收益率平方和。

表 2 展示了不同空间的市场与行业因子对总波动的日均解释度,可见市场因子的解释度在不同空间中均为 25%左右,但行业因子解释度的差异较大。对于大盘空间(如沪深 300空间),行业因子解释度达 20%以上,说明大盘空间同一行业的股票趋同度较高;但随着空间所覆盖市值的减小,同一行业股票分化也逐渐加大,中证 1000 空间的行业因子解释度不足 5%。

表 2: 市场与行业因子日均对总波动的解释度

	沪深 300 空间	中证 500 空间	中证 800 空间	中证 1000 空间
市场因子	25.36%	25.73%	24.93%	24.21%
行业因子	20.96%	9.91%	11.46%	4.85%

资料来源:中信证券数量化投资分析系统

在行业因子的考察上,我们也可根据公式(7)定义行业因子对行业内波动率的解释度。 由图 9 可见,行业因子对行业内波动率的解释度与行业的同质性有较大关系。对于银行、煤炭、非银、钢铁等同质化较高的行业,行业因子对行业内总波动解释度均较高;但对于基础化工、商贸零售、机械等行业内同质化较低的行业,行业因子的解释度较低。





图 9: 行业因子对行业内波动率解释度(按中证 800 空间降虚排序)

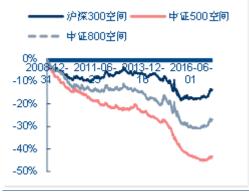
技术类因子的考察

使用等权加权的复合方式构建因子可以提升因子的总体解释度

对于概念相似、数量较多、且暴露度的相关性较高的部分因子来说,我们建议对因子进 行复合,以复合后的综合因子进行最终的组合构建。如果暂不考虑权重优化问题的话,选择 等权加权即可。

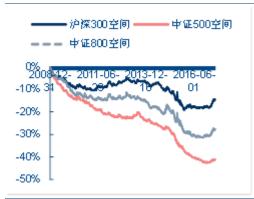
以规模因子为例,设规模复合因子=自由流通市值*50%+流通市值*50%。其中在加权过 程中,首先要对单一因子按照上文的标准化处理方法进行处理,然后对处理后的指标进行加 权,得到复合因子值后还要再采用标准化的方法处理,最终才得到回归方程使用的因子。以 上文中得到的市场中性、行业中性超额收益作为本步处理的被解释变量,以规模复合因子、 以及原始自由流通市值因子和流通市值因子分别作为解释变量进行回归,得到历史累计收益 率序列如图 10 至图 12 所示。可见用这三个因子得到的因子收益率走势非常相近,但是复合 因子对总波动的解释度要高于使用单一因子(见表3),因此我们建议使用复合因子。





资料来源:中信证券数量化投资分析系统

图 11: 自由流通市值因子累计收益



资料来源:中信证券数量化投资分析系统





图 12: 流通市值因子累计收益

表 3: 规模、自由流通市值、流通市值因子的日均解释度

	规模因子	自由流通市值因子	流通市值因子
沪深 300 空间	0.40%	0.36%	0.40%
中证 500 空间	0.36%	0.34%	0.34%
中证 800 空间	0.47%	0.45%	0.47%

资料来源:中信证券数量化投资分析系统

除了规模因子以外,其他几个技术类因子也需要进行复合处理,构建方式如表 4 所示。

表 4: 技术类复合因子的构建方式

因子	构建方式
规模因子	自由流通市值*50%+流通市值*50%
流动性因子	过去 20 日日均成交额*50%+过去 20 日日均换手率*50%
流动性动量因子	20 日/60 日日均成交额*50%+20 日/60 日日均换手率*50%
动量因子	20 日收益率*1/3+60 日收益率*1/3+120 日收益率*1/3
波动率复合因子	20 日波动率*1/3+60 日波动率*1/3+120 日波动率*1/3

资料来源:中信证券研究部

对单一因子进行逐步回归、单步回归的对比考察

由于技术类中的很多因子与规模因子具有相关性,因此本文在处理技术类因子时,首先处理规模因子,再以规模因子的残差处理其他的技术类因子,这样其他因子的 alpha 收益可视作剔除了规模因子影响的 alpha 收益。

