

2017 年 12 月 28 日

证券研究报告—金融工程专题报告

# 基于“长记忆性”特征改进因子择时的研究

调研日期 2017 年 11 月 18 日  
分析日期 2017 年 12 月 18 日

证券分析师： 丁竞渊  
执业证书编号： S0630516080001  
电话： 021-20333723  
邮箱： djy@longone.com.cn

## ◎ 投资要点：

### ◆ 主要内容。

“长记忆性”是金融时间序列的普遍现象，Hurst 指数可以有效地表征金融时间序列的记忆性特征。本报告通过构建移动 Hurst 指数来捕捉金融时间序列非周期循环过程中的择时机会，尝试将其应用到单因子/多因子选股模型中，探索改进因子择时效果的新模式。

根据金融时间序列中移动 Hurst 指数值大多数时候较高（长记忆性），移动 Hurst 指数值处于低位后会迅速回升至高位的特点，我们结合 Hurst 指数与均线系统构建了 HS 调节因子模型。利用 HS 调节因子与其他因子加权组合为 HS 复合因子，以改进其他因子的选股效果。

针对 Hurst 指数与其他因子组合方式无先验函数关系的问题，我们采取分类决策树的形式构建了 Hurst 因子分类树模型。

验证结果表明 HS 调节因子模型对因子选股模型没有明显改进效果，而 Hurst 因子分类树模型对部分因子的选股效果有明显提升，但其效果的发挥受原始因子的效果，以及 Hurst 指数与原始因子结合方式的影响很大。利用机器学习技术寻找 Hurst 指数与原始因子相互关系的函数形式是我们未来的主要研究方向。。

### ◆ 主要结论。

1. 将 Hurst 指数应用到因子选股模型中，对改进因子选股的效果具有重要意义。
2. Hurst 指数对因子选股模型的改进效果，一方面取决于原始因子的有效性，另一方面取决于 Hurst 指数与原始因子的组合方式。在没有理论支持的先验函数形式的情况下，直接采用因子评分加权组合无法达到预期效果。机器学习将是未来寻找因子组合方式的重要途径。
3. Hurst 因子分类树模型对市值因子的选股效果有明显提升，最大回撤大幅下降，期望收益率/最大回撤比大幅提高，区分效果明显增强。

相关研究报告

## 正文目录

|                                 |    |
|---------------------------------|----|
| 1. 基于“长记忆性”特征的研究思路.....         | 4  |
| 2. Hurst 模型的本质与意义.....          | 6  |
| 2.1. R/S 分析方法.....              | 6  |
| 2.2. 移动 Hurst 指数的择时意义.....      | 8  |
| 3. 基于 Hurst 指数的因子择时模型.....      | 9  |
| 3.1. HS 调节因子模型.....             | 10 |
| 3.1.1. HS 调节因子的构建.....          | 10 |
| 3.1.2. HS 调节因子的评分.....          | 10 |
| 3.1.3. 基于 HS 调节因子的因子分组检验方法..... | 11 |
| 3.2. Hurst 因子分类树模型.....         | 11 |
| 4. 主要实证结果分析.....                | 12 |
| 4.1. 回测检验方案.....                | 13 |
| 4.2. 原始单因子组合实证结果分析.....         | 14 |
| 4.3. 原始多因子组合实证结果分析.....         | 16 |
| 4.4. HS 调节因子模型实证结果分析.....       | 17 |
| 4.5. Hurst 因子分类树模型实证结果分析.....   | 20 |
| 5. 结论与进一步研究.....                | 24 |

## 图表目录

|   |    |
|---|----|
| 图 1 (上证综指周收益率分布).....                     | 4  |
| 图 2 (沪深 300 周收益率分布).....                  | 4  |
| 图 3 (上证 50 周收益分布).....                    | 5  |
| 图 4 (中证 500 周收益分布).....                   | 5  |
| 图 5 (各主要指数收益率 HURST 指数 (2007-2017) )..... | 6  |
| 图 6 (上证综指 HURST 指数 $H=0.68$ (斜率) ).....   | 8  |
| 图 7 (上证综指及其移动 HURST 指数).....              | 9  |
| 图 8 (中证 500 及其移动 HURST 指数).....           | 9  |
| 图 9 (Hurst 因子分类树).....                    | 12 |
| 图 10 (单因子最优组合收益 (除市值因子外) ).....           | 14 |
| 图 11 (市值因子 30 组组合收益).....                 | 14 |
| 图 12 (各单因子对冲收益 (除市值因子外) ).....            | 15 |
| 图 13 (市值因子 30 组对冲组合收益).....               | 15 |
| 图 14 (EPS_R+ MKTVAL 组合及对冲收益曲线).....       | 16 |
| 图 15 (ROE+ MKTVAL 组合及对冲收益曲线).....         | 16 |
| 图 16 (EPS_R+PTBR 组合及对冲收益曲线).....          | 17 |
| 图 17 (ROE + PCF 组合及对冲收益曲线).....           | 17 |
| 图 18 (各原始单因子对应的 HS 复合因子组合收益).....         | 18 |
| 图 19 (各原始单因子对应的 HS 复合因子组合对冲收益).....       | 18 |
| 图 20 (各原始多因子对应的 HS 复合因子组合收益).....         | 19 |
| 图 21 (各原始多因子对应的 HS 复合因子组合对冲收益).....       | 19 |
| 图 22 (Hurst 分类树过滤后各单因子最优分组收益).....        | 20 |
| 图 23 (Hurst 分类树过滤后各单因子最优分组对冲收益).....      | 20 |
| 图 24 (Hurst 分类树过滤后市值因子各分组收益).....         | 21 |
| 图 25 (Hurst 分类树过滤后市值因子各分组对冲收益).....       | 21 |
| 图 26 (Hurst 分类树过滤后各多因子最优分组收益).....        | 22 |

|  |    |
|--|----|
| 图 27 (Hurst 分类树过滤后各多因子最优分组对冲收益).....               | 22 |
| 图 28 (原始市值因子分组效果回归).....                           | 23 |
| 图 29 (Hurst 分类树过滤后的市值因子分组效果回归).....                | 23 |
| 图 30 (原始 EPS_R+ MKTVAL 因子回归).....                  | 24 |
| 图 31 (Hurst 分类树过滤的 EPS_R+ MKTVAL 因子回归).....        | 24 |
| 表 1 原始单因子.....                                     | 13 |
| 表 2 原始多因子.....                                     | 13 |
| 表 3 各原始单因子最优组合回测表现及因子分组区分效果.....                   | 14 |
| 表 4 各原始单因子最优对冲组合回测表现及因子分组区分效果.....                 | 15 |
| 表 5 各原始多因子最优组合回测表现及因子分组区分效果.....                   | 17 |
| 表 6 各原始多因子最优对冲组合回测表现及因子分组区分效果.....                 | 17 |
| 表 7 各原始单因子对应 HS 复合因子最优组合回测表现及因子分组区分效果.....         | 18 |
| 表 8 各原始单因子对应 HS 复合因子最优对冲组合回测表现及因子分组区分效果.....       | 18 |
| 表 9 各原始多因子对应 HS 复合因子最优组合回测表现及因子分组区分效果.....         | 19 |
| 表 10 各原始多因子对应 HS 复合因子最优对冲组合回测表现及因子分组区分效果.....      | 19 |
| 表 11 各原始单因子经 Hurst 因子分类树过滤最优组合回测表现及因子分组区分效果.....   | 21 |
| 表 12 各原始单因子经 Hurst 因子分类树过滤最优对冲组合回测表现及因子分组区分效果..... | 21 |
| 表 13 各原始多因子经 Hurst 因子分类树过滤最优组合回测表现及因子分组区分效果.....   | 22 |
| 表 14 各原始多因子经 Hurst 因子分类树过滤最优对冲组合回测表现及因子分组区分效果..... | 22 |

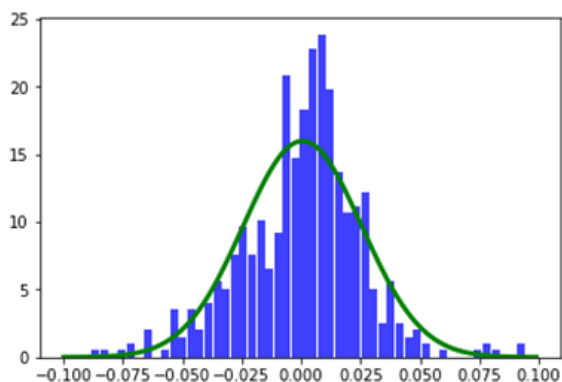
## 1. 基于“长记忆性”特征的研究思路

以“理性人假说”与“有效市场假说”为基础的现代资本市场理论是金融市场研究的重要里程碑。“理性人假说”的实质在于将投资者描述为追求“均值-方差有效性”的个体，即追求同等风险下收益最高的资产，或同等收益下风险最低的资产，也就是所谓“期望效用定理”。而“有效市场假说”则假定任何时刻所有（公开）信息均正确地反映到了价格之上，从而确保了模型中的投资者可以做到“理性”，即追求“均值-方差有效性”。既然价格已经反映了所有信息，则收益率时间序列具备了独立性特征，也就具备了运用概率论工具建立模型的条件。收益率序列是独立的（短期记忆性）、线性的、具有有限方差、服从正态分布，是该理论对于市场的基本观点。

然而，对于有效市场理论的质疑从未停息，主要反驳的依据集中在微观与宏观两个层面。在微观层面，Amos Tversky 与 2002 年诺贝尔经济学奖得主 Kahneman 提出的“前景理论”挑战了“期望效用定理”。“前景理论”认为大多数人在面临获得时是风险规避的，在面临损失时是风险偏爱的，人对损失比对获得敏感。因而真实的投资者不可能如模型中所描述的那样理性。2017 年诺贝尔经济学奖得主 Thaler 则进一步发展了“行为金融学”理论，认为人的决策并非在最大化最终全部收益，人不同的“心理账户”之间不能互换，这也让真实投资者的行为迥异于“理性人假说”。

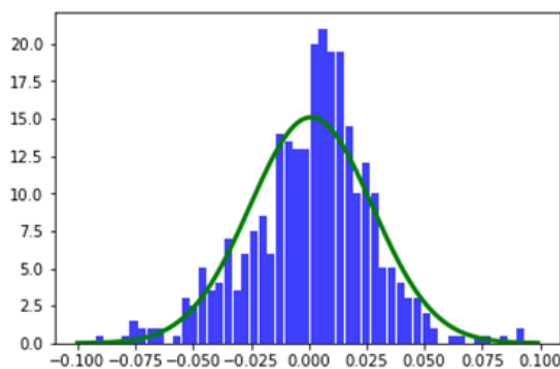
“理性人假说”在微观层面上的缺陷是否会在宏观层面上有所体现，从而最终影响到有效市场理论的实际运用价值呢？人们也在宏观层面，即市场数据统计的层面寻找答案——这也是金融市场中最容易检验的角度。通过对市场收益率数据的统计，几乎无论在哪个市场，哪个时间段，收益率分布普遍呈现出“尖峰厚尾”现象，与正态分布具有显著差异。如以下图 1 至图 4 展示了上证综指、沪深 300、上证 50、中证 500 的收益率分布与正态分布的对比，均不同程度呈现出“尖峰厚尾”现象。

图 1（上证综指周收益率分布）



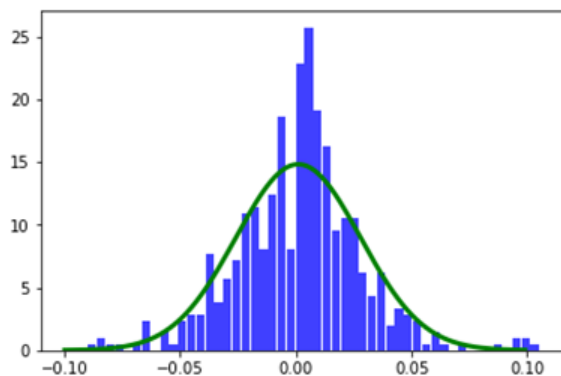
资料来源：东海证券

图 2（沪深 300 周收益率分布）



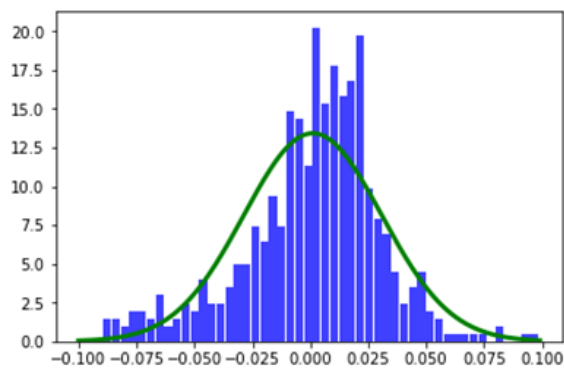
资料来源：东海证券

图 3 (上证 50 周收益分布)



