

使用 Thompson Sampling 算法的策略混合模型

——多因子模型研究系列之六

分析师：宋旻

SAC NO: S1150517100002

2018 年 12 月 28 日

证券分析师

宋旻

022-28451131

18222076300

songyang@bhzq.com

相关研究报告

《多因子模型研究之一：单因子测试》20171011

《多因子模型研究之二：收益预测模型》20171229

《多因子模型研究之三：风险模型与组合优化》20180416

《随机森林多因子模型与传统多因子模型的选股风格对比——多因子模型研究系列之四》
20180726《使用 bandit learning 算法的多因子模型——多因子模型研究系列之五》
20180925

核心观点：

- 本篇报告中，我们介绍了 **Thompson Sampling** 算法，并将其应用于改进的多因子模型。Thompson Sampling 算法是在线学习算法的一种，2017 年年末，市值因子失效，导致大多数传统多因子模型回撤。我们想要尝试使用在线学习算法，构建可以适应这种风格转变的多因子模型
- 针对沪深 300 和中证 500 成分股，选取估值、盈利、成长、动量、反转、波动率、流动性、市值八大类因子建立多因子模型。使用传统多因子模型估计投资组合的未来收益，使用分层抽样方法剔除了行业的影响，单独检测市值因子在历史不同时间段的表现。我们发现，在过去较长的一段时间内，小市值组合一直占据着较为稳定的优势地位，直到 2017 年，市值因子风格才完全反转。
- 我们检测了买入持有、定期调整资产比例、Greedy、Epsilon-Greedy 和 Thompson Sampling 五种策略应用于抽样多因子模型的历史表现。Thompson Sampling 在各项指标中都处于领先地位，尤其是在 2017 年没有收到市场风格转换影响，表现出了较好的适应性，实现了持续盈利。
- 我们发现 **Thompson Sampling** 更适用于具有周期性的资产投资将会更具优势。我们又构建了股票和债券的混合模型，更明显的体现出了该方法在周期轮动情况下的适应能力。
- 我们认为，**Thompson Sampling** 在现实中的应用可以与市场判断相结合。当判断市场未来更多体现震荡行情时，可以使用这种方法来平滑风险；当判断市场未来更多体现趋势行情时，可以使用其他更偏重动量的策略，以取得更高的短期投资收益。在长期投资中，总的来说 Thompson Sampling 方法会优于其他方法，因为没有任何趋势可以一直持续，市场的周期性已经被很多研究论证过了。
- 未来，我们会继续研究更多的在线学习模型，探索其适用范围、运行机制、收益来源以及成果的延续能力，并将其应用扩展到更多领域，如行业轮动、资产配置等。
- 风险提示：随着市场环境变化，模型存在失效风险。

目 录

1. 概述	4
2. 理论简介	4
3. 模型建立	5
3.1 模型建立	5
3.2 对照组构建	7
3.3 算法介绍	8
4. 回测结果	10
4.1 沪深 300 回测结果	10
4.2 中证 500 回测结果	12
4.3 股债混合模型回测结果	14
5. 总结与未来研究方向展望	16

表目录

表 1: 模型入选因子汇总	6
表 2: 模型历史回测结果 (沪深 300)	11
表 3: 模型分年度收益情况 (沪深 300)	12
表 4: 模型历史回测结果 (中证 500)	13
表 5: 模型分年度收益情况 (中证 500)	14
表 6: 模型历史回测结果 (股债混合)	15

图目录

图 1: 沪深 300 市值因子回测收益曲线 (小市值/大市值)	7
图 2: 中证 500 市值因子回测收益曲线 (小市值/大市值)	7
图 3: Beta 分布概率密度曲线图	9
图 4: 选股模型相对沪深 300 超额收益曲线	11
图 5: 沪深 300 选股模型回测净值曲线	12
图 6: 选股模型相对中证 500 超额收益曲线	13
图 7: 中证 500 选股模型回测净值曲线	14
图 8: 股债混合模型超额收益曲线	16
图 9: 股债混合模型回测收益曲线	16

1. 概述

传统的多因子模型理论来源于 Markowitz 的风险收益模型，其目的在于使用历史因子数据，预测投资组合下期收益与风险，找到横截面上风险收益比最小的组合。可想而知，在这一过程中，对于未来预测的准确度十分关键。

在此之前，对于多因子模型的改进，大部分目标都在于提高预测的准确率上，不论是寻找新的因子，或者改进模型细节，都是为了使预测更加准确。然而，想要使用历史数据预测未来并不容易，历史并非永远重复，即使重复，也很难准确把握关键时点。例如 2017 年以来，小市值因子失效，很多多因子模型来不及适应，产生较大回撤。针对这种情况，一些模型引入了因子择时策略，但是因子择时本质上依然是在预测未来，所以依然有出错的可能。也有人为了规避风险，对于市值因子做完全中性处理。然而市值因子在大部分时间段内，alpha 收益是非常明显的，如果不是对于回撤有严格要求的模型，放弃市值因子的 alpha 暴露，会导致较大的收益损失，有可能得不偿失。

寻找一种新的选股方式变得十分必要。在 Li 和 Hoi 的论文《Online Portfolio Selection: A Survey》中，作者为我们介绍了一种新的选股方法：**在线学习选股**。在线学习算法中，我们处理一个时间序列的决策问题，每一期的选择，根据目前已知的信息实时反馈更新算法。我们把未来看成一个概率分布，不去预测单次下注的胜负，而是关注于在不确定的情况下，如何分配手中的筹码，使长期投资总收益期望的几何平均值最大。这也是在线学习（Online Learning）模型与传统多因子模型的最大不同。

在线学习的算法很多，在上一篇报告《使用 Bandit Learning 算法的多因子模型——多因子模型研究系列之五》中，我们选用多臂赌博机算法（Multi-armed Bandit Learning）构建选股模型，取得了较好的效果。本篇报告中，我们介绍另一种在线学习算法：Thompson Sampling，并将其应用在策略混合上。本文参考了 Shen 的论文《Portfolio Blending via Thompson Sampling》，作者将 Thompson Sampling 算法应用于美国股市，取得了不错的投资结果。本文仿照该方法，将 Thompson Sampling 算法应用于 A 股，回测表明，在投资周期性质的标的时，Thompson Sampling 算法表现出了较好的适应性。

2. 理论简介

在线学习的原理可以用一个经典例子：多臂老虎机（Multi-armed Bandit）问题来

阐述。上一篇报告中介绍的 UCB 算法和这篇报告中介绍的 Thompson Sampling 算法都是该问题的一个分支。多臂老虎机问题的提出和研究最早可以追述到上世纪三十年代，该问题可以使用一种较为形象的场景进行解释。