接下来进行对流动性、流动性动量、动量、波动率因子的考察。但在这一步中,我们需要进行一个对比考察:

- ▶ 考察(1): 逐步回归。以技术类因子的残差作为被解释变量,按照流动性、流动性动量、动量、波动率的顺序进行逐步回归,每步回归时都以前一步的残差作为被解释变量。
- ▶ 考察(2): 单步回归: 以技术类因子的残差作为被解释变量,以上述四个因子分别作为 该残差项的解释变量进行回归,得到各自的回归结果。

在逐步、单步考察中,我们通过对比同一因子两种不同方式得到的收益率序列的差异,以及对总波动解释度的差异来判断其与前序因子共线性的高低。全部因子中任意两因子间均具有一定的相关性,而逐步回归方法相当于将属于前一个因子的全部可解释收益归于该因子,后一个因子则用于解释剔除了前一个因子的剩余部分。但如果某一因子的剩余部分占该因子原本总解释度较低的话,则说明该因子与前序因子的共线性较高,应予剔除。

从表 5 和表 6 的结果来看,波动率因子进行逐步回归的解释度要比单步回归低很多,而其他三个因子相差较小,这说明波动率因子对总波动的解释中有很大一部分已经被前序的因



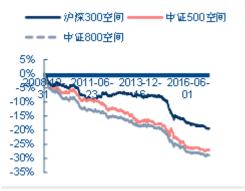
子所代替,而其他三个因子被各自前序因子代替的比例则较小。由图 13 至图 18 可见,流动性、流动性动量、动量三个因子的两种方式的考察结果非常类似,但由图 19 和图 20 所示的波动率因子的两种方式结果相差较多,这也是由于波动率因子高共线性导致伪回归的一种表现。综上我们认为应保留流动性、流动性动量、动量三个因子,剔除波动率因子。

图 13: 流动性因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 15: 流动性动量因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 17: 动量因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 14: 流动性因子的单步回归累计收益



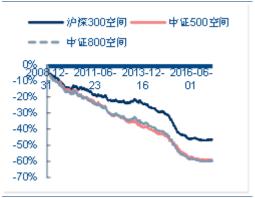
资料来源:中信证券数量化投资分析系统

图 16: 流动性动量因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 18: 动量因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统



图 19: 波动率因子的逐步回归累计收益



图 20: 波动率因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

表 5: 技术类因子的逐步回归日均解释度

	规模	流动性	流动性动量	动量	波动率
沪深 300 空间	0.40%	0.40%	0.31%	0.56%	0.28%
中证 500 空间	0.36%	0.53%	0.24%	0.42%	0.23%
中证 800 空间	0.47%	0.40%	0.23%	0.50%	0.25%

资料来源:中信证券数量化投资分析系统

表 6: 技术类因子的单步回归日均解释度

	规模	流动性	流动性动量	动量	波动率
沪深 300 空间	0.40%	0.40%	0.41%	0.79%	0.58%
中证 500 空间	0.36%	0.53%	0.37%	0.68%	0.61%
中证 800 空间	0.47%	0.40%	0.34%	0.77%	0.59%

资料来源:中信证券数量化投资分析系统

一致预期类因子的考察

一致预期类因子中,我们选取的是 Wind 一致预期数据,并采用经 Wind 调整的滚动未来一年的数据进行测算。下表展示了备选一致预期因子的具体定义。

表 7: 备选一致预期因子的定义

A Parameter				
因子名称	定义			
一致预期 E2P	市盈率的倒数			
一致预期 PEG	PEG			
一致预期 ROE	净资产收益率			
一致预期 CF2P	预期每股现金流除以当前股价			
一致预期 B2P	市净率的倒数			
一致预期 D2P	预期每股分红除以当前股价			