资料来源：东海证券

图 4 (中证 500 周收益分布)



资料来源：东海证券

如果收益率满足随机游走与正态分布，则收益率序列的标准差具有以下标度关系：

eq. 1

$$\sigma(T\tau) = T^{\frac{1}{2}}\sigma(\tau)$$

即累计波动率随规模（时间长度） $T$  的变化速度为  $T$  的平方根。如果对上式两边取对数：

eq. 2

$$\log(\sigma(T\tau)) = \frac{1}{2}\log(T) + \log(\sigma(\tau))$$

也就是说累计波动率的对数相对于规模的对数的增长速度为斜率 0.5。而不管产生的原因如何，实际市场收益率的“尖峰厚尾”现象意味着正态分布模型对于累计风险的估计是不足的。如风险管理中常用的 VaR 模型，其基本假设就是有效市场与随机游走，对 VaR 模型的过度依赖也是引起 2008 年美国市场量化投资普遍亏损的原因之一。

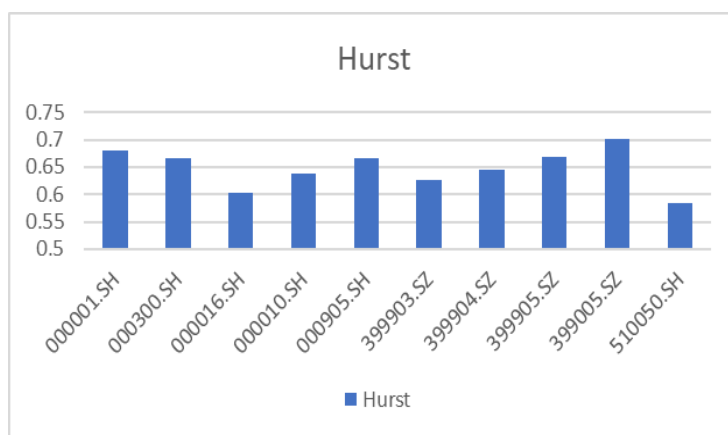
20 世纪初 Hurst 在研究尼罗河水文资料时提出了 R/S 分析法，成为刻画时间序列累计波动率相对于规模的增长速度的一般方法。所谓“一般”是指该方法对时间序列的独立性与分布没有强制假设，从而具有普遍适用性。通过 R/S 分析法计算出的 Hurst 指数取值范围为(0,1)，当 Hurst 指数为 0.5 时意味着时间序列是随机游走的，符合前述 1/2 法则；Hurst 指数大于 0.5 时意味着时间序列具有持续沿当前方向发展的倾向，即长记忆性；Hurst 指数小于 0.5 时，意味着时间序列具有反持续性倾向，即短记忆性。

20 世纪 60 年代 Mandelbrot 首次将该方法用于证券市场收益率分析上。上世纪 90 年代 Peters 在此基础上发展了分形市场理论，认为金融市场本质上是一个复杂的非线性动力系统。不同投资者在投资期限、投资行为、投资周期、对信息的反应等方面的差异决定了市场的流动性特征。市场具有长记忆性：过去信息对未来产生长期影响；与自相似性：不同时间标度下具有相同或相似的统计规律。

我们对市场各主要指数近 10 年收益率序列的 Hurst 指数进行统计，验证了长记忆性特征（Hurst 指数大于 0.5）在市场中是普遍存在的。近年来，随着计算机运算能力的提升，研究者开始尝试利用计算机建立“人工金融市场模型”，运用高性能计算机仿真金融市场的运行与演化，如圣塔菲研究院的“人工股票市场模型”，也从实验金融学的角度验证了上述结论。



图 5 (各主要指数收益率 HURST 指数 (2007-2017) )



资料来源：东海证券

R/S 分析法与 Hurst 指数为探索金融市场的本质特征，发展金融市场理论提供了重要的工具。但在实际的投资中如何运用这一工具，是我们更感兴趣的话题。在已有的相关研究中，Hurst 指数主要用于改进趋势跟踪或反转交易中的择时效果。由于 Hurst 指数仅反映长记忆性特征，并不反映价格运动的方向，因此通常与均线系统或价格通道配合使用。

因子选股是量化策略研究的另一个重要领域。通过长期研究，人们发掘出了规模、价值、成长、情绪等众多因子，并按多种方式对因子进行组合，形成多因子模型。各类因子实质上从不同的角度表征了影响市场的信息因素，多因子的组合方式代表了策略对因子之间相互关系的理解。然而，市场本身对因子的反应在不同的时间范围内是不断变化的，这也导致许多因子模型即使能取得很好的平均收益，也必须付出极大的回撤。因此，因子择时又成了因子模型运用时的一个重要课题，如按照某种标准将市场划分为强市、弱势、震荡市，在不同市场情况下选择运用不同的因子模型。

我们注意到，因子数据的更新频率通常慢于市场价格的变化，且影响市场价格的因素众多，相互关系复杂，市场价格与因子之间绝非简单的线性关系。在因子值一定的情况下，市场价格的变动反映了市场对因子或信息因素的反应的强度和持续程度。而 Hurst 指数本身又是对时间序列记忆性的一种度量工具，是波动强度与持续程度的体现，其本身也是一种择时的指标。因此，我们对能否利用 Hurst 指数来改进因子择时，进行了可行性研究。

## 2.Hurst 模型的本质与意义

### 2.1.R/S 分析方法

英国水文学家 Hurst 在参与尼罗河水坝工作时，需要根据雨水的流入量来决定每年水库的放水量。他研究了公元 622~1469 年埃及人保存的水文资料，发现记录并没有呈现出随机性，大流量的泛滥往往跟在大流量泛滥之后——即具有长记忆性的非周期循环。因此他发展了自己的方法论来解决时间序列记忆性特征的测量问题。Mandelbrot 和 Wallis 对 Hurst 理论（尤其是 R/S 分析方法）作了较为系统的研究，并给出严格的数学描述。20 世纪 90 年代，Peters 则将该方法引入资本市场领域，发展了分形市场理论。

方法（重标极差法）计算 Hurst 指数的过程如下：

1. R/S 分析对长度为  $M$  的时间序列进行预先处理，按下面公式转换为长度为  $N = M - 1$  的对数比时间序列：

eq. 3

$$N_i = \log(M_{i+1}/M_i), \quad i = 1, 2, \dots, (M-1)$$

在处理金融时间序列时，通常采用收益率时间序列代替上述处理。

2. 将长度为  $N$  的时间序列均分为  $A$  个长度为  $n$  的连续子区间，记为  $I_a$ ,  $a = 1, 2, \dots, A$ ,  $I_a$  中的每个点记为  $N_{k,a}$ 。计算每个子区间  $I_a$  的均值：

eq. 4

$$e_a = (1/n) \sum_{k=1}^n N_{k,a}, \quad a = 1, 2, \dots, A$$

3. 每个子区间  $I_a$  对于均值的累积截距的时间序列定义如下：

eq. 5

$$X_{k,a} = \sum_{i=1}^k (N_{i,a} - e_a), \quad k = 1, 2, \dots, n$$

4. 每个子区间  $I_a$  的极差被定义为该子区间中  $X_{k,a}$  的最大值与最小值之差：

eq. 6

$$R_{I_a} = \max_k (X_{k,a}) - \min_k (X_{k,a}), \quad k = 1, 2, \dots, n$$

5. 计算每个子区间  $I_a$  的标准差：

eq. 7

$$S_{I_a} = \left( (1/n) \sum_{k=1}^n (N_{k,a} - e_a)^2 \right)^{\frac{1}{2}}, \quad a = 1, 2, \dots, A$$

6. 每个子区间  $I_a$  的重标极差值  $(R/S)_a = R_{I_a}/S_{I_a}$ ，则所有长度为  $n$  的  $A$  个子区间的平均  $R/S$  值为：

eq. 8

$$(R/S)_n = (1/A) \sum_{a=1}^A (R_{I_a}/S_{I_a})$$

7. 对不同的时间尺度（即  $n$  取不同的值）重复上述步骤 2~6，从  $n$  的最大值递减直至  $n$  的最小值  $n = N/2$ ，也就是划分为 2 个子区间。这样我们就得到了一组对应于不同  $n$  值的  $(R/S)_n$  值。

8. 样本的平均重标极差值与样本长度之间存在标度关系：

eq. 9

$$(R/S)_n = c \times n^H$$

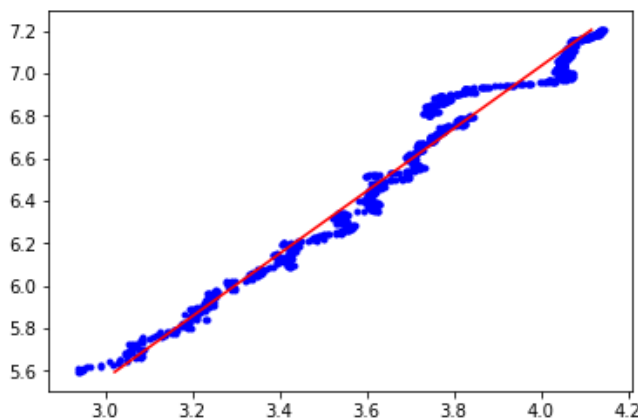
等式两边取对数：

eq. 10

$$\log(R/S)_n = \log(C) + H \log(n)$$

对 $\log(R/S)_n$ 序列与 $\log(n)$ 序列进行线性回归，即可得到斜率 H，也就是 Hurst 指数。

图 6 (上证综指 HURST 指数 H=0.68 (斜率))



资料来源：东海证券

从 Hurst 指数的计算过程可以看出，累积截距 $X_{k,a}$ 刻画了时间序列中样本对均值的累积偏离程度。样本朝同一方向持续发展，极差 $R_{I_a}$ 值就较大，对应于价格时间序列中的向上或向下趋势；反之极差 $R_{I_a}$ 值较低，则对应于价格时间序列中的震荡行情。如果将时间序列中样本的次序打乱，则 Hurst 指数的结果会完全不同；而样本次序对于标准差的计算却没有任何影响。也就是说 Hurst 指数的计算中保留了时间序列的“过程”信息——记忆性。

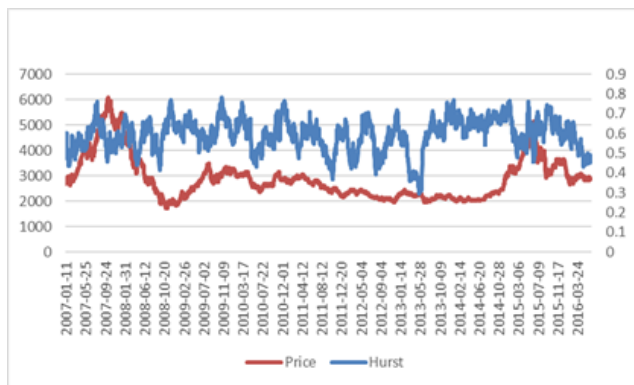
通过反复变换子区间长度 n，我们可以获得累积波动对于规模（时间长度）的函数关系。对比回归方程（eq. 10）与公式（eq. 2）所示的 1/2 法则，可以发现两者极为相似。但 R/S 分析法的优点在于不必假设所测度的时间序列的分布特征：无论是否正态分布，R/S 分析的稳健性均不受影响。对于正态分布的时间序列，其结论与 1/2 法则一致：Hurst 指数为 0.5。