假设你进入了一家赌场，这家赌场的大厅里有 n 个老虎机，当你往老虎机中投入一枚硬币，老虎机有一定概率掉落奖励，每个老虎机掉落奖励的概率不同且未知，那么在现有硬币有限的情况下，使用什么策略才能使你最后所得的总体奖励最大？

在尝试了几次后，你对每个赌博机的掉落概率有了一个初步的估计，但是受尝试次数的限制，这个估计可能是不准确的，且未来该概率还有可能改变。接下来，是选择一直坚持目前已知的最好选择（exploitation，守成），还是继续尝试，以达到更准确的估计，找到奖励概率更高的老虎机（exploration，探索），这便是多臂老虎机理论解决的问题。通过一系列在线学习反馈算法，平衡守成与探索的比例，以达到最后总体收益的最大化。

在量化投资领域，我们把每个老虎机的“臂”理解为不同的资产或资产组合，每种资产的回报率并非一成不变的，我们假设它和历史 M 期的表现有关，在线学习算法可以使程序自我学习，根据数据的改变实时更新算法。在投资环节的每个时刻，根据目前已知的信息做出决策，目的并不是最大化即时收益，而是最大化远期收益。

3. 模型建立

3.1 分层抽样多因子模型的建立

2017 年年末，市值因子失效，导致大多数传统多因子模型回撤。我们想要尝试使用在线学习算法，构建可以适应这种风格转变的多因子模型。

首先，针对沪深 300/中证 500 成分股，以月度调仓的频率构建投资组合。回测数据为 2010 年-2018 年 11 月。

我们选取估值、盈利、成长、动量、反转、波动率、流动性、市值八大类因子构建多因子模型（见表 1），经过缺失值处理、去极值、标准化、中性化等前期准备步骤，采用半衰期加权移动平均方法构建收益预测模型。模型具体构建方法与细节可参考我们之前发表的报告：《多因子模型研究之一：单因子测试》、《多因子模

型研究之二：收益预测模型》和《多因子模型研究之三：风险模型与组合优化》。

接着，我们以 29 个中信一级行业为蓝本，将非银行金融行业拆分为证券、保险、信托及其他三个子行业，采用共 31 个行业的分类，**构建行业中性分层抽样指数增强策略**。将基准指数按行业划分为 31 个子集，每个子集中用市值因子将股票划分为数目相等的两组：大市值组与小市值组，在每个组中各自选取预期收益最高的两只股票，令其均分所在行业在基准指数中的权重。

这样，我们就得到了各包含约 62 只股票的大市值组与小市值组，每组股票中的行业权重与基准指数相同。从而剔除了行业的影响，单独检测市值因子在历史不同时间段的表现。

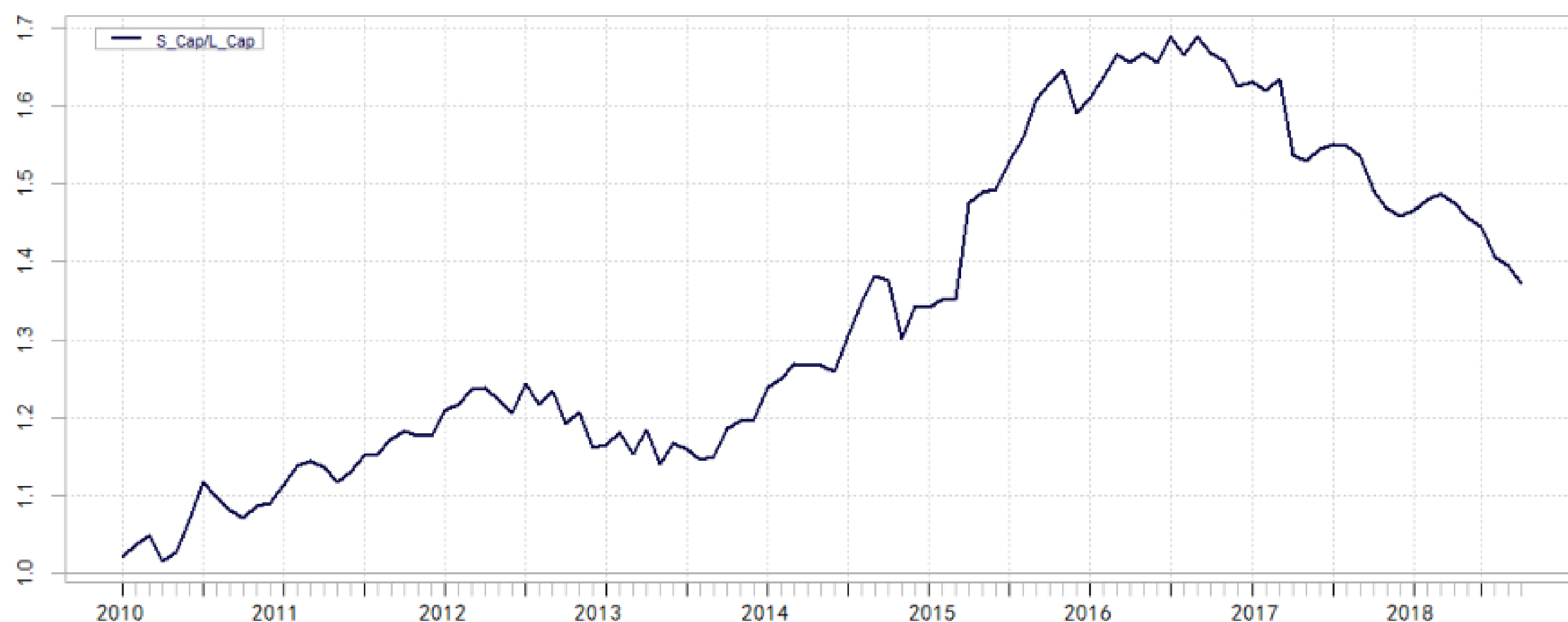
表 1：模型入选因子汇总

因子大类	最终入选因子
估值因子	BP、扣非 EP_ttm
盈利因子	单季度 ROE
成长因子	单季度营业收入增长率、单季度归母净利润增长率
动量因子	指数加权一年收益率/上月收益率
反转因子	上月收益率
波动率因子	月度波动率、季度波动率、年度波动率
流动性因子	月度换手率、季度换手率、年度换手率
市值因子	流通市值对数