资料来源:中信证券研究部

按照与上文相同的标准进行考察,发现一致预期 B2P、D2P 两因子在逐步回归和单步回归下的日均解释度相差过大,且历史因子收益率序列的表现差异也较大,因此剔除;一致预期 ROE、CF2P 虽然在单步回归和逐步回归下表现差异不大,但因子收益率表现不尽如人意,因此也予以剔除;最终我们保留了一致预期 E2P、PEG 两因子,两者与前序因子的共线性不大,且历史因子收益率也较为显著。



图 21: 一致预期 E2P 因子的逐步回归累计收益



图 23: 一致预期 PEG 因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 25:一致预期 ROE 因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 22: 一致预期 E2P 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 24: 一致预期 PEG 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 26:一致预期 ROE 因子的单步回归累计收益





图 27: 一致预期 CF2P 因子的逐步回归累计收益

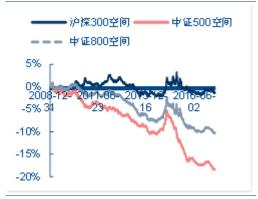
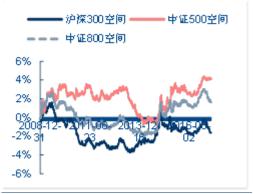
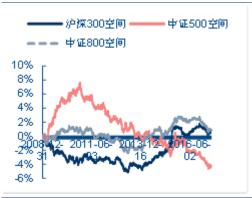


图 29: 一致预期 B2P 因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 31: 一致预期 D2P 因子的逐步回归累计收益



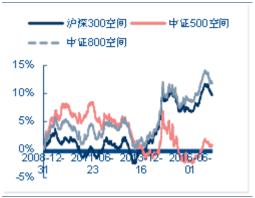
资料来源:中信证券数量化投资分析系统

图 28: 一致预期 CF2P 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 30: 一致预期 B2P 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 32: 一致预期 D2P 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

表 8: 一致预期类因子的逐步回归日均解释度

	一致预期 E2P	一致预期 PEG	一致预期 ROE	一致预期 CF2P	一致预期 B2P	一致预期 D2P
沪深 300 空间	0.17%	0.18%	0.24%	0.15%	0.07%	0.11%
中证 500 空间	0.19%	0.12%	0.15%	0.11%	0.07%	0.09%
中证 800 空间	0.14%	0.09%	0.14%	0.08%	0.05%	0.06%



表 9: 一致预期类因子的单步回归日均解释度

	一致预期 E2P	一致预期 PEG	一致预期 ROE	一致预期 CF2P	一致预期 B2P	一致预期 D2P
沪深 300 空间	0.17%	0.20%	0.29%	0.17%	0.21%	0.16%
中证 500 空间	0.19%	0.14%	0.19%	0.12%	0.17%	0.14%
中证 800 空间	0.14%	0.10%	0.18%	0.10%	0.16%	0.10%

财务类因子的考察

类似的,按照上文的考察过程,我们对以下备选财务因子也进行了考察。下表展示了我们备选因子的定义。

表 10: 备选静态财务类因子的定义

因子名称	定义
E2P	TTM 市盈率的倒数
B2P	MRQ 市净率的倒数
S2P	TTM 市销率的倒数
ROE	TTM 净资产收益率
ROA	TTM 总资产收益率
销售净利率	TTM 净利润/TTM 营业总收入
销售毛利率	TTM 营业成本/TTM 营业收入
净利润增长率	TTM 净利润/TTM 净利润_Lag-1

资料来源:中信证券研究部

由测算结果可见,B2P、S2P、ROA、销售净利率这四个因子的历史收益表现较差,因此不作为备选因子;而 E2P、ROE、销售毛利率、净利润增长率、总收入增长率这 4 个因子因子的历史超额收益相对较好,且在单步回归和逐步回归下的差异也不大,因此将其作为备选因子。

图 33: E2P 因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 34: E2P 因子的单步回归累计收益





图 35: B2P 因子的逐步回归累计收益

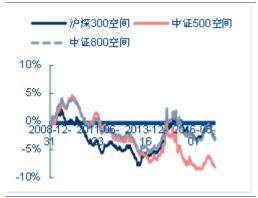
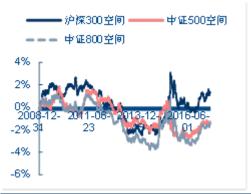
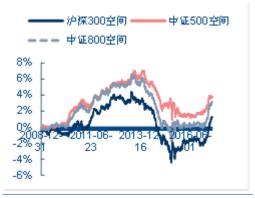