## 2.2.移动 Hurst 指数的择时意义

对一个时段的时间序列样本的 Hurst 指数研究有助于探索样本整体的非周期循环特征，具有重要的学术意义。然而对于不同时段、不同频率的时间序列样本，对应的 Hurst 指数也是不断变化的。在实际的投资领域，更值得关注的是在某一时刻 Hurst 指数所表征的信息。采用类似于移动平均线的构造方式，对某个时刻将该时刻前一个固定周期 T 的时间序列作为样本，即可构造出当前时刻的 Hurst 指数值。对每个时刻进行类似的处理，即可获得移动 Hurst 指数序列。我们希望在整体样本 Hurst 指数展现的总体记忆性特征的基础上，通过移动 Hurst 指数序列所展现的记忆性的时变特征，寻找挖掘市场机会的新途径。

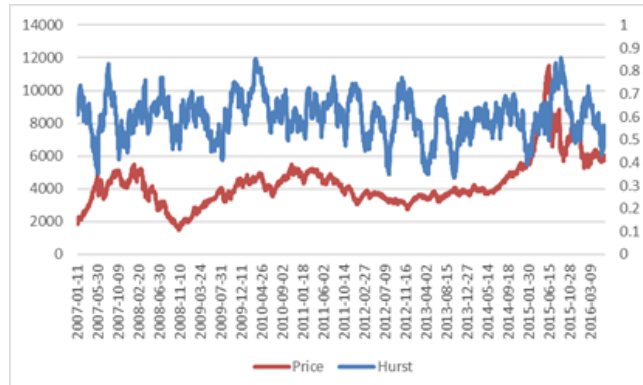


图 7 (上证综指及其移动 HURST 指数)



资料来源：东海证券

图 8 (中证 500 及其移动 HURST 指数)



资料来源：东海证券

注：周期 $T=66$ 。

从图 7、图 8 可以看出，根据上证综指和中证 500 的收益率数据计算的移动 Hurst 指数序列基本在 $[0.55, 0.7]$ 范围内浮动。这意味着价格时间序列相对于随机游走具有更长的记忆性。移动 Hurst 指数较大时，其对应周期的时间序列经历了较大的波幅，通常被认为表征了趋势的形成。移动 Hurst 指数较小时，其对应周期的时间序列在小幅度范围内震荡，通常被认为表征了反转信号。然而需要注意的是，在震荡上行的趋势中遭遇压力位，或震荡下行的趋势中遭遇支撑位时，价格也会出现横盘，也可能导致 Hurst 指数数值较小。此外，构造移动 Hurst 指数序列时，设置的参数周期 $T$ ， $T$ 的不同也可能在微观层面上造成移动 Hurst 指数具有不同的值。

由于金融时间序列的长记忆性，Hurst 指数通常较大。以上证综指为例，以 66 个交易日 ( $T=66$ , 1 季度) 为周期计算的移动 Hurst 指数的均值为 0.59，标准差 0.087，中位数为 0.60。然而移动 Hurst 指数也并非持续保持高位，甚至有数次短暂下探突破 0.5，这体现了价格时间序列在趋势与震荡之间的非周期循环。由于 Hurst 指数较高是常态，较低值出现后会迅速向正常水平回归，因此我们认为移动 Hurst 指数出现较小值时体现的信号强度比较大值时更强，也更值得关注。

### 3. 基于 Hurst 指数的因子择时模型

单因子/多因子选股模型的最大风险在于因子信号的滞后，这是因子数据更新频率通常低于价格数据决定的，这造成即使表现很好的因子也有较大的回撤，也会影响因子的区分效果。因子择时的传统做法是，研究不同因子在不同板块、行业、市场行情下的表现，在实际投资中首先按照一定方法形成对市场情况的判断，然后选择对应的因子，在因子、板块、行业、风格之间轮动。

从已有的研究资料来看，这样的因子择时可以起到一定的效果，但仍然存在一些缺陷：对市场情况的判断的方式有很多，有定性的也有定量的，但准确判断市场情况本身就是一项艰巨的任务；一些因子择时模型中，对市场情况的主观判断依赖较大，降低了验证结果的可信度；择时判断的表现形式繁杂，描述方式随意性很大，不同因子择时模型很难相互比较。

Hurst 指数是表征时间序列非周期循环特征——记忆性的有效指标，而且具备几个优越的特性：R/S 分析方法对样本数据没有任何先验性的假设，使之具有普遍适用性；Hurst 指数的取值范围为 $[0, 1]$ ，是一个有界的数值表示，一个 Hurst 指数值的含义不受样本的规模、尺度、独立性等因素的影响，具有作为一般统计量的良好特性；Hurst 指数是一个简单的数值值，便于进行比较计算，也便于与各种数值型因子结合。

有鉴于此，我们考虑利用 R/S 分析方法的优良特性，将具有择时意义的 Hurst 指数应用到单因子/多因子选股模型中，作为独立因子或其他因子的调节参数。这样就在单因子/多因子选股模型中纳入了择时因素，而不是让选股和择时两个过程截然分开。此外，这种因子择时的方式可以直接应用原有的选股模型回溯框架进行检验，避免了对市场情况判断和分类方式的依赖，从而可以具有更好的客观性和可操作性。

### 3.1.HS 调节因子模型

在本模型中，我们尝试基于移动 Hurst 指数构建一个调节因子，在实际运用中将其与其他因子组合为一个复合因子，从而将移动 Hurst 指数体现的择时信息融入选股因子中，以期改进原因子的选股效果。

#### 3.1.1.HS 调节因子的构建

由于 Hurst 指数仅体现了时间序列的记忆性特征，是一种价格运动的“速度”指标，并不包含价位与价格运动方向信息。因此，利用 Hurst 指数构造因子时可以考虑将价位和方向因素与之结合。在本模型中我们仅考虑了价位因素。

移动 Hurst 指数较低时（接近或低于 0.5），样本时间序列相对处于窄幅震荡的形态。根据技术分析的观点这可能有几种情况：下跌趋势结束底部调整；下跌过程中遇支撑位低位整理；上涨趋势结束顶部调整；上涨过程中遇阻力位高位整理。移动 Hurst 指数较高时（大于 0.6），样本时间序列波幅较大，处于向上趋势，向下趋势，或宽幅震荡的过程中。前已述及（图 7、图 8），移动 Hurst 指数处于低位的频率较低，持续时间较短，所体现的信号强度更强；移动 Hurst 指数短暂处于低位后往往快速回升，即价格时间序列短暂窄幅震荡结束后会转而向上向下趋势运动。从技术分析角度，均线系统中的支撑位阻力位对投资者的心理影响巨大，是多空反复争夺的“战线”，因此发生价格震荡所处的价位也是重要因素。

基于此我们借助均线系统的思想，引入“当前相对均线价差”结合移动 Hurst 指数构建了“HS 调节因子”。另需说明的是，“HS 调节因子”是用于其他因子选股模型的改进，并不包括方向性判断信息，不能作为独立因子使用；方向性判断信息仍由原因子承担。“HS 调节因子”的构建步骤如下：

1. 构建移动 Hurst 指数：在每个交易日  $i$ ，对每个股票  $a$  向前取  $T$  个交易日的收益率序列作为样本，按 2.1 所述 R/S 分析方法计算 Hurst 指数。所有交易日的移动 Hurst 指数形成移动 Hurst 指数序列  $\{MH_{i_a}\}$ 。

2. 对股票  $a$  每个交易日  $i$  的收盘价  $P_{i_a}$  移动均线  $MA_{i_a}$  计算相对均线价差  $S_{i_a}$ ：

eq. 11

$$S_{i_a} = (P_{i_a} - MA_{i_a}) / MA_{i_a}$$

3. 对股票  $a$  每个交易日  $i$  计算 HS 调节因子  $HS_{i_a}$ ：

eq. 12

$$HS_{i_a} = MH_{i_a} \times S_{i_a}$$

4. 所有交易日的 HS 调节因子形成 HS 调节因子序列  $\{HS_{i_a}\}$

#### 3.1.2.HS 调节因子的评分

从 3.1.1 的定义可以看出, HS 调节因子体现了价格时间序列靠近均线的程度和震荡的幅度。当  $HS_{i_a} \approx 0$  时, 价格  $P_{i_a}$  在均线  $MA_{i_a}$  附近小幅震荡;  $HS_{i_a} \ll 0$  时价格  $P_{i_a}$  远低于  $MA_{i_a}$ , 且 (或) 移动 Hurst 指数  $MA_{i_a}$  较大, 即前期价格以较大幅度下跌。  $HS_{i_a} \gg 0$  时, 价格  $P_{i_a}$  远高于  $MA_{i_a}$ , 且 (或) 移动 Hurst 指数  $MA_{i_a}$  较大, 即前期价格以较大幅度上涨。根据前面的分析, Hurst 指数处于低位后大幅回升的概率较大, 即此后出现趋势性行情的可能性较大, 因此, 我们对 HS 调节因子的评分方式也遵循将 HS 调节因子绝对值越小的股票给予越高的评分的原则。具体评分方式为:

1. 在每个交易日  $i$  对样本股票池中每个股票  $a$  计算 HS 调节因子  $HS_{i_a}$ 。
2. 将样本股票池中所有  $n$  个股票, 按 HS 调节因子的绝对值  $|HS_{i_a}|$  按下面公式进行归一化调整:

eq. 13

$$HS\_Adj_{i_a} = \frac{\max_a(|HS_{i_a}| - \min_a(|HS_{i_a}|) + 1)}{|HS_{i_a}| - \min_a(|HS_{i_a}|) + 1}$$

3. 每个交易日  $i$ , 对样本股票池中所有  $n$  个股票  $a$  按下面公式计算因子分:

eq. 14

$$SC(HS)_{i_a} = HS\_Adj_{i_a} \times 100 / \max_a(HS\_Adj_{i_a})$$

经过以上步骤的处理, 样本股票池中所有  $n$  个股票按因子的绝对值  $|HS_{i_a}|$  评分,  $|HS_{i_a}|$  越小分值越高, 所有分值被归一化为百分制。

### 3.1.3. 基于 HS 调节因子的因子分组检验方法

设计 HS 调节因子的目的是用于优化其他因子选股模型的效果。我们将被调节的因子称为原因子。我们希望由原因子负责对价格运动方向的判断, 加入 HS 调节因子后, 利用其包含的记忆性和价位信息得到更优的投资组合。

在交易日  $i$ , 我们首先对样本股票池中的每只股票  $a$ , 按原因子其自身的意义进行因子评分, 并同样被归一化为百分制分值。以市值因子为例, 在交易日  $i$  股票  $a$  的市值因子评分记为  $SC(MktVal)_{i_a}$ 。我们同样对该股票计算其 HS 调节因子评分, 记为  $SC(HS)_{i_a}$ 。然后, 我们将原因子与 HS 调节因子的评分按一定权重加权, 得到调节后的 HS 复合因子分值。如, 市值因子评分对应的权重记为  $w_{MktVal}$ , HS 调节因子评分记为  $w_{HS}$ , 则市值因子经调节后的 HS 复合因子分值为:

eq. 15

$$SC(MktVal\_HS)_{i_a} = SC(MktVal)_{i_a} \times w_{MktVal} + SC(HS)_{i_a} \times w_{HS}$$

在回测中, 我们以一般因子选股模型的检验方式检验调节后的 HS 复合因子: 在回测框架中设定的每个建仓日  $i$ , 计算样本股票池中每只股票调节后的 HS 复合因子分值; 按分值降序排列样本股票池中的所有股票, 并按分值高低等分为  $G$  组。每个股票分组作为一个投资组合, 其中每个股票在该组中所占权重按其因子分占该组股票总分数的比例配置。

## 3.2. Hurst 因子分类树模型

在 3.1 所述的 HS 调节因子模型中，原因子与 HS 调节因子的组合方式本身可能会影响 HS 复合因子的显著性。例如，在建仓日  $i$  股票  $a$  的原因子分值较高，HS 调节因子分值较低；而股票  $b$  的原因子分值较低，HS 调节因子分值较高，这使得股票  $a$  和股票  $b$  的 HS 复合因子分值可能非常接近，造成了信息的丢失。