数据来源：渤海证券研究所

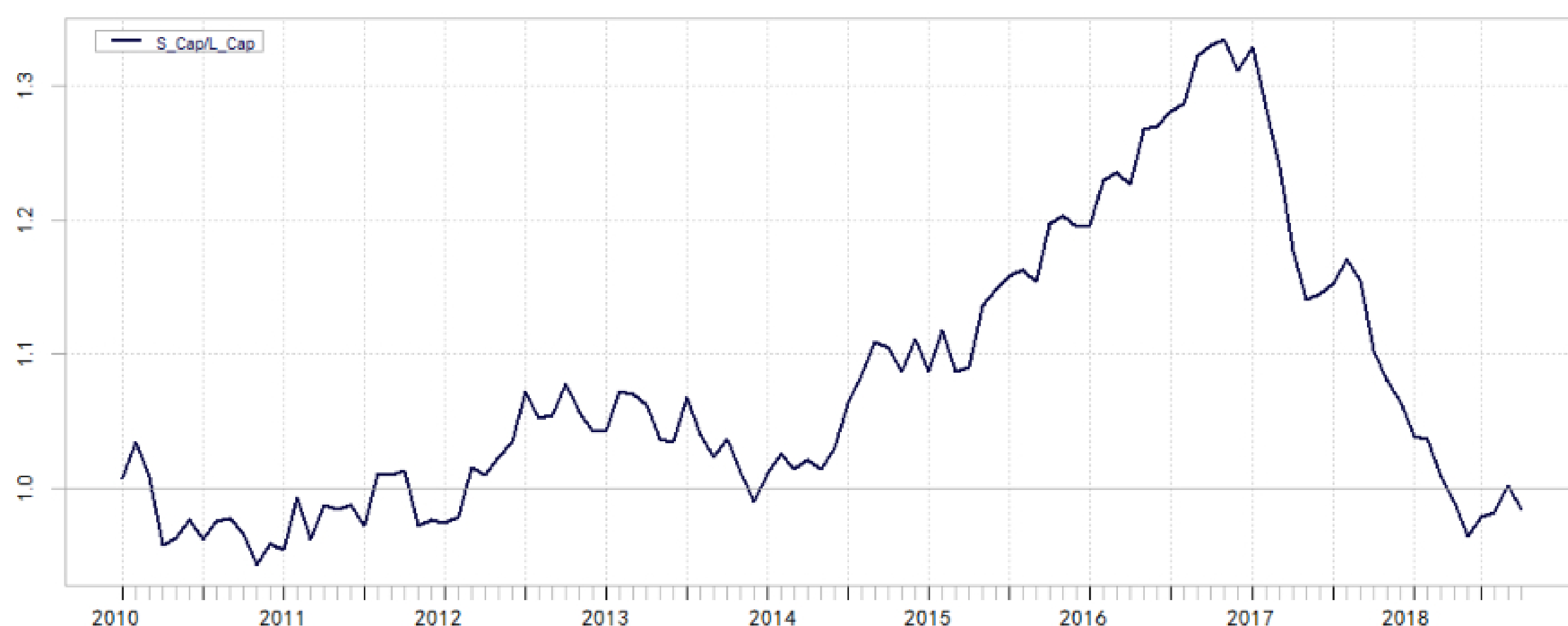
通过回测可以看出，在过去较长的一段时间内，小市值组合一直占据着较为稳定的优势地位，虽然中间有一些小段的反复，但总体来说直到 2017 年，市值因子风格才完全反转。由此可见，模型能否适应因子的风格变化，主要看其能否在 2017 年时迅速适应市场转变。

图 1: 沪深 300 市值因子回测收益曲线 (小市值/大市值)



资料来源: Wind, 渤海证券研究所

图 2: 中证 500 市值因子回测收益曲线 (小市值/大市值)



资料来源: Wind, 渤海证券研究所

3.2 对照组构建

关于业绩基准, 虽然模型是在指数的基础上进行增强, 但因为其本质是大市值组合和小市值组合的混合与轮动, 单纯使用指数作为参照组显得不太合适。在这里, 我们在原始指数之外又引入了两组新的对照组, 以客观反映模型的收益能力。

3.2.1 买入持有组 (Buy and Hold, BAH)

在策略之初以相等的权重买入大市值组合与小市值组合两组资产, 在之后的持有期间内不再进行其他操作, 保持这种组合。这也是最直接的对照组构建方法。

3.2.2 定期调整资产比例组 (Constant Rebalanced Portfolios, CRP)

请务必阅读正文之后的免责条款部分

在策略之初以相等的权重买入大市值组合与小市值组合两组资产，之后在每个调仓期调整资产配比，使得两种资产的总值相等。有研究表明，相比于买入持有策略，定期调整资产比例可以显著降低投资组合的不稳定性，从而更直接的体现两个组合的平均收益。

3.3 算法介绍

我们测试了三种在线学习算法的表现，分别为 Greedy Algorithm、Epsilon-Greedy Algorithm 和 Thomson Sampling，我们首先把每种算法推广到投资领域，然后根据算法构建投资策略。

3.3.1 Greedy Algorithm

又称贪婪算法或贪心算法，其核心思想为计算每一个臂的回报，然后选择回报最大的臂进行操作。在具体实施时，我们设置 M ($M=24$) 期历史数据为观测窗口，计算两种组合的历史收益，并在下一期选择收益高的一组。贪心算法的问题在于没有完全探索其它奖励概率的可能性，很容易丢掉最优解。

3.3.2 Epsilon-Greedy Algorithm

该算法为贪心算法的改进。在原始算法中，首先选定一个较小的常数 ϵ ，让程序每次生成一个 0 到 1 之间的随机数，如果随机数落在 $(0, \epsilon)$ 区间内，则随机选择一个臂，如果随机数落在 $(\epsilon, 1)$ 区间内，则选择当前情况下回报率最大的臂。即每次操作有 ϵ 的概率随机选择，剩下 $1-\epsilon$ 的概率依然选择预期回报率最大的臂。在离散空间里，这是一个不确定的算法，因为并不能确定每次生成的随机数落在哪个空间。但是当算法被应用于投资领域，我们面对的问题变成了连续空间，即每次可以直接赋予预期回报率较大组合 $1-\epsilon$ 的权重（我们设 $\epsilon = 0.3$ ），赋予另外一组投资组合 ϵ 的权重。这样，最终的投资组合表达式与离散空间下的投资期望表达式是一致的，算法也变成了一个确定性的算法。

这种算法相对于贪心算法保持了一定对于探索的开放性，但回测效果依然不令人满意，究其原因，是因为这种算法不能根据每次的试验结果进行自我调整，最优组合和其他组合的权重一直是恒定的，这样会赋予了掉落概率低的臂过大权重，从而影响最终的整体收益。