图 37: S2P 因子的逐步回归累计收益



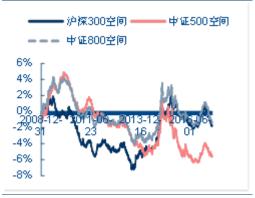
资料来源:中信证券数量化投资分析系统

图 39: ROE 因子的逐步回归累计收益



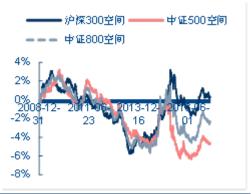
资料来源:中信证券数量化投资分析系统

图 36: B2P 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 38: S2P 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 40: ROE 因子的单步回归累计收益

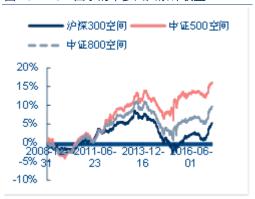




图 41: ROA 因子的逐步回归累计收益

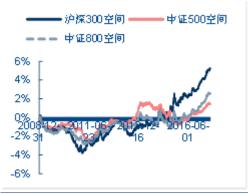
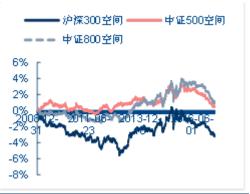


图 43: 销售净利率因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 45: 销售毛利率因子的逐步回归累计收益



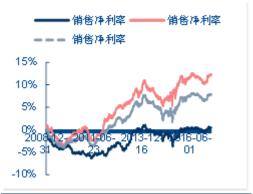
资料来源:中信证券数量化投资分析系统

图 42: ROA 因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 44: 销售净利率因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

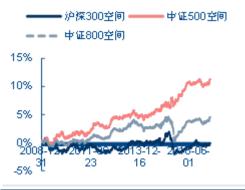
图 46: 销售毛利率因子的单步回归累计收益



图 47: 净利润增长率因子的逐步回归累计收益



图 49: 总收入增长率因子的逐步回归累计收益



资料来源:中信证券数量化投资分析系统

图 48: 净利润增长率因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

图 50: 总收入增长率因子的单步回归累计收益



资料来源:中信证券数量化投资分析系统

表 11: 财务类因子的逐步回归日均解释度

	E2P	B2P	S2P	ROE	ROA	销售净利率	销售毛利率	净利润 增长率	总收入增 长率
沪深 300 空间	0.07%	0.18%	0.12%	0.13%	0.07%	0.13%	0.08%	0.14%	0.15%
中证500空间	0.11%	0.16%	0.09%	0.06%	0.03%	0.05%	0.06%	0.12%	0.11%
中证800空间	0.06%	0.14%	0.07%	0.07%	0.03%	0.07%	0.05%	0.08%	0.07%

资料来源:中信证券数量化投资分析系统

表 12: 财务类因子的单步回归日均解释度

	E2P	B2P	S2P	ROE	ROA	销售净利率	销售毛利率	净利润 増长率	总收入增 长率
沪深 300 空间	0.07%	0.16%	0.18%	0.25%	0.27%	0.21%	0.17%	0.18%	0.21%
中证500空间	0.11%	0.16%	0.13%	0.20%	0.23%	0.19%	0.16%	0.14%	0.15%
中证800空间	0.06%	0.13%	0.12%	0.18%	0.21%	0.14%	0.12%	0.10%	0.11%

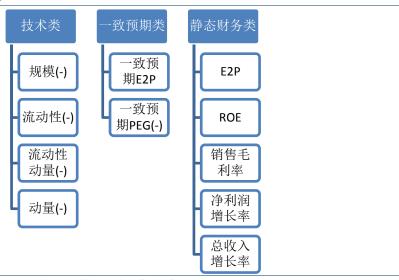
资料来源:中信证券数量化投资分析系统

多因子跟踪框架体系的最终构成

最终,我们筛选出来的多因子框架的构成,以及各因子的日均解释度分别如图 51 和表 13 所示。关于这些因子的具体使用和组合优化等相关研究,请参阅《量化选股系列专题研究——多因子模型构建中的组合优化与风险预算研究》。