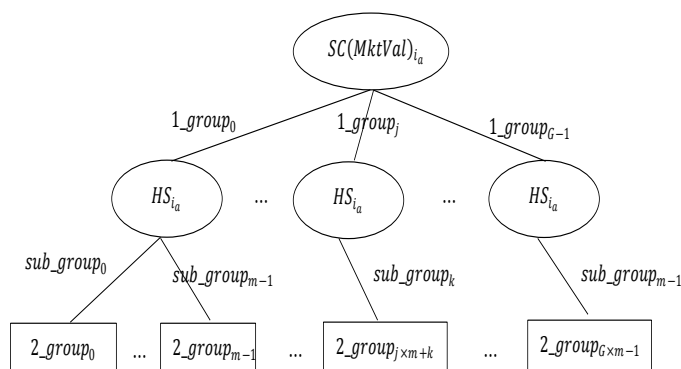
因此，我们考虑将因子择股的过程分为两个阶段：原因子分组阶段和 Hurst 指数过滤阶段，以期在保留原因子区分效果的同时，利用 Hurst 指数增强择股的效果。我们构建了一个“Hurst 因子分类树”来实现这个过程。Hurst 因子分类树是一个简单的分类决策树，由于决策树是一种非参数的层次模型，没有对参数形式的任何假定，因此适用于因子之间相互关系不确定的情形。

Hurst 因子分类树的原因子分组步骤如下：

1. 在回测框架的每个建仓日  $i$ ，对每只股票  $a$  计算其原因子评分值与移动 Hurst 指数。
2. 将样本股票池中所有股票按原因子评分值的降序进行排列，按因子评分值等分为  $G$  组。
3. 对按原因子分组的每个分组中的股票，按 Hurst 指数的升序排列，再等分为  $m$  个子组，共形成  $g = G \times m$  个子分组。
4. 每个子分组中的股票按等权重分配。

仍以原因子为市值因子为例，其 Hurst 因子分类树如图 9 所示。在建仓日  $i$  对股票  $a$  按其市值因子分值  $SC(MktVal)_{i_a}$  在整个样本股票池中的排序分至第 1 层分组的第  $j$  组  $1\_group_j$ ，其中  $j \in [0, G-1]$ 。然后按其移动 Hurst 指数值  $HS_{i_a}$ ，在其所属的第 1 层分组中分至子分组  $k$ ，其中  $k \in [0, m-1]$ 。这样就得到了股票  $a$  在建仓日  $i$  的第 2 层分组  $2\_group_{j \times m + k}$ ，第 2 层分组为最终分组。

图 9 (Hurst 因子分类树)



资料来源：东海证券

在这样的分组方式下，属于第 1 层分组中不同组的股票，在第 2 层分组中也属于不同分组，其相对分组次序没有发生变化，原因子的区分效果得到了保留。

另需说明的是，Hurst 因子分类树模型还不是一个真正的决策树模型。简单起见，其每层分组的组数，每个分组的股票数，每只股票在组合中的权重均是我们先验设定的，具有一定的局限性。但 Hurst 因子分类树已经具备了决策树的形式，这为我们未来利用机器学习方法构造真正的决策树模型打下了基础。

## 4. 主要实证结果分析



## 4.1.回测检验方案

为了检验基于 Hurst 指数的因子择时模型对因子选股的改进效果，我们选择了一些常见的单因子与多因子作为原始因子。选择这些原始因子不是为了验证其本身的选股效果，而是为了研究利用 Hurst 指数对各类单因子/多因子选股模型进行改进的可行性。因此，我们首先对原始单因子/多因子形成的股票分组进行回测，然后对采用 HS 调节因子模型和 Hurst 因子分类树模型改进后的股票分组进行回测，最后比较不同原始因子在改进前后的效果。我们总共选取了 10 个单因子（表 1）和 4 个多因子（表 2）作为原始因子。

表 1 原始单因子

| 类型  | 名称                  | 说明                                |
|-----|---------------------|-----------------------------------|
| 规模类 | 市值 (MktVal)         | 流通市值=流通股本×收盘价                     |
| 估值类 | 市盈率 (PE TTM)        | 总市值/最近 12 个月净利润                   |
|     | 市销率 (PS)            | 总市值/最新报告期主营收入                     |
|     | 市现率 (PCF)           | 总市值/最新报告期经营活动产生现金流量净额             |
|     | 市净率 (PB)            | 总市值/最新年报净资产                       |
|     | 股价有形净值比 (PTBR)      | 总市值/(最近报告期股东权益-最近报告期无形资产)         |
| 成长类 | 每股收益 (EPS TTM)      | 最近 12 个月归属母公司净利润/总股本              |
|     | 每股收益增长率 (EPS_R)     | (本期每股收益-上年同期每股收益)/abs(上年同期每股收益)   |
|     | 每股净资产收益率 (ROE)      | 净利润/股东权益                          |
|     | 每股净资产收益率增长率 (ROE_R) | (本期净资产收益率-上期净资产收益率)/abs(上期净资产收益率) |

资料来源：东海证券

表 2 原始多因子

| 名称            | 说明                            |
|---------------|-------------------------------|
| EPS_R+ MktVal | 选择小市值且高收益增长率的股票               |
| ROE+ MktVal   | 选择小市值且高净资产收益率的股票              |
| EPS_R+PTBR    | 戴维斯双击变种：选择高收益增长率，且股价有形净值比低的股票 |
| ROE + PCF     | 选择净资产收益率高，市现率低的股票             |

资料来源：东海证券

我们以 2007 年 4 月 26 日至 2017 年 10 月 9 日的 A 股全市场数据为基准，过滤掉 ST 股、新股、次新股，作为样本股票池。针对 HS 调节因子模型和 Hurst 因子分类模型，我们计算了样本股票池中所有股票的移动 Hurst 指数，周期为 66 个交易日。自 2007 年 4 月 26 日起，至 2017 年 10 月 9 日止，每隔 20 个交易日为 1 个换股日，共计 121 期。

对于原始因子，在每个换股日  $i$  根据表 1、表 2 中的原始因子对样本股票池中的股票计算因子评分值，以因子分降序排列并等分为 30 个股票组合  $Port_{ji}, j \in [0, 29]$ 。将每个换股日  $i, i \in [0, 120]$  的第  $j$  个股票组合依次排列，就得到了分组  $j$  的投资组合序列：

eq. 16

$$Group_j = \{Port_{j0}, \dots, Port_{j120}\}, \quad j = 0, \dots, 29$$

对于 HS 调节因子模型，在每个换股日  $i$  首先按 3.1.1、3.1.2 所述方式计算样本股票池中的股票的 HS 调节因子分值。然后结合原始因子评分值按 3.1.3 所述方式得到 HS 复合因子分值。对各原始因子对应的 HS 复合因子分值，按与原因子相同的方式构建 30 个分组的投资组合序列。

对 Hurst 因子分类树模型, 对各原因子在每个换股日*i*得到的 30 个股票组合 $Port_{ji}$ , 每个股票组合按 3.2 所述方式进一步细分为 3 个股票组合 $Port_{ki}$ 。每个原始因子改进后共计获得共 90 组投资组合序列:

eq. 17

$$Group_k = \{Port_{k_0}, \dots, Port_{k_{120}}\}, \quad k = 0, \dots, 89$$

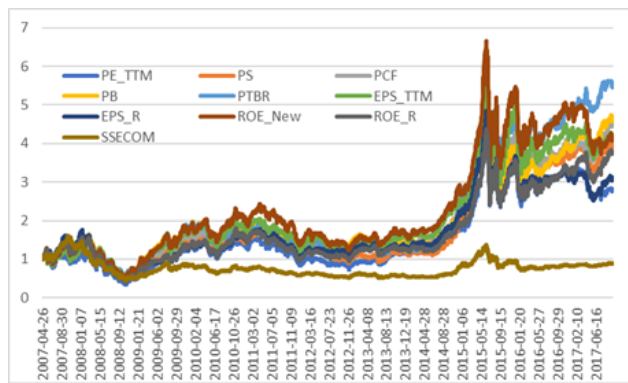
对各原始因子得到的各分组投资组合序列, 及其对应的采用 HS 调节因子模型和 Hurst 因子分类树模型得到的各分组投资组合序列进行回测, 其中包括投资组合序列的回测结果以及采用各主要组合对冲后的结果。对冲方案采用的指数包括上证综指、沪深 300、中证 200、中证 500 等。最后根据回测结果对比分析模型的改进效果。

对于回测结果, 我们除了统计期望收益率、最大回撤、期望收益/回撤比、胜率、Sharpe 等情况外, 还特别关注各因子的区分效果。即因子产生的各分组投资组合序列的表现是否与分组呈现出单调依赖关系, 这是衡量择股模型的重要标准。因此, 我们对各分组的序号与期望收益率的相关性, 序号与期望收益/回撤比的相关性进行了回归分析, 并采用 P 值来检验回归分析的显著性。

## 4.2.原始单因子组合实证结果分析

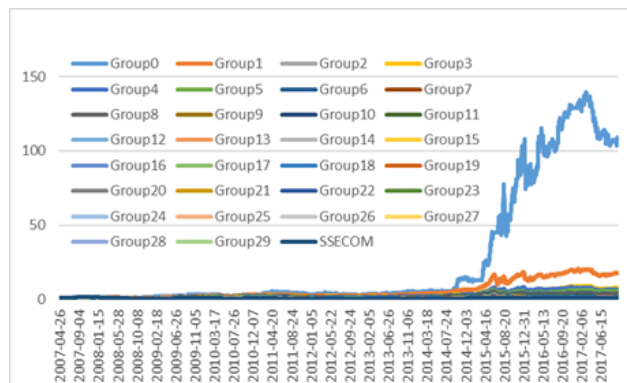
经过回测, 各原始单因子 (除市值因子外) 最优<sup>1</sup>分组所生成投资组合序列的收益曲线与上证综指 (SSE COM) 对比如图 10 所示, 市值因子各分组与上证综指 (SSE COM) 对比如图 11 所示:

图 10 (单因子最优组合收益 (除市值因子外))



资料来源: 东海证券

图 11 (市值因子 30 组组合收益)



资料来源: 东海证券

各单因子最优组合的回测结果, 及因子分组区分效果统计如下:

表 3 各原始单因子最优组合回测表现及因子分组区分效果

| 因子     | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回撤<br>比 Coeff. | P 值      |
|--------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| MktVal | 59.73%       | 52.60% | 1.135 | 55.11% | 0.8599          | 0.7852              | 1.14E-09 |
| PE TTM | 10.89%       | 73.53% | 0.148 | 58.38% | 0.0863          | 0.0549              | 6.76E-06 |
| PS     | 14.87%       | 70.92% | 0.210 | 56.92% | 0.8845          | 0.8864              | 1.94E-10 |
| PCF    | 16.69%       | 69.15% | 0.241 | 56.22% | 0.5850          | 0.6030              | 8.93E-06 |

<sup>1</sup> 所谓“最优”并非指回测结果最优, 而是指按因子评分值形成的各分组中序号最小的, 下同。



|         |        |        |       |        |        |        |          |
|---------|--------|--------|-------|--------|--------|--------|----------|
| PB      | 16.72% | 68.07% | 0.246 | 56.49% | 0.9269 | 0.9301 | 3.64E-17 |
| PTBR    | 18.64% | 68.04% | 0.274 | 55.31% | 0.9240 | 0.9273 | 1.15E-16 |
| EPS TTM | 15.48% | 68.60% | 0.226 | 58.22% | 0.9383 | 0.9409 | 5.39E-04 |
| EPS_R   | 12.06% | 75.72% | 0.159 | 56.81% | 0.6132 | 0.5765 | 1.13E-09 |
| ROE     | 16.73% | 69.17% | 0.242 | 58.03% | 0.8808 | 0.8805 | 3.39E-02 |
| ROE_R   | 14.27% | 74.28% | 0.192 | 57.55% | 0.6996 | 0.6778 | 2.43E-10 |