3.3.3 Thompson Sampling

在 Thompson Sampling 算法中，我们将多臂赌博机每一个臂的回报概率看作一个 Beta 分布。Beta 分布是一个定义在 $[0, 1]$ 区间上的连续概率分布族，它有两个正值参数，称为形状参数，一般用 α 和 β 表示。在贝叶斯推断中，Beta 分布是 Bernoulli 分布、二项式分布、负二项式分布和几何分布的共轭先验分布。

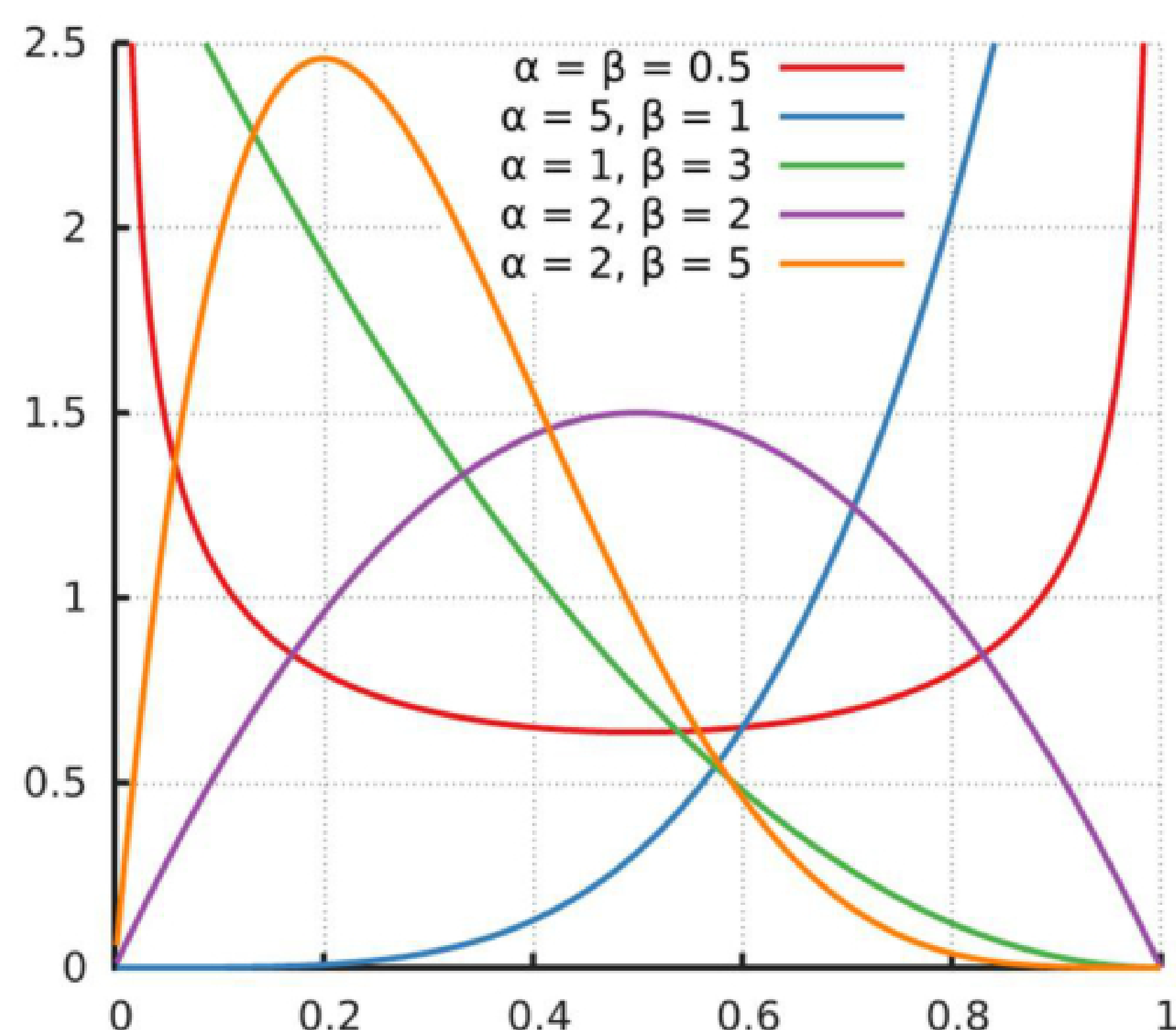
Beta 分布的概率密度函数形式如下：

$$f(x; \alpha, \beta) = \text{constant} \cdot x^{\alpha-1} (1-x)^{\beta-1}$$

Beta 分布可以看作一个概率 p 的概率分布，当某项实验成功的具体概率未知时，它可以给出了所有概率出现的可能性大小。

Beta 分布有两个参数， α 和 β ，一般记为 $\text{Beta}(\alpha, \beta)$ 。其中参数 α 可以看成实验成功的次数，参数 β 可以看成实验失败的次数。当 α 和 β 等于不同值时概率密度函数如下图：

图 3: Beta 分布概率密度曲线图



资料来源: Wind, 渤海证券研究所

从图中可以看出， $\alpha + \beta$ 的值越大，分布曲线越窄。这代表当实验次数增加时，我们对于成功概率预测的确定性会增强；

而 $\alpha / (\alpha + \beta)$ 的值是 beta 分布的均值（期望值），同时它也代表了之前实验中成功次数所占的比例。它的值越大，beta 分布的中心越靠近 1，实验成功的概率越大，反之，则 beta 分布的中心越靠近越靠近 0，实验成功的概率越小。

回到 Thompson Sampling 算法,我们将每一个臂的回报概率看做一个 Beta 分布,通过不断测试调整每个 Beta 分布的参数 α 和 β 的值,来确定最后的权重。

在离散的空间,我们在每次试验中选中一个臂,有收益则该臂的 Beta 分布的参数 α 增加 1,否则该臂的 Beta 分布的参数 β 增加 1。然后使每个臂现有的 beta 分布各产生一个随机数,下一次实验则选择所有臂中产生的随机数中最大的那个臂。

在离散空间中,这依然是一个不确定的算法。但是扩展到连续空间后,只需在每一期计算之前 M 期历史收益情况,当大市值组合收益>小市值组合收益时, α 值增加 1;当大市值组合收益<小市值组合收益时, β 值增加 1。再将当期大市值组合的权重设为 $\frac{\alpha}{\alpha+\beta}$;小市值组合的权重设为 $\frac{\beta}{\alpha+\beta}$,我们便得到了在该设定下两个组合混合的最优比例。和 Epsilon-Greedy Algorithm 的情况类似,当扩展到一个连续空间时,这也变成了一个确定性的算法。和前面两种算法相比, **Thompson Sampling** 在自我调整的及时性上有更大的优势,因此在下文回测中表现也最好。