图 51: 单因子基本架构



资料来源:中信证券研究部。注:"-"表示因子值越小正超额收益越高

表 13: 已保留各因子在不同空间的日均解释度

	沪深 300 空间	中证 500 空间	中证 800 空间
市场因子	25.36%	25.73%	24.93%
行业因子	20.96%	9.91%	11.46%
规模	0.40%	0.36%	0.47%
流动性	0.40%	0.53%	0.40%
流动性动量	0.31%	0.24%	0.23%
动量	0.56%	0.42%	0.50%
一致预期 E2P	0.17%	0.19%	0.14%
一致预期 PEG	0.18%	0.12%	0.09%
E2P	0.07%	0.11%	0.06%
ROE	0.20%	0.10%	0.11%
销售毛利率	0.13%	0.11%	0.07%
净利润增长率	0.14%	0.11%	0.07%
总收入增长率	0.15%	0.11%	0.08%



分析师声明

主要负责撰写本研究报告全部或部分内容的分析师在此声明: (i) 本研究报告所表述的任何观点均精准地反映了上述每位分析师个人对标的证券和发行人的看法; (ii) 该分析师所得报酬的任何组成部分无论是在过去、现在及将来均不会直接或间接地与研究报告所表述的具体建议或观点相联系。

评级说明

投资建议的评级标准 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 6 到 12 个月内的相对市场表现,也即:以报告发布日后的 6 到 12 个月内的公司股价(或行业指数)相对同期相关证券 市场代表性指数的涨跌幅作为基准。其中:A股市场以沪 深 300 指数为基准,新三板市场以三板成指(针对协议转 让标的)或三板做市指数(针对做市转让标的)为基准; 香港市场以摩根士丹利中国指数为基准;美国市场以纳斯 达克综合指数或标普 500 指数为基准。

	评级	说明
	买入	相对同期相关证券市场代表性指数涨幅 20%以上;
DV. 1885 2575 647.	增持	相对同期相关证券市场代表性指数涨幅介于5%~20%之间
股票评级	持有	相对同期相关证券市场代表性指数涨幅介于-10%~5%之间
	卖出	相对同期相关证券市场代表性指数跌幅 10%以上;
	强于大市	相对同期相关证券市场代表性指数涨幅 10%以上;
行业评级	中性	相对同期相关证券市场代表性指数涨幅介于-10%~10%之间;
	弱于大市	相对同期相关证券市场代表性指数跌幅 10%以上

其他声明

本研究报告由中信证券股份有限公司或其附属机构制作。中信证券股份有限公司及其全球的附属机构、分支机构及联营机构(仅就本研究报告免责条款而言,不含 CLSA group of companies),统称为"中信证券"。

法律主体声明

中国:本研究报告在中华人民共和国(香港、澳门、台湾除外)由中信证券股份有限公司(受中国证券监督管理委员会监管,经营证券业务许可证编号:**72**03**74**000)分发。

新加坡:本研究报告在新加坡由 CLSA Singapore Pte Ltd(公司注册编号: 198703750W)分发。作为资本市场经营许可持有人及受豁免的财务顾问,CLSA Singapore Pte Ltd 仅向新加坡《证券及期货法》s.4A(1)定义下的"机构投资者、认可投资者及专业投资者"提供证券服务。根据新加坡《财务顾问法》下《财务顾问(修正)规例(2005)》中关于机构投资者、认可投资者、专业投资者及海外投资者的第 33、34、35 及 36 条的规定,《财务顾问法》第 25、27 及 36 条不适用于 CLSA Singapore Pte Ltd。如对本报告存有疑问,还请联系 CLSA Singapore Pte Ltd(电话: +65 6416 7888)。MCI (P) 024 11 2017.