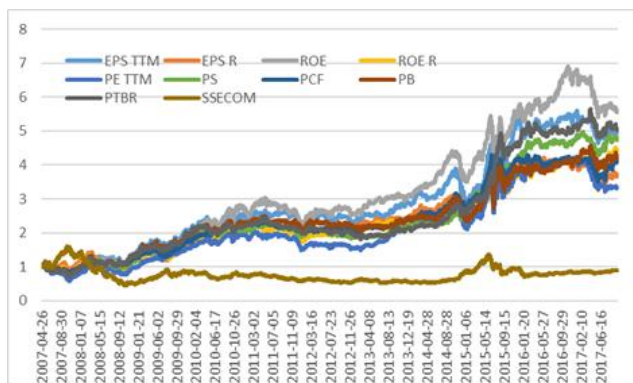
资料来源：东海证券

注：在表格中期望收益率/最大回撤比，被简称为“收益回撤比”，下同。

从原始单因子组合的回测结果来看，市值因子的期望收益率（年化）最好，其他因子效果一般，PE 因子效果最差。原始单因子组合的最大回撤普遍太大，期望收益/最大回撤比普遍偏低，仅市值因子的收益回撤比超过 1。从各因子组合的分组区分度<sup>2</sup>来看，我们测量了分组序列与期望收益率之间的相关系数，以及分组序列与收益回撤比的相关系数，市值、PS、PB、PTBR、EPS TTM、ROE 的区分效果较好，ROE R、EPS R、PCF 的区分效果一般、PE TTM 的区分效果最差。另外比较有趣的是，对 EPS TTM 和 ROE 因子，因子值大的组合效果较差，因子值小的组合效果反而较好，并具有比较明显的区分效果。而它们的增长率，ROE R、EPS R 却正好相反，因子值大的组合效果较好，因子值小的组合效果较差，但两者的区分效果相对较差。这样的结果可能是由财报数据的滞后性导致的，但此问题不属本次研究的范围。

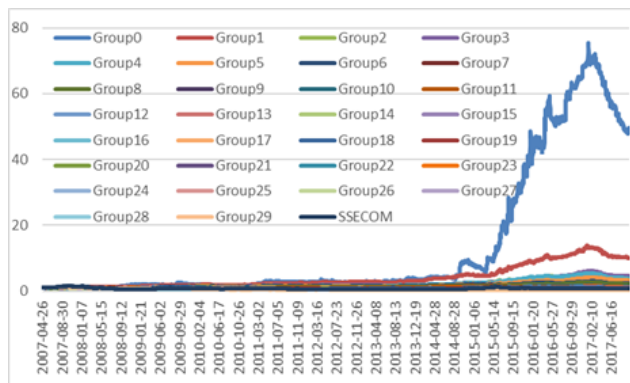
我们将单因子组合以市场主要指数进行对冲并进行回测，以上证综指为例，对冲组合的收益曲线与上证综指（SSE COM）对比如图 12（各单因子对冲收益（除市值因子外））所示，市值因子各分组对冲组合收益曲线与上证综指（SSE COM）对比如图 13（市值因子 30 组对冲组合收益）所示：

图 12（各单因子对冲收益（除市值因子外））



资料来源：东海证券

图 13（市值因子 30 组对冲组合收益）



资料来源：东海证券

注：分组序号越小的组合序列因子分值越高，下同。

各原始单因子最优对冲组合回测表现及因子分组区分效果统计如下：

表 4 各原始单因子最优对冲组合回测表现及因子分组区分效果

| 因子     | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回撤<br>比 Coeff. | P 值      |
|--------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| MktVal | 58.90%       | 44.38% | 1.327 | 48.90% | 0.8752          | 0.8785              | 9.97E-13 |
| PE TTM | 11.95%       | 27.83% | 0.430 | 55.59% | 0.1164          | 0.4358              | 7.32E-10 |

<sup>2</sup> 因子组合的分组区分度，即评价指标与因子分组序列的单调性，用以批判因子的区分能力。我们采用评价指标与因子分组序列的相关系数来衡量分组的区分度。

|         |        |        |       |        |        |        |          |
|---------|--------|--------|-------|--------|--------|--------|----------|
| PS      | 17.03% | 33.11% | 0.514 | 56.49% | 0.8975 | 0.8899 | 1.22E-11 |
| PCF     | 18.40% | 19.34% | 0.942 | 56.73% | 0.5635 | 0.6630 | 2.47E-06 |
| PB      | 18.80% | 31.22% | 0.602 | 57.08% | 0.9279 | 0.9039 | 1.77E-16 |
| PTBR    | 20.55% | 29.52% | 0.696 | 53.15% | 0.9274 | 0.9080 | 2.3E-16  |
| EPS TTM | 15.99% | 37.76% | 0.423 | 58.14% | 0.9374 | 0.8871 | 6.96E-01 |
| EPS_R   | 12.92% | 30.08% | 0.430 | 55.78% | 0.6244 | 0.6414 | 1.15E-11 |
| ROE     | 15.39% | 26.19% | 0.588 | 55.86% | 0.8894 | 0.7858 | 6.47E-01 |
| ROE_R   | 13.50% | 26.77% | 0.504 | 56.53% | 0.7116 | 0.7243 | 1.35E-12 |

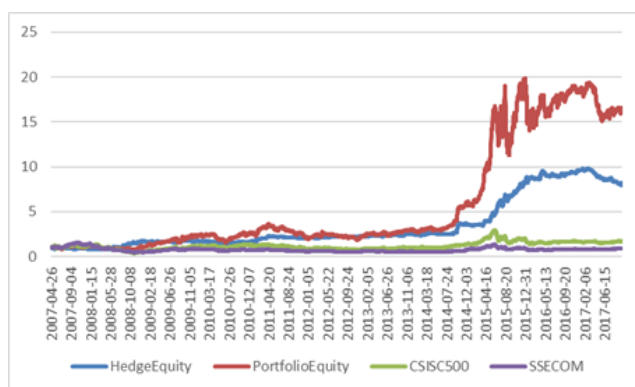
资料来源：东海证券

采用上证综指对冲以后，各单因子对冲组合相对于非对冲组合在期望收益率方面变化不大，少数略有提升，个别略有降低。市值因子的期望收益率仍然大幅领先。对冲组合的最大回撤大幅降低，收益回撤比大幅提高，但整体上回撤仍然较大。市值因子的回撤降低幅度不大，但收益回撤比仍然是唯一为正值的因子。值得注意的是市值因子在 2017 年以来收益持续回落，造成大幅回撤。从因子分组的区分效果来看，各因子的提升幅度不大，仅 PE 因子的区分效果大幅提升，但与其他因子相比，其期望收益、收益回撤比、区分效果均垫底。

### 4.3.原始多因子组合实证结果分析

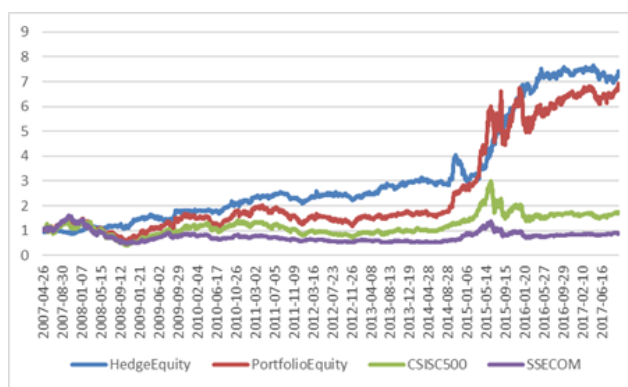
我们对表 2 所列的各原始多因子所生成的原始多因子投资组合序列进行了回测，并对各组合的最优对冲方案进行了统计。其中，EPS\_R+ MktVal 采用中证 500 (CSISC500) 对冲；ROE+ MktVal 采用上证综指对冲 (SSE COM)；EPS\_R+PTBR 采用上证综指对冲 (SSE COM)；ROE + PCF 采用上证综指对冲 (SSE COM)。以下图 14、图 15、图 16、图 17 分别展示了上述各原始多因子产生的最优投资组合序列及其对冲结果的收益曲线，和上证综指、中证 500 的对比。

图 14 (EPS\_R+ MKTVAL 组合及对冲收益曲线)



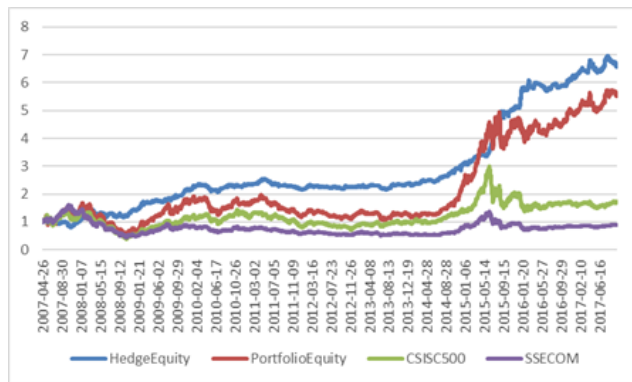
资料来源：东海证券

图 15 (ROE+ MKTVAL 组合及对冲收益曲线)



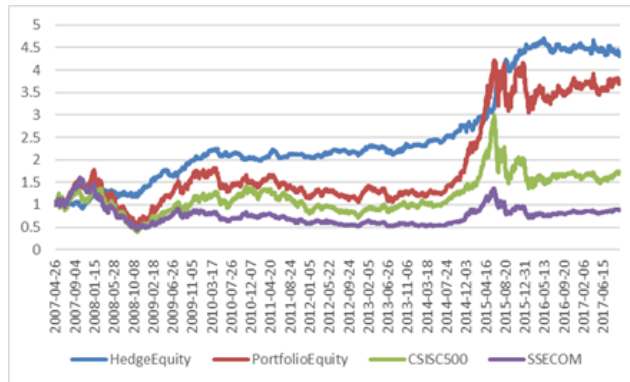
资料来源：东海证券

图 16 (EPS\_R+PTBR 组合及对冲收益曲线)



资料来源：东海证券

图 17 (ROE + PCF 组合及对冲收益曲线)



资料来源：东海证券

各原始多因子最优组合回测表现及因子分组区分效果统计如下：

表 5 各原始多因子最优组合回测表现及因子分组区分效果

| 因子            | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回<br>撤比 Coeff. | P 值      |
|---------------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| EPS_R+ MKTVAL | 32.38%       | 49.41% | 0.655 | 55.94% | 0.6623          | 0.5732              | 1.44E-06 |
| ROE+ MKTVAL   | 21.35%       | 63.91% | 0.334 | 54.68% | 0.3625          | 0.3137              | 2.46E-02 |
| EPS_R+PTBR    | 19.82%       | 68.50% | 0.289 | 57.55% | 0.9172          | 0.9155              | 1.37E-12 |
| ROE + PCF     | 14.07%       | 70.38% | 0.200 | 54.72% | 0.6839          | 0.6990              | 1.02E-10 |

资料来源：东海证券

各原始多因子最优对冲组合回测表现及因子分组区分效果统计如下：

表 6 各原始多因子最优对冲组合回测表现及因子分组区分效果

| 因子            | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回<br>撤比 Coeff. | P 值      |
|---------------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| EPS_R+ MKTVAL | 23.31%       | 23.07% | 1.009 | 48.23% | 0.6711          | 0.6729              | 7.00E-06 |
| ROE+ MKTVAL   | 22.20%       | 26.95% | 0.824 | 53.38% | 0.4179          | 0.1403              | 5.40E-04 |
| EPS_R+PTBR    | 13.46%       | 10.50% | 1.282 | 53.03% | 0.9173          | 0.6996              | 2.47E-06 |
| ROE + PCF     | 15.71%       | 13.57% | 1.160 | 55.58% | 0.6676          | 0.6649              | 3.48E-09 |

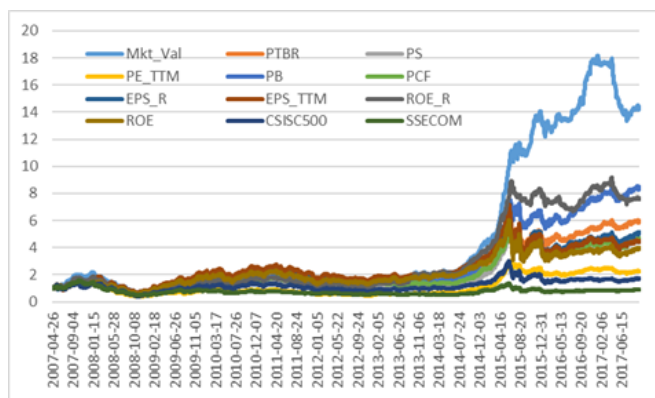
资料来源：东海证券

从上述各原始多因子产生组合的回测结果来看，各原始多因子的期望收益率和收益回撤比总体均好于除市值因子以外的其他原始单因子组合。尤其是采取对冲之后，收益回撤比大幅提升，普遍大于 1 或接近 1。部分原始多因子在采取对冲后，回撤大幅降低的同时，期望收益率还有所提高。从因子区分效果来看，原始多因子没有明显改进，甚至有的还有下降。如前面在原始单因子组合的测试结果中我们发现，ROE 因子，其因子区分效果与通常的预期并不一致，ROE 值小的分组反而能取得更好的收益。而市值因子和 PCF 因子的区分效果则符合预期，即因子值小的分组收益更好。这导致我们按照一定权重配比将 ROE 因子与市值因子或 PCF 因子进行组合，将其因子评分值加权求和作为原始多因子评分值，进而进行分组的做法，使得原始多因子的区分效果明显下降。但这一问题不属于本次研究的范畴，不再进行深入探讨。