4. 回测结果

我们分别针对沪深 300 和中证 500 指数,测试了买入持有组 (BAH)、定期调整资产比例组 (CRP)、Greedy 算法、Epsilon-Greedy 算法和 Thompson Sampling 算法下的大市值组合与小市值组合混合结果。

4.1 沪深 300 回测结果

针对沪深 300 指数的回测情况, Thompson Sampling 方法在五种混合策略中表现最佳,取得了 7.02%的年化收益,信息比率 1.41,相对指数月度胜率 66.67%,均为最高。

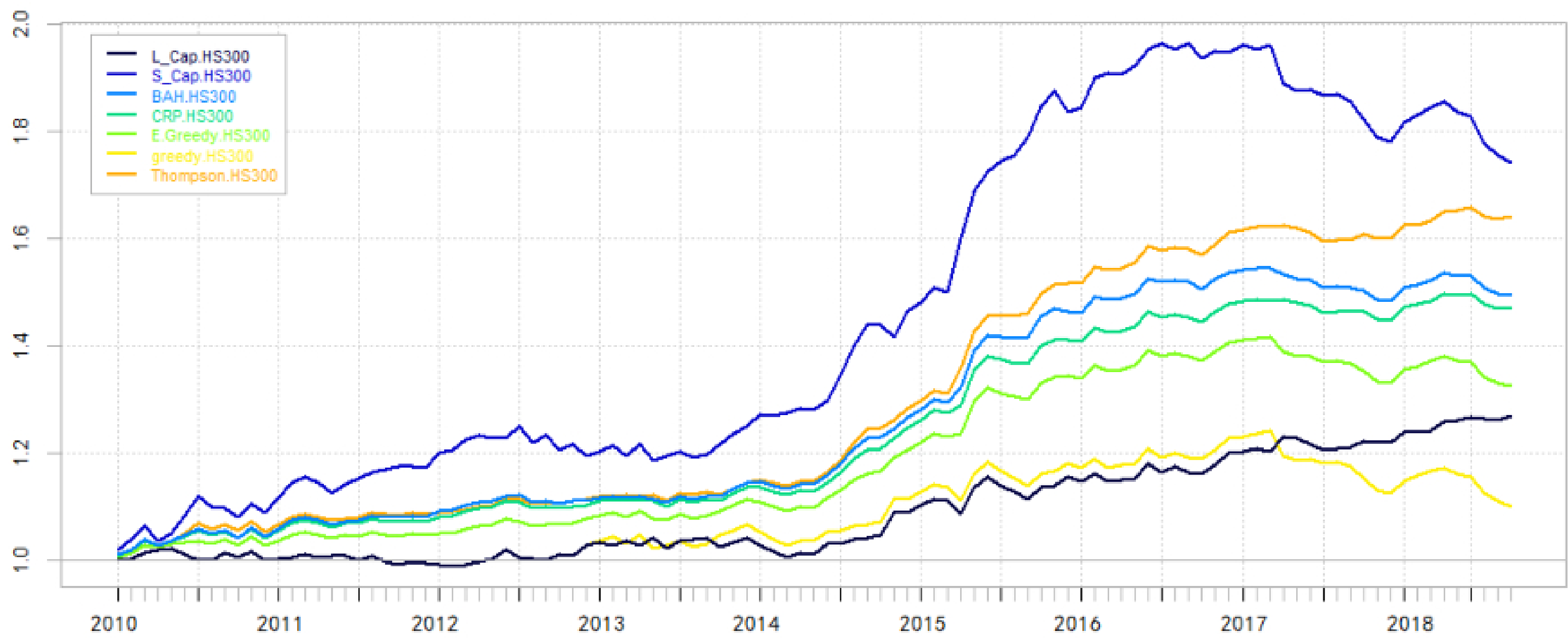
表 2：模型历史回测结果（沪深 300）

	L_Cap	S_Cap	BAH	CRP	Greedy	E-Greedy	Thompson	HS300
累计收益	49.79%	105.63%	76.54%	73.09%	30.02%	65.55%	81.10%	18.15%
年化收益	4.73%	8.59%	6.71%	6.47%	3.05%	5.93%	7.02%	1.92%
波动率	24.78%	25.43%	24.82%	24.78%	25.18%	24.73%	24.72%	24.16%
最大回撤	36.33%	30.43%	33.34%	34.14%	36.33%	34.49%	32.04%	39.62%
夏普比率	0.19	0.34	0.27	0.26	0.12	0.24	0.28	0.08
信息比率	0.70	1.10	1.37	1.39	0.25	1.25	1.41	--
胜率	58.10%	63.81%	60.00%	60.95%	55.24%	61.90%	66.67%	--