针对不同司法管辖区的声明

中国:根据中国证券监督管理委员会核发的经营证券业务许可,中信证券股份有限公司的经营范围包括证券投资咨询业务。

新加坡: 监管法规或交易规则要求对研究报告涉及的实际、潜在或预期的利益冲突进行必要的披露。须予披露的利益冲突可依照相关法律法规要求在特定报告中获得,详细内容请查看 https://www.clsa.com/disclosures.html。该等披露内容仅涵盖 CLSA group, CLSA Americas 及 CL Securities Taiwan Co., Ltd 的情况,不涉及中信证券及/或其附属机构的情况。如投资者浏览上述网址时遇到任何困难或需要过往日期的披露信息,请联系compliance_hk@clsa.com.

美国: 本研究报告由中信证券编制。本研究报告在美国由中信证券(CITIC Securities International USA, LLC(下称"CSI-USA")除外)和 CLSA group of companies(CLSA Americas, LLC(下称"CLSA Americas")除外)仅向符合美国《1934 年证券交易法》下 15a-6 规则定义且分别与 CSI-USA 和 CLSA Americas 进行交易的"主要美国机构投资者"分发。对身在美国的任何人士发送本研究报告将不被视为对本报告中所评论的证券进行交易的建议或 对本报告中所载任何观点的背书。任何从中信证券与 CLSA group of companies 获得本研究报告的接收者如果希望在美国交易本报告中提及的任何证券 应当分别联系 CSI-USA 和 CLSA Americas.

英国:本段"英国"声明受英国法律监管并依据英国法律解释。本研究报告在英国须被归为营销文件,它不按《英国金融行为管理手册》所界定、旨在提升投资研究报告独立性的法律要件而撰写,亦不受任何禁止在投资研究报告发布前进行交易的限制。本研究报告在欧盟由 CLSA (UK)发布,该公司由金融行为管理局授权并接受其管理。本研究报告针对《2000年金融服务和市场法 2005年(金融推介)令》第 19 条所界定的在投资方面具有专业经验的人士,且涉及到的任何投资活动仅针对此类人士。若您不具备投资的专业经验,请勿依赖本研究报告的内容。

一般性声明

本研究报告对于收件人而言属高度机密,只有收件人才能使用。本研究报告并非意图发送、发布给在当地法律或监管规则下不允许该研究报告发送、发布的人员。本研究报告仅为参考之用,在任何地区均不应被视为出售任何证券或金融工具的要约,或者证券或金融工具交易的要约邀请。中信证券并不因收件人收到本报告而视其为中信证券的客户。本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要,不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具的分析,本报告的收件人须保持自身的独立判断。

本报告所载资料的来源被认为是可靠的,但中信证券不保证其准确性或完整性。中信证券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有关的其他损失承担任何责任。本报告提及的任何证券均可能含有重大的风险,可能不易变卖以及不适用所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

本报告所载的资料、观点及预测均反映了中信证券在最初发布该报告日期当日分析师的判断,可以在不发出通知的情况下做出更改,亦可因使用不同假设和标准、采用不同观点和分析方法而与中信证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。中信证券并不承担提示本报告的收件人注意该等材料的责任。中信证券通过信息隔离墙控制中信证券内部一个或多个领域的信息向中信证券其他领域、单位、集团及其他附属机构的流动。负责撰写本报告的分析师的薪酬由研究部门管理层和中信证券高级管理层全权决定。分析师的薪酬不是基于中信证券投资银行收入而定,但是,分析师的薪酬可能与投行整体收入有关,其中包括投资银行、销售与交易业务。

若中信证券以外的金融机构发送本报告,则由该金融机构为此发送行为承担全部责任。该机构的客户应联系该机构以交易本报告中提及的证券或要求获悉更详细信息。本报告不构成中信证券向发送本报告金融机构之客户提供的投资建议,中信证券以及中信证券的各个高级职员、董事和员工亦不为(前述金融机构之客户)因使用本报告或报告载明的内容产生的直接或间接损失承担任何责任。

未经中信证券事先书面授权,任何人不得以任何目的复制、发送或销售本报告。

中信证券 2018 版权所有。保留一切权利。