#### 4.4.HS 调节因子模型实证结果分析

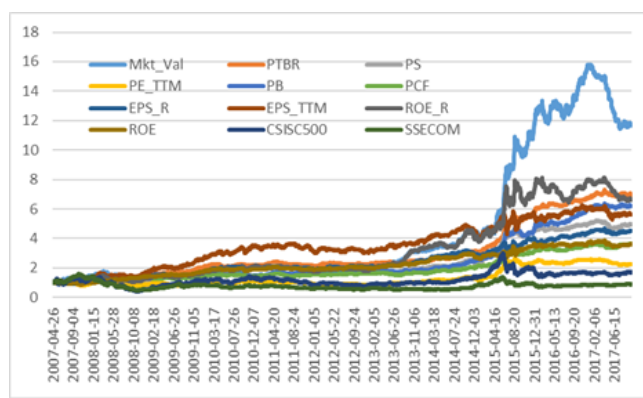
我们将各原始单因子、原始多因子按 HS 调节因子模型进行处理，得到各原始因子对应的 HS 复合因子评分值，并对样本股票池进行重新分组测试。图 18 展示了各原始单因子对应的 HS 复合因子产生的最优分组组合的收益曲线，图 19 展示了它们采用指数对冲方案后的收益曲线。为描述方便，各原始因子对应的 HS 复合因子在图表中均以原始因子标识，下同。

图 18 (各原始单因子对应的 HS 复合因子组合收益)



资料来源：东海证券

图 19 (各原始单因子对应的 HS 复合因子组合对冲收益)



资料来源：东海证券

各原始单因子对应的 HS 复合因子形成的最优分组投资组合回测表现，及其因子分组区分效果统计如表 7 所示。它们对应的指数对冲回测表现，及其因子分组区分效果统计如表 8 所示。

表 7 各原始单因子对应 HS 复合因子最优组合回测表现及因子分组区分效果

| 因子      | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回撤<br>比 Coeff. | P 值      |
|---------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| MktVal  | 30.57%       | 70.56% | 0.433 | 59.17% | 0.1749          | 0.1759              | 1.45E-03 |
| PE TTM  | 11.20%       | 70.99% | 0.158 | 57.24% | 0.3152          | 0.3508              | 5.91E-02 |
| PS      | 16.25%       | 71.32% | 0.228 | 57.71% | 0.9485          | 0.9475              | 1.10E-06 |
| PCF     | 16.78%       | 70.36% | 0.238 | 56.73% | 0.6984          | 0.7120              | 7.92E-04 |
| PB      | 23.82%       | 67.86% | 0.351 | 57.87% | 0.9450          | 0.9466              | 9.29E-08 |
| PTBR    | 19.55%       | 69.46% | 0.281 | 56.49% | 0.9776          | 0.9767              | 1.43E-07 |
| EPS TTM | 16.20%       | 67.47% | 0.240 | 57.24% | 0.2302          | 0.2263              | 7.68E-05 |
| EPS_R   | 17.78%       | 67.21% | 0.265 | 57.51% | 0.2765          | 0.2610              | 2.84E-05 |
| ROE     | 14.76%       | 65.26% | 0.226 | 57.12% | 0.2449          | 0.2353              | 4.33E-05 |
| ROE_R   | 22.56%       | 72.16% | 0.313 | 56.25% | 0.0926          | 0.0796              | 8.22E-04 |

资料来源：东海证券

表 8 各原始单因子对应 HS 复合因子最优对冲组合回测表现及因子分组区分效果

| 因子     | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回撤<br>比 Coeff. | P 值      |
|--------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| MktVal | 27.93%       | 32.17% | 0.868 | 53.93% | 0.0447          | 0.0913              | 3.27E-04 |
| PE TTM | 13.43%       | 30.89% | 0.435 | 58.30% | 0.4945          | 0.4467              | 3.22E-02 |
| PS     | 17.39%       | 28.31% | 0.614 | 56.57% | 0.9545          | 0.9282              | 1.68E-06 |
| PCF    | 13.84%       | 12.30% | 1.125 | 53.85% | 0.5975          | 0.7215              | 3.41E-03 |
| PB     | 20.02%       | 13.52% | 1.481 | 54.21% | 0.9463          | 0.8092              | 5.91E-05 |
| PTBR   | 21.42%       | 28.05% | 0.764 | 56.02% | 0.9808          | 0.9485              | 1.66E-06 |

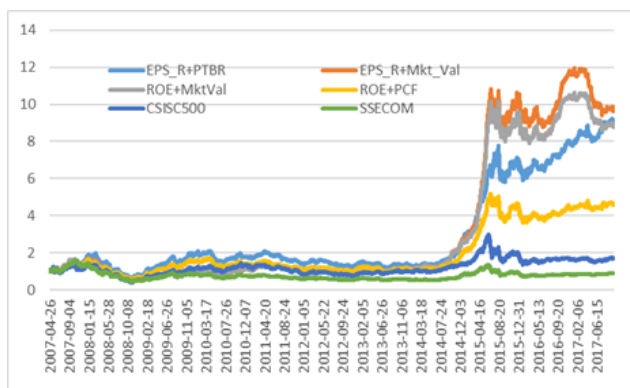


|         |        |        |       |        |        |        |          |
|---------|--------|--------|-------|--------|--------|--------|----------|
| EPS TTM | 18.89% | 24.13% | 0.783 | 57.39% | 0.0564 | 0.0926 | 8.16E-05 |
| EPS_R   | 16.34% | 16.41% | 0.996 | 56.69% | 0.0428 | 0.0037 | 2.18E-02 |
| ROE     | 13.66% | 18.29% | 0.747 | 56.37% | 0.0072 | 0.0864 | 1.87E-02 |
| ROE_R   | 20.80% | 31.17% | 0.667 | 54.45% | 0.0611 | 0.0244 | 2.25E-04 |

资料来源：东海证券

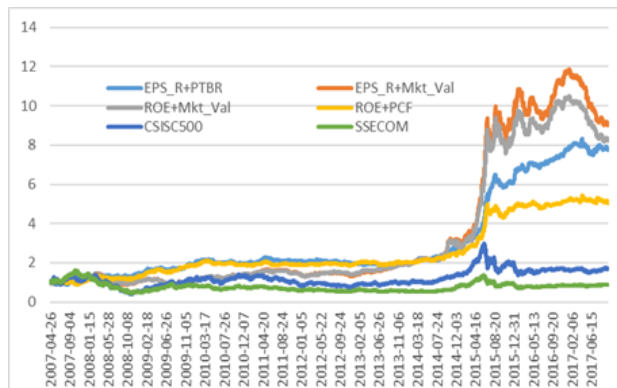
我们对各原始多因子同样采用 HS 调节因子模型进行处理，对应 HS 复合因子得到的最优分组组合的收益曲线如图 20 所示；采用指数对冲方案后的收益曲线如图 21 所示。

图 20 (各原始多因子对应的 HS 复合因子组合收益)



资料来源：东海证券

图 21 (各原始多因子对应的 HS 复合因子组合对冲收益)



资料来源：东海证券

各原始多因子对于的 HS 复合因子形成的最优分组投资组合回测表现，及其因子分组区分效果统计如表 9 所示。它们对应的指数对冲回测表现，及其因子分组区分效果统计如表 10 所示。

表 9 各原始多因子对应 HS 复合因子最优组合回测表现及因子分组区分效果

| 因子            | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回<br>撤比 Coeff. | P 值      |
|---------------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| EPS_R+ MKTVAL | 25.70%       | 74.00% | 0.347 | 57.87% | 0.1891          | 0.2091              | 4.97E-02 |
| ROE+ MKTVAL   | 24.54%       | 72.77% | 0.337 | 58.03% | 0.2446          | 0.2637              | 7.42E-02 |
| EPS_R+PTBR    | 24.73%       | 69.89% | 0.354 | 56.53% | 0.9074          | 0.9013              | 8.30E-15 |
| ROE + PCF     | 16.59%       | 69.53% | 0.239 | 55.11% | 0.7374          | 0.7447              | 1.83E-11 |

资料来源：东海证券

表 10 各原始多因子对应 HS 复合因子最优对冲组合回测表现及因子分组区分效果

| 因子            | 期望收益率<br>(年) | 最大回撤   | 收益回撤比 | 胜率(日)  | 分组-收益<br>Coeff. | 分组-收益回<br>撤比 Coeff. | P 值      |
|---------------|--------------|--------|-------|--------|-----------------|---------------------|----------|
| EPS_R+ MKTVAL | 24.77%       | 39.14% | 0.633 | 54.56% | 0.3052          | 0.3574              | 2.13E-02 |
| ROE+ MKTVAL   | 23.59%       | 36.22% | 0.651 | 54.45% | 0.3614          | 0.2860              | 1.17E-02 |
| EPS_R+PTBR    | 22.71%       | 18.93% | 1.200 | 53.86% | 0.9147          | 0.8111              | 3.04E-10 |
| ROE + PCF     | 17.63%       | 14.60% | 1.208 | 54.09% | 0.6974          | 0.6367              | 1.32E-08 |

资料来源：东海证券

从上述回测结果来看，就原始单因子而言，采用 HS 调节因子模型后，在提升期望收益率方面没有太大改进；在降低最大回撤方面，特别是采用对冲之后，对部分因子有一定效果；在因子分组的区分效果方面，对多数单因子甚至区分效果反而降低。尤其是对市值因子而言，仅最大回撤有所降低，其他各方面效果均显著降低。就原始多因子而言，采用

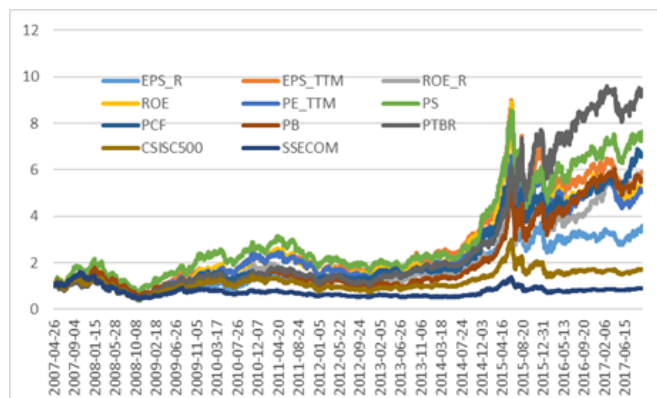
HS 调节因子模型后，在期望收益率、最大回撤、期望收益/最大回撤比方面，没有明显效果；因子分组的区分效果反而降低。

总体来说 HS 调节因子模型的效果并不理想。我们对各组测试结果进行分析后发现，主要原因不在于 HS 调节因子本身，而在于其与原始因子组合形成 HS 复合因子的方式上。采用因子评分值加权形成的复合因子评分，对原始因子和 HS 调节因子所包含的信息都造成了一定程度的丢失，弱化了因子的显著性。在某只股票的原始因子评分与 HS 调整因子评分出现一高一低的情况下，其 HS 复合因子评分值在样本股票池中出现的位置很难估计，这对 HS 复合因子的分组区分效果造成了严重影响。另需说明的是，多个因子加权组合也是多因子选股模型中常用的方式，这一方式虽然易于处理，但其缺陷在其他多因子选股模型中也同样存在。

## 4.5.Hurst 因子分类树模型实证结果分析

按 3.2 所述方式，我们在原始单因子/多因子形成的 30 个分组的基础上，采用 Hurst 因子分类树进行进一步细分，形成 90 个子分组。各原始单因子（除市值因子外）经 Hurst 因子分类树模型处理后形成的最优分组组合的收益曲线如图 22 所示；各最优分组组合采用指数对冲后的收益曲线如图 23 所示。