数据来源：渤海证券研究所、Wind

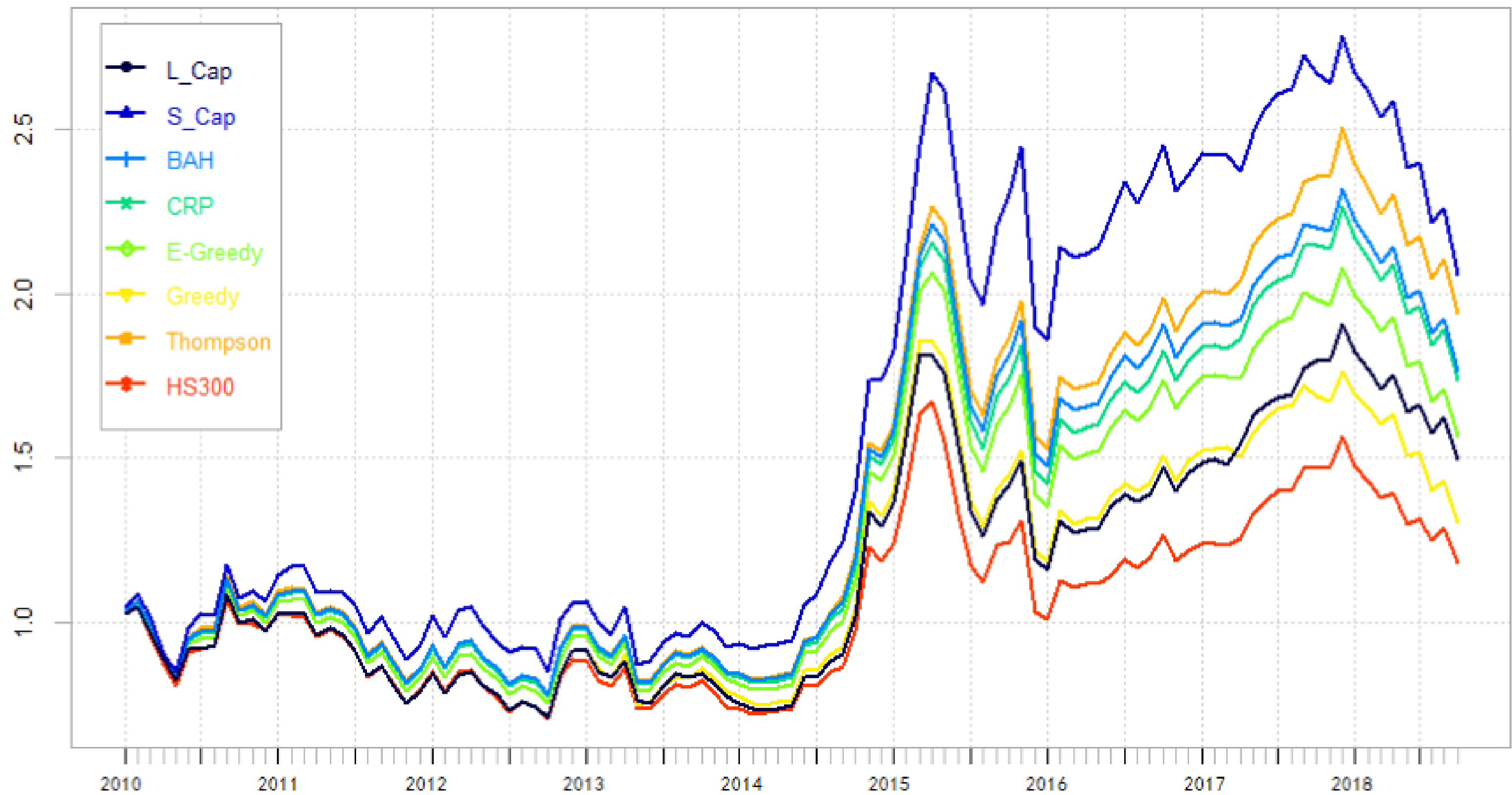
通过观察组合相对沪深 300 的超额收益曲线，可以发现小市值组合在 2017 年前大幅跑赢大市值组合，而从 2017 年开始，市场风格发生转变，小市值组合产生了较大回撤，导致大部分组合表现欠佳，其中 Thompson Sampling 表现出了较好的适应性，没有受到市场风格转换影响，实现了持续盈利。

图 4：选股模型相对沪深 300 超额收益曲线



资料来源：Wind，渤海证券研究所

图 5：沪深 300 选股模型回测净值曲线



资料来源：Wind，渤海证券研究所

在组合分年度收益统计中，可以发现，在大部分年度中，Thompson Sampling 保持了领先的收益，而在大小市值分化较为明显的年度，如 2010、2014、2015、2017 年，Thompson Sampling 的优势尤其明显，证明该方法在风格不确定的市场中表现更灵活，而在其余年度，也能维持一定的收益。

表 3：模型分年度收益情况（沪深 300）

	L_Cap	S_Cap	BAH	CRP	Greedy	E-Greedy	Thompson	HS300
2010	-2.7%	6.1%	1.7%	1.4%	-2.7%	-0.1%	2.6%	-2.6%
2011	-18.8%	-12.3%	-15.6%	-15.9%	-18.8%	-16.8%	-16.0%	-18.4%
2012	15.2%	13.6%	14.5%	14.7%	15.2%	14.8%	14.5%	11.5%
2013	-15.1%	-12.5%	-13.8%	-13.9%	-13.2%	-13.5%	-14.0%	-16.3%
2014	67.3%	87.5%	77.3%	75.7%	67.3%	73.3%	79.3%	60.1%
2015	-7.8%	9.1%	0.4%	-1.5%	-7.8%	-2.9%	2.8%	-13.0%
2016	22.2%	24.8%	23.5%	23.1%	22.2%	23.0%	25.0%	17.7%
2017	31.2%	17.9%	24.4%	26.3%	18.3%	22.0%	28.1%	28.9%
2018	-21.6%	-26.2%	-23.9%	-23.3%	-26.2%	-24.8%	-22.6%	-24.5%

数据来源：渤海证券研究所、Wind

4.2 中证 500 回测结果

中证 500 的回测情况与沪深 300 类似，Thompson Sampling 方法仍在多项指标中表现最佳，取得了 12.43%的年化收益，信息比率 2.24，相对指数月度胜率 75.24%。仅在相对指数月度胜率一项输给了买入持有策略，而考虑到买入持有策

略在 17 年不尽如人意的表现，可以认定其高胜率具有一定偶然性，毕竟，在中证 500 指数中，小市值股票前期的优势更明显，买入持有策略不调整投资比例，滚动一定时间后小市值组合会占到一个非常高的比重，此时如果小市值行情持续，买入持有策略会受益较多，但一旦风格转变，也面临着更大的风险。

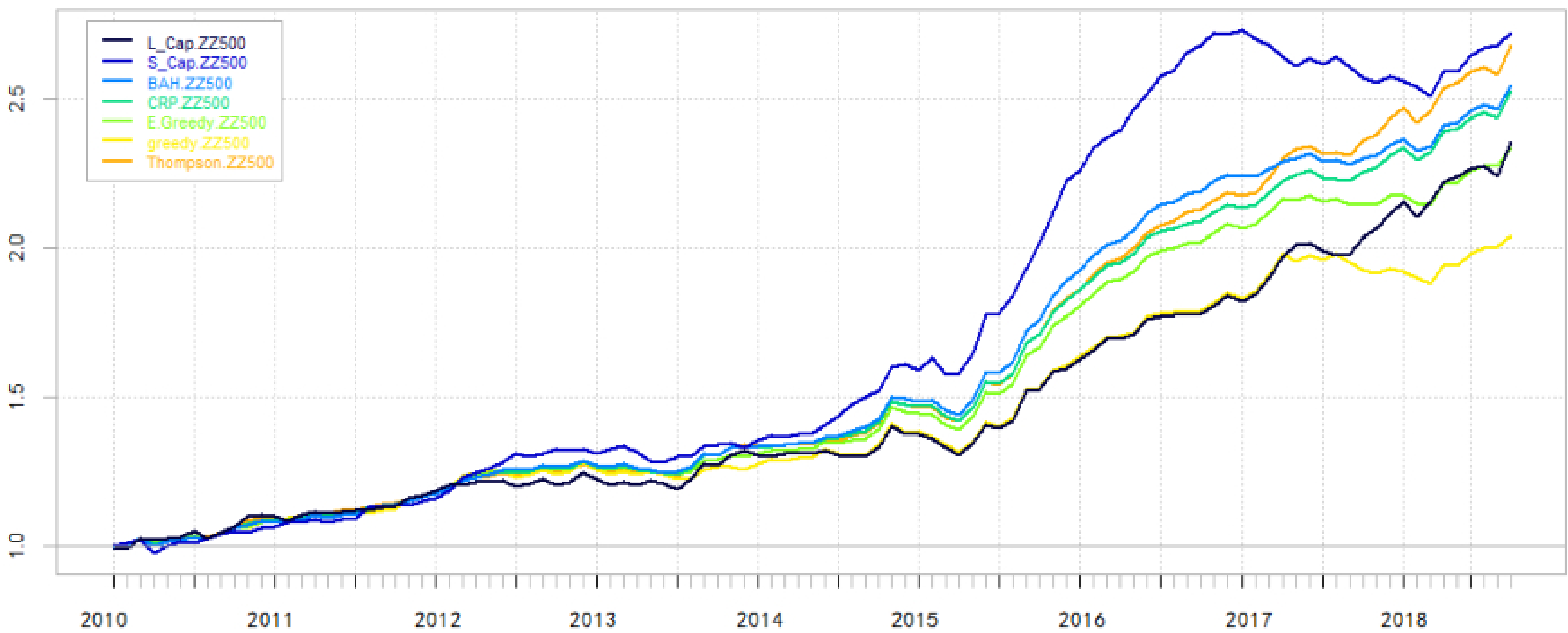
表 4：模型历史回测结果（中证 500）

	L_Cap	S_Cap	BAH	CRP	Greedy	E-Greedy	Thompson	HS300
累计收益	145.81%	184.00%	165.77%	163.53%	143.71%	113.26%	178.80%	4.34%
年化收益	10.83%	12.67%	11.82%	11.71%	10.72%	9.04%	12.43%	0.49%
波动率	27.10%	27.89%	27.25%	27.18%	27.33%	27.58%	27.16%	27.49%
最大回撤	38.88%	32.60%	34.77%	35.42%	35.28%	36.06%	36.16%	55.55%
夏普比率	0.40	0.45	0.43	0.43	0.39	0.33	0.46	0.02
信息比率	1.65	1.94	2.23	2.19	2.00	1.45	2.24	--
胜率	70.48%	69.52%	79.05%	74.29%	73.33%	66.67%	75.24%	--

数据来源：渤海证券研究所、Wind

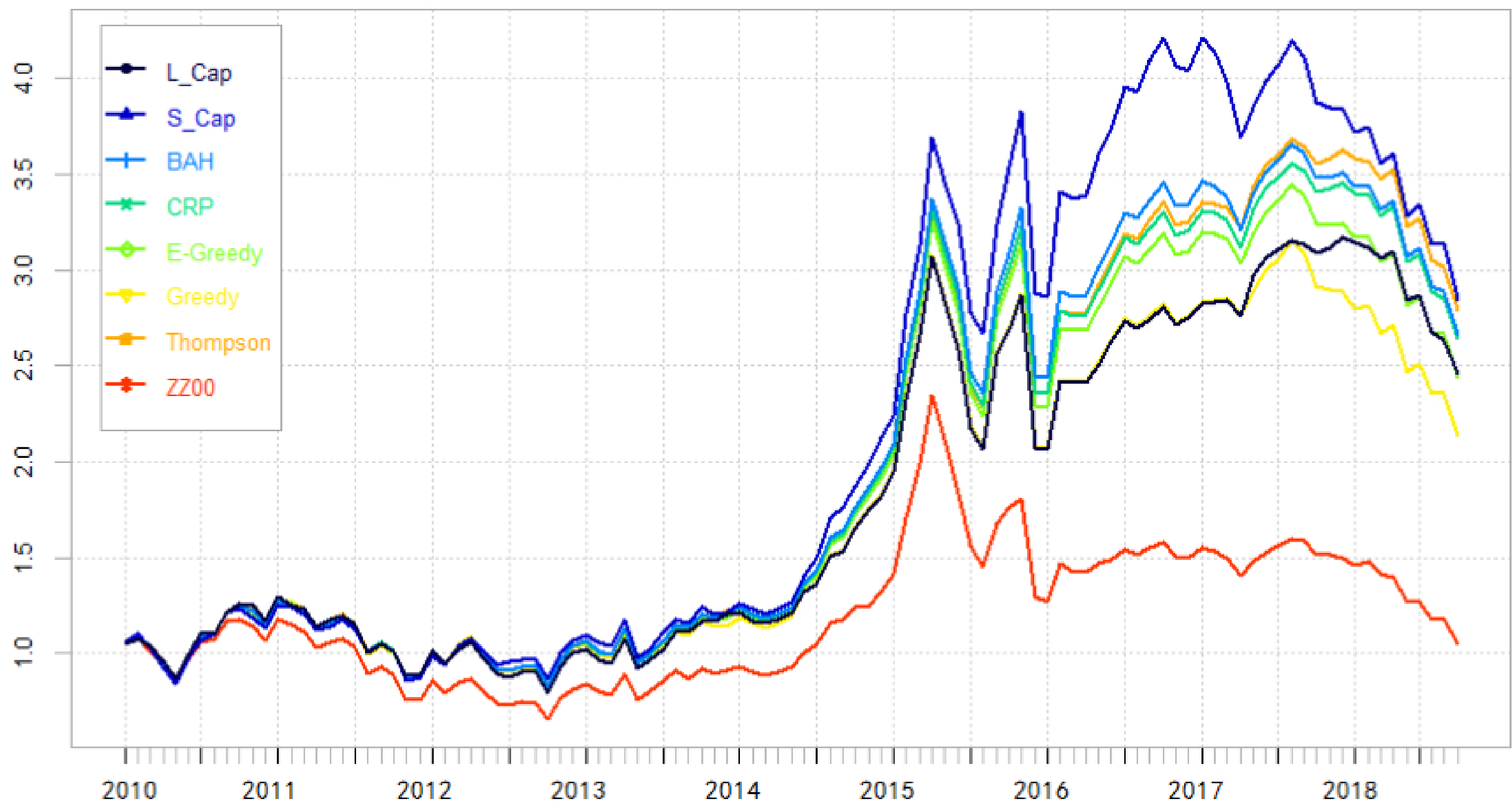
通过观察组合超额收益曲线，可以发现小市值组合在 2017 前年大幅跑赢大市值组合，而从 2017 年开始，市场风格发生转变，小市值组合产生了较大回撤，18 年后，又重新占据上风。在这一波风格转变中，大部分组合表现欠佳，只有 Thompson Sampling 表现出了较好的适应性，没有受到市场风格转换影响，实现了持续盈利。

图 6：选股模型相对中证 500 超额收益曲线



资料来源：Wind，渤海证券研究所

图 7：中证 500 选股模型回测净值曲线



资料来源：Wind，渤海证券研究所

在组合分年度收益统计中,可以发现,在大小市值风格轮动表现较为明显的年度,如 2010、2016、2017 年, Thompson Sampling 的优势明显,而在其余年度,当某一种风格过于强势的时候, Thompson Sampling 可能暂时跑输买入持有策略,其原因上一段已经论述,但总的来说, Thompson Sampling 可以在各种类型的市场中都维持较好的收益。

表 5：模型分年度收益情况（中证 500）

	L_Cap	S_Cap	BAH	CRP	Greedy	E-Greedy	Thompson	HS300
2010	17.1%	12.8%	15.0%	15.4%	14.9%	14.5%	16.2%	6.3%
2011	-24.1%	-22.6%	-23.3%	-23.4%	-23.2%	-23.0%	-23.7%	-28.6%
2012	13.0%	22.4%	17.7%	17.0%	17.4%	17.0%	16.1%	6.4%
2013	19.4%	13.2%	16.3%	16.9%	14.0%	10.5%	18.4%	12.6%
2014	51.5%	75.6%	63.2%	60.9%	62.0%	60.2%	59.2%	45.5%
2015	13.7%	35.3%	24.2%	21.4%	19.9%	13.7%	21.8%	-2.0%
2016	32.9%	40.5%	36.7%	35.5%	35.1%	32.9%	37.5%	15.2%
2017	15.2%	-4.9%	4.8%	7.9%	4.7%	4.4%	11.7%	0.2%
2018	-22.3%	-26.0%	-24.2%	-23.7%	-24.9%	-26.0%	-23.1%	-30.2%

数据来源：渤海证券研究所、Wind

4.3 股债混合模型回测结果

在之前关于市值因子的测试中,我们大概看出了 Thompson Sampling 方法在风

格轮换中的适应性,但因为这段回测期市值因子的大规模风格轮换其实只有一次,所以并不能充分发挥 Thompson Sampling 的优势。我们猜测,相对于趋势类资产, Thompson Sampling 方法更适合应用于具有周期性的资产投资。于是我们又将其应用于股票和债券的混合模型。

我们使用沪深 300 指数和中证国债指数作为投资标的,提取 2007 年以来的日度收益,取 60 日为移动窗口,重复上文的建模步骤,最后得到在线学习算法关于股债混合模型的回测结果。可以看出, Thompson Sampling 各项指标均大幅领先。

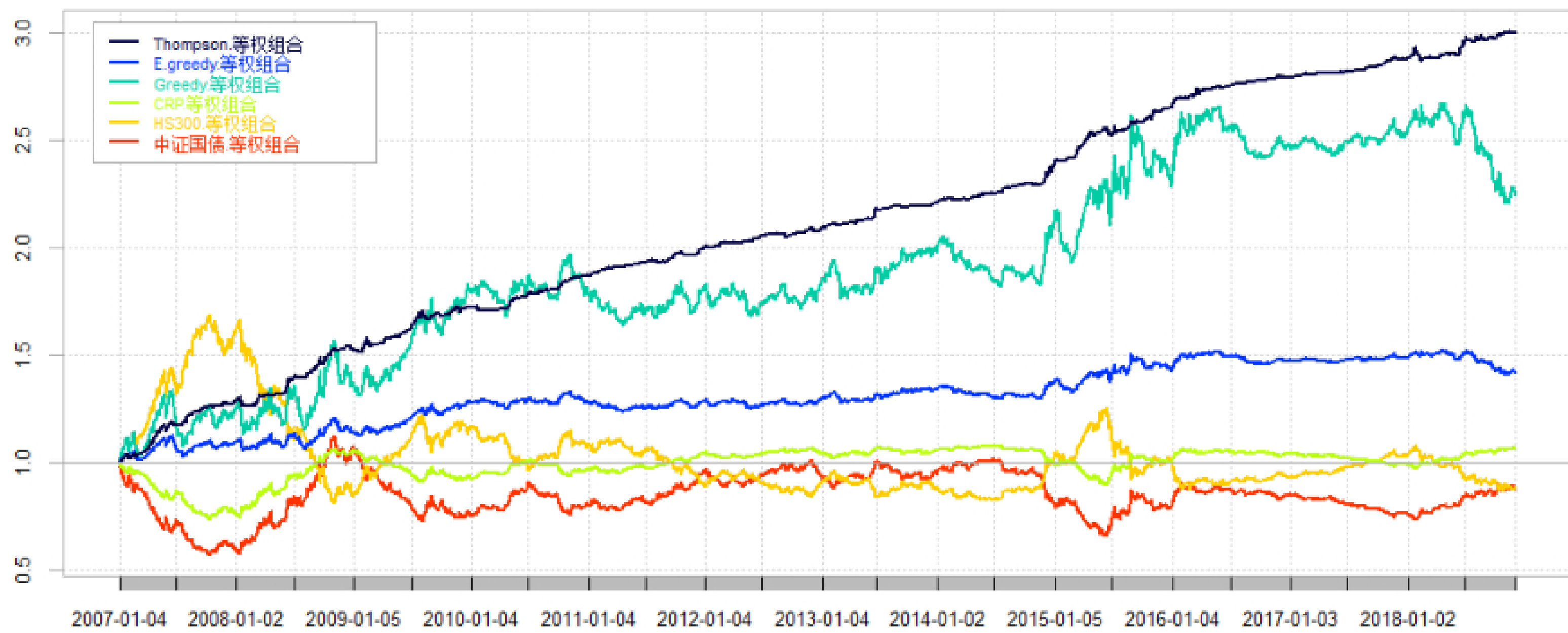
表 6: 模型历史回测结果 (股债混合)

	Thompson	E-greedy	Greedy	CRP	HS300	中证国债	等权组合
累计收益	437.30%	154.59%	301.70%	90.54%	59.77%	57.62%	79.01%
年化收益	15.73%	8.46%	12.84%	5.76%	4.16%	4.03%	5.19%
波动率	15.09%	14.69%	18.84%	7.92%	28.68%	2.16%	14.35%
最大回撤	32.13%	37.19%	35.90%	17.67%	72.30%	5.87%	42.49%
夏普比率	1.0381	0.5734	0.6785	0.7249	0.1443	1.8575	0.3602
信息比率	3.0044	0.5649	0.5289	0.0825	-0.0714	-0.0799	--
胜率	53.24%	51.90%	51.90%	47.41%	52.59%	47.41%	--

数据来源: 渤海证券研究所、Wind