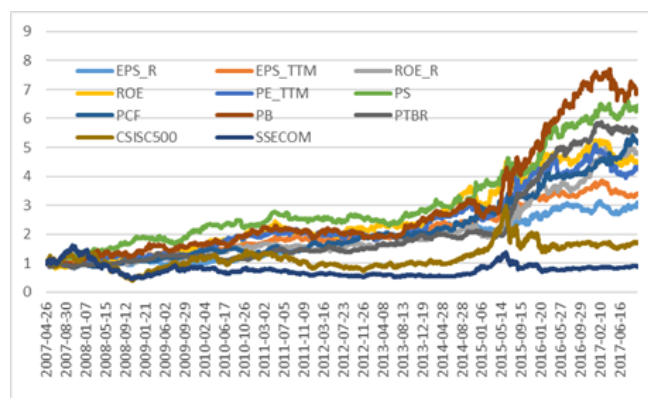
图 22 (Hurst 分类树过滤后各单因子最优分组收益)



资料来源：东海证券

注：除市值因子外

图 23 (Hurst 分类树过滤后各单因子最优分组对冲收益)



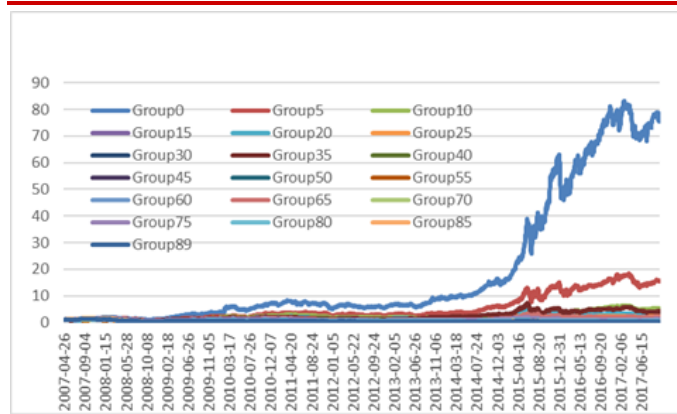
资料来源：东海证券

注：除市值因子外

市值因子经 Hurst 因子分类树模型处理后形成的各分组收益如图 24 所示，各分组采用指数对冲后的收益如图 25 所示。因分组数过多，为展示清晰，我们以 5 组为间隔选取分组进行展示。

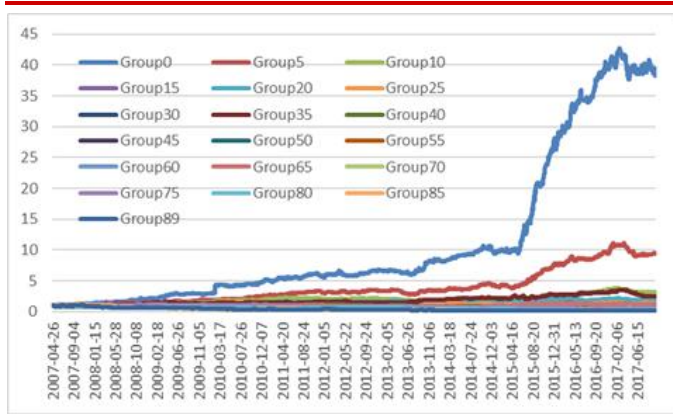


图 24 (Hurst 分类树过滤后市值因子各分组收益)



资料来源：东海证券

图 25 (Hurst 分类树过滤后市值因子各分组对冲收益)



资料来源：东海证券

各原始单因子经 Hurst 因子分类树模型处理后形成的最优分组组合回测表现及因子分组区分效果如表 11 所示；它们对应的指数对冲回测表现，及其因子分组区分效果统计如表 12 所示。

表 11 各原始单因子经 Hurst 因子分类树过滤最优组合回测表现及因子分组区分效果

| 因子      | 期望收益率 (年) | 最大回撤   | 收益回撤比 | 胜率(日)  | Sharpe | 分组-收益 Coeff. | 分组-收益回撤比 Coeff. | P 值      |
|---------|-----------|--------|-------|--------|--------|--------------|-----------------|----------|
| MktVal  | 54.40%    | 57.30% | 0.950 | 59.52% | 2.5749 | 0.8398       | 0.8108          | 9.00E-26 |
| PE TTM  | 17.71%    | 69.22% | 0.256 | 58.34% | 2.0641 | 0.1938       | 0.1988          | 3.99E-04 |
| PS      | 22.58%    | 67.75% | 0.333 | 56.57% | 1.7317 | 0.0548       | 0.0427          | 1.27E-05 |
| PCF     | 20.82%    | 69.42% | 0.300 | 55.00% | 1.6009 | 0.0946       | 0.0917          | 4.43E-04 |
| PB      | 18.73%    | 70.50% | 0.266 | 56.49% | 2.1715 | 0.0529       | 0.0346          | 1.65E-04 |
| PTBR    | 25.15%    | 69.36% | 0.363 | 56.88% | 1.1509 | 0.0226       | 0.0001          | 7.76E-04 |
| EPS TTM | 19.47%    | 70.27% | 0.277 | 58.62% | 2.0684 | 0.1144       | 0.1090          | 4.52E-03 |
| EPS_R   | 13.69%    | 74.43% | 0.184 | 56.84% | 1.3052 | 0.0804       | 0.0887          | 3.75E-06 |
| ROE     | 17.91%    | 70.93% | 0.253 | 57.47% | 2.9241 | 0.1234       | 0.1137          | 6.14E-03 |
| ROE_R   | 18.92%    | 74.29% | 0.255 | 58.18% | 2.2520 | 0.1104       | 0.1169          | 6.50E-04 |

资料来源：东海证券

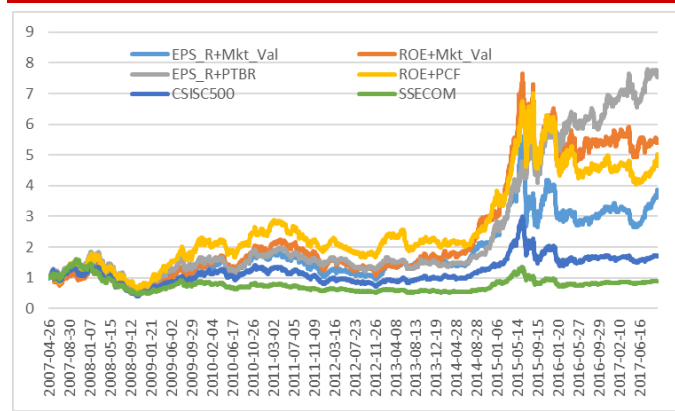
表 12 各原始单因子经 Hurst 因子分类树过滤最优对冲组合回测表现及因子分组区分效果

| 因子      | 期望收益率 (年) | 最大回撤   | 收益回撤比 | 胜率(日)  | Sharpe | 分组-收益 Coeff. | 分组-收益回撤比 Coeff. | P 值      |
|---------|-----------|--------|-------|--------|--------|--------------|-----------------|----------|
| MktVal  | 43.98%    | 13.64% | 3.224 | 53.11% | 2.1353 | 0.8550       | 0.6900          | 1.42E-15 |
| PE TTM  | 15.56%    | 22.53% | 0.691 | 53.81% | 1.4773 | 0.2024       | 0.1970          | 4.10E-02 |
| PS      | 20.47%    | 16.99% | 1.205 | 51.89% | 0.9874 | 0.0559       | 0.0050          | 1.28E-03 |
| PCF     | 17.81%    | 17.31% | 1.029 | 50.94% | 0.2583 | 0.1129       | 0.0862          | 8.54E-02 |
| PB      | 21.34%    | 21.99% | 0.970 | 56.10% | 0.3041 | 0.0539       | 0.0057          | 1.10E-05 |
| PTBR    | 18.77%    | 19.19% | 0.978 | 52.12% | 3.3591 | 0.0198       | 0.0933          | 5.722-01 |
| EPS TTM | 13.04%    | 14.89% | 0.875 | 51.65% | 1.7002 | 0.1018       | 0.0683          | 5.18E-01 |
| EPS_R   | 12.01%    | 20.70% | 0.580 | 52.64% | 3.1530 | 0.1194       | 0.1194          | 2.49E-03 |
| ROE     | 16.13%    | 22.51% | 0.716 | 52.75% | 3.7638 | 0.1075       | 0.0790          | 2.05E-02 |
| ROE_R   | 16.91%    | 18.22% | 0.928 | 54.23% | 1.5279 | 0.1159       | 0.1487          | 4.58E-02 |

资料来源：东海证券

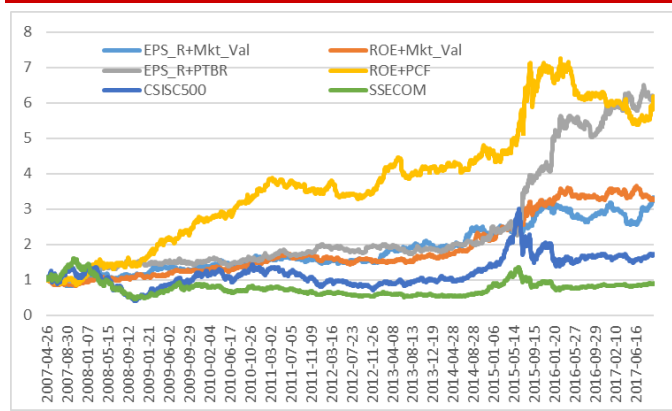
各原始多因子按Hurst分类树模型处理后形成的最优分组组合的收益曲线如图26所示；各最优分组组合采用指数对冲后的收益曲线如图27所示。

图 26 (Hurst 分类树过滤后各多因子最优分组收益)



资料来源：东海证券

图 27 (Hurst 分类树过滤后各多因子最优分组对冲收益)



资料来源：东海证券

各原始单因子经Hurst因子分类树模型处理后形成的最优分组组合回测表现及因子分组区分效果如表13所示；它们对应的指数对冲回测表现，及其因子分组区分效果统计如表14所示。

表 13 各原始多因子经Hurst因子分类树过滤最优组合回测表现及因子分组区分效果

| 因子            | 期望收益率 (年) | 最大回撤   | 收益回撤比 | 胜率(日)  | Sharpe | 分组-收益 Coeff. | 分组-收益回撤比 Coeff. | P 值      |
|---------------|-----------|--------|-------|--------|--------|--------------|-----------------|----------|
| EPS_R+ MKTVAL | 14.44%    | 74.19% | 0.195 | 56.69% | 1.6423 | 0.0783       | 0.0866          | 3.60E-05 |
| ROE+ MKTVAL   | 18.59%    | 70.81% | 0.263 | 57.00% | 2.5633 | 0.0015       | 0.0063          | 3.03E-06 |
| EPS_R+PTBR    | 22.58%    | 69.64% | 0.324 | 55.82% | 1.7724 | 0.0436       | 0.0277          | 3.55E-04 |
| ROE + PCF     | 17.54%    | 65.21% | 0.269 | 55.31% | 1.884  | 0.1164       | 0.1091          | 2.14E-04 |

资料来源：东海证券

表 14 各原始多因子经Hurst因子分类树过滤最优对冲组合回测表现及因子分组区分效果

| 因子            | 期望收益率 (年) | 最大回撤   | 收益回撤比 | 胜率(日)  | Sharpe | 分组-收益 Coeff. | 分组-收益回撤比 Coeff. | P 值      |
|---------------|-----------|--------|-------|--------|--------|--------------|-----------------|----------|
| EPS_R+ MKTVAL | 12.80%    | 20.07% | 0.638 | 53.11% | 0.5976 | 0.0856       | 0.1230          | 6.09E-03 |
| ROE+ MKTVAL   | 12.59%    | 18.26% | 0.690 | 50.00% | 0.4304 | 0.0031       | 0.0252          | 5.52E-01 |
| EPS_R+PTBR    | 19.74%    | 15.21% | 1.298 | 50.75% | 0.9929 | 0.0421       | 0.0265          | 5.72E-03 |
| ROE + PCF     | 19.98%    | 26.07% | 0.766 | 54.24% | 0.0712 | 0.1316       | 0.0164          | 3.05E-07 |

资料来源：东海证券

从回测结果来看，经Hurst因子分类树模型处理后，各原始单因子在对冲后的期望收益率有一定程度提升，而最大回撤则有较大幅度的降低，从而获得了较高的期望收益率/最大回撤比。同时各最优组合也取得了较好的Sharpe值。但对于各原始多因子而言，期望收益率/最大回撤比反而有所下降，各最优组合Sharpe值也不如单因子。

就原始因子中表现最好的市值因子而言，Hurst因子分类树模型的改进效果明显，对冲后的最大回撤从44.38%降至13.64%，期望收益率/最大回撤比从1.327提升至3.224。对比采用Hurst因子分类树模型过滤前（图13）后（图25）的各分组收益曲线，最优分组

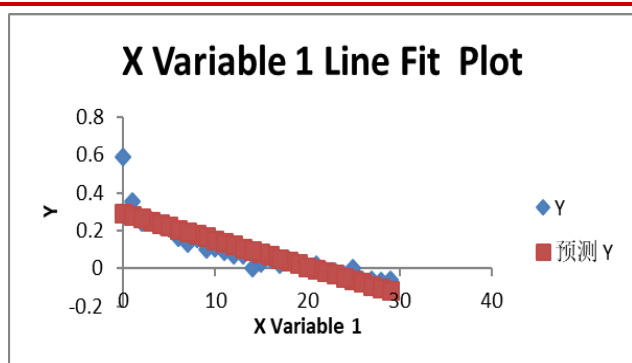
的平滑程度明显提高，2015 年股灾与 2016 年熔断造成的回撤明显降低，特别是 2017 年以来市值因子的持续性大幅回撤被抑制。

从由各分组的序号与期望收益率的相关性，序号与期望收益/回撤比的相关性，所展示的因子分组区分效果来看，除市值因子外，其他因子的分组区分效果明显降低。但通过对测试数据的详细比较，我们认为需要更细致地看待这个问题。

衡量因子分组区分效果是为了展示根据因子分值的高低，是否可以将好的股票组合和坏的股票组合区分开来。我们从（年化）期望收益率和期望收益率/最大回撤比来衡量各分组的好坏。因此，我们采用线性回归的方式获得分组序号与期望收益率，分组序号与期望收益率/最大回撤比的相关系数，以此来检验因子分组效果的单调性。

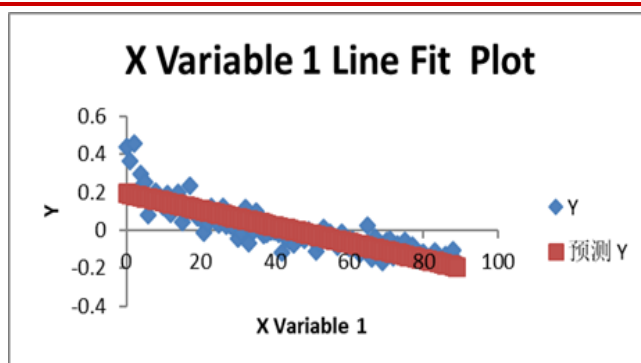
我们首先来看因子分组区分效果较好的市值因子，原因子分组序号与期望收益率的相关系数为 0.8752，经 Hurst 分类树模型改进后为 0.8550。图 28 为原始市值因子的分组序号与期望收益率的回归分析图，图 29 为市值因子经 Hurst 分类树模型改进后的分组序号与期望收益率的回归分析图，两者均展现出很好的分组效果。但由于经 Hurst 分类树模型处理后，分组数由原来的 30 组变为 90 组，线性回归拟合的效果有所减弱，但这并不能说明分组效果减弱。事实上，经 Hurst 分类树模型处理后市值因子的分组效果呈现出一种非线性的单调关系：序号越靠前的分组，期望收益率提高的速度越快，实际分组区分效果好于回归分析所展现的结果，同时也优于原市值因子分组。

图 28 (原始市值因子分组效果回归)



资料来源：东海证券

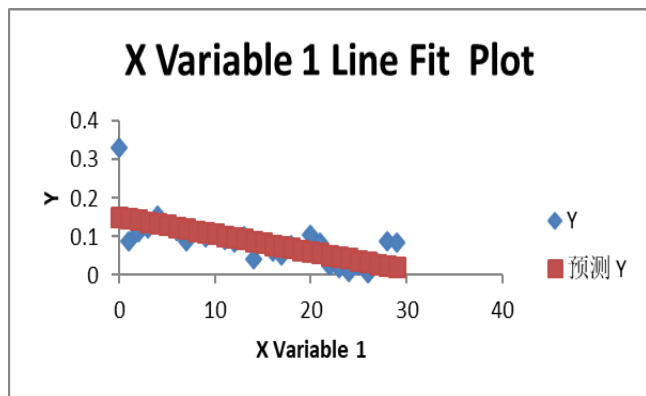
图 29 (Hurst 分类树过滤后的市值因子分组效果回归)



资料来源：东海证券

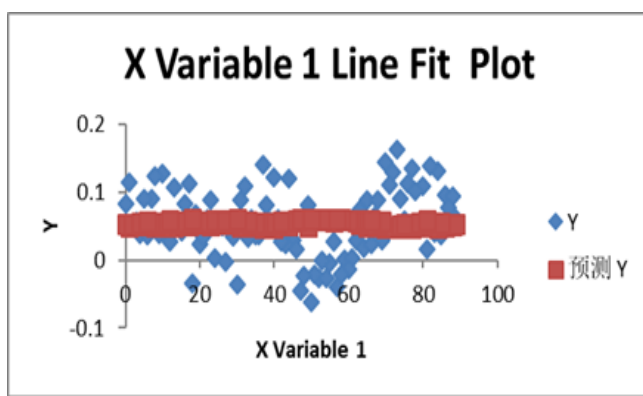
我们再来考察其他因子分组区分效果下降的情况，发现经 Hurst 分类树模型改进后，因子分组序号与期望收益率，分组序号与期望收益率/最大回撤比的散点分布普遍呈现出“W”型。以原始多因子 EPS\_R+ MKTVAL 为例，图 30 为原始市值因子的分组序号与期望收益率的回归分析图，图 31 为市值因子经 Hurst 分类树模型改进后的分组序号与期望收益率的回归分析图。原始因子因分组数较少，呈现出了较好的拟合效果（0.6711）。经 Hurst 分类树模型处理后，拟合效果大幅降低（0.0856）。这一方面是由于分组数增加天然导致的，但另一方面，看似紊乱的散点分布普遍呈现出的“W”型，说明经 Hurst 分类树模型处理后，实际分组效果是提升的。只是由于 Hurst 分类树模型对分组次序的细分调整仅作用于原分组的内部，因此只对一定的子分组区间有效。此时，通过线性回归得到的分组序号与期望收益率的相关系数，不再具有对分组区分情况的解释性。

图 30 (原始 EPS\_R+ MKTVAL 因子回归)



资料来源：东海证券

图 31 (Hurst 分类树过滤的 EPS\_R+ MKTVAL 因子回归)



资料来源：东海证券

## 5. 结论与进一步研究

Hurst 指数是时间序列记忆性特征的测度，就金融时间序列而言这也是一种择时因素。本次研究中，我们着重进行了将 Hurst 指数所体现的择时因素融入到单因子/多因子选股模型中的尝试。同时这也是将分形市场理论运用到实际，以改进原有有效市场理论不足的尝试。通过 2 种基于 Hurst 指数的因子择时模型的构建与测试验证，我们得到以下几个结论：

1. 将 R/S 分析法与 Hurst 指数应用到单因子/多因子选股模型中具有重要意义，合理运用可以有效改善选股模型的效果。但因其本身不包含价格运动方向性的信息，因此不适于作为独立因子应用。
2. Hurst 指数对单因子/多因子选股模型的改进效果，一方面取决于原始因子的有效性，对效果较差的原始因子改进不明显，另一方面取决于 Hurst 指数与原始因子的组合方式，不合理的组合方式甚至会弱化原始因子的有效性。
3. 传统以因子评分值加权的多因子组合方式在一定条件下会丢失原有因子的信息，使因子效果不显著。采用分类决策树形式对因子进行组合，可以避免对参数形式的假定，具有很好的研究前景。
4. 通过 Hurst 分类树模型可以有效降低市值因子的回撤，提高期望收益率/最大回撤比，平滑收益曲线，并改进因子分组的区分效果。
5. 以线性回归方式测量因子分组的区分效果并不全面，在一定条件下会失去解释性，非线性的方法值得尝试。

本次研究是我们对实际应用 Hurst 指数的初步研究，通过本次研究的结果，我们认为寻找 Hurst 指数与其他因子的相互作用关系是成功应用的关键。然而在已有的理论基础上，我们并没有两者关系先验的函数形式。

在 Hurst 因子分类树模型中，我们采用了决策树的形式，这使得我们脱离了对参数假设和先验函数形式的依赖。虽然在本模型中我们的分支节点是先验设定的，但这是未来真正利用决策树理论的一个良好开端。在进一步的研究中，我们将运用监督学习的智能算法实现规则的学习，寻找 Hurst 指数与其他因子的相互作用关系。同时这也是我们未来构造其他多因子模型的基本方法框架。



## 分析师简介:

丁竞渊, 计算机应用技术博士, 数学与应用数学学士, 东海证券研究所高级研究员, 证券从业10年, 具有丰富的量化投资研究、量化平台开发与金融数据分析经验。

## 附注:

### 一、市场指数评级

看多——未来6个月内上证综指上升幅度达到或超过20%

看平——未来6个月内上证综指波动幅度在-20%—20%之间

看空——未来6个月内上证综指下跌幅度达到或超过20%

### 二、行业指数评级

超配——未来6个月内行业指数相对强于上证指数达到或超过10%

标配——未来6个月内行业指数相对上证指数在-10%—10%之间

低配——未来6个月内行业指数相对弱于上证指数达到或超过10%

### 三、公司股票评级

买入——未来6个月内股价相对强于上证指数达到或超过15%

增持——未来6个月内股价相对强于上证指数在5%—15%之间

中性——未来6个月内股价相对上证指数在-5%—5%之间

减持——未来6个月内股价相对弱于上证指数5%—15%之间

卖出——未来6个月内股价相对弱于上证指数达到或超过15%

### 四、风险提示

本报告所载的全部内容只提供给客户做参考之用, 并不构成对客户的投资建议, 并非作为买卖、认购证券或其它金融工具的邀请或保证, 建议客户如有任何疑问应当咨询独立财务顾问并独自进行投资判断。

### 五、免责条款

本报告基于本公司研究所及研究人员认为可信的公开资料或实地调研的资料, 但对这些信息的真实性、准确性和完整性不做任何保证。本报告仅反映研究员个人出具本报告当时的分析和判断, 并不代表东海证券股份有限公司, 或任何其附属或联营公司的立场, 本公司可能发表其他与本报告所载资料不一致及有不同结论的报告。本报告可能因时间等因素的变化而变化从而导致与事实不完全一致, 敬请关注本公司就同一主题所出具的相关后续研究报告及评论文章。在法律允许的情况下, 本公司的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易, 并可能为这些公司正在提供或争取提供多种金融服务, 本公司的关联机构或个人可能在本报告公开发布之间已经了解或使用其中的信息。

分析师承诺“本人及直系亲属与本报告中涉及的内容不存在利益关系”。本报告仅供“东海证券股份有限公司”客户、员工及经本公司许可的机构与个人阅读。

本报告版权归“东海证券股份有限公司”所有, 未经本公司书面授权, 任何人不得对本报告进行任何形式的翻版、复制、刊登、发表或者引用。

### 六、资格说明

东海证券股份有限公司是经中国证监会核准的合法证券经营机构, 已经具备证券投资咨询业务资格。我们欢迎社会监督并提醒广大投资者, 参与证券相关活动应当审慎选择具有相当资质的证券经营机构, 注意防范非法证券活动。

## 上海 东海证券研究所

地址: 上海市浦东新区东方路1928号 东海证券大厦

网址: [Http://www.longone.com.cn](http://www.longone.com.cn)

电话: (8621) 20333619

传真: (8621) 50585608

邮编: 200215

## 北京 东海证券研究所

地址: 北京市西三环北路87号国际财经中心D座15F

网址: [Http://www.longone.com.cn](http://www.longone.com.cn)

电话: (8610) 66216231

传真: (8610) 59707100

邮编: 100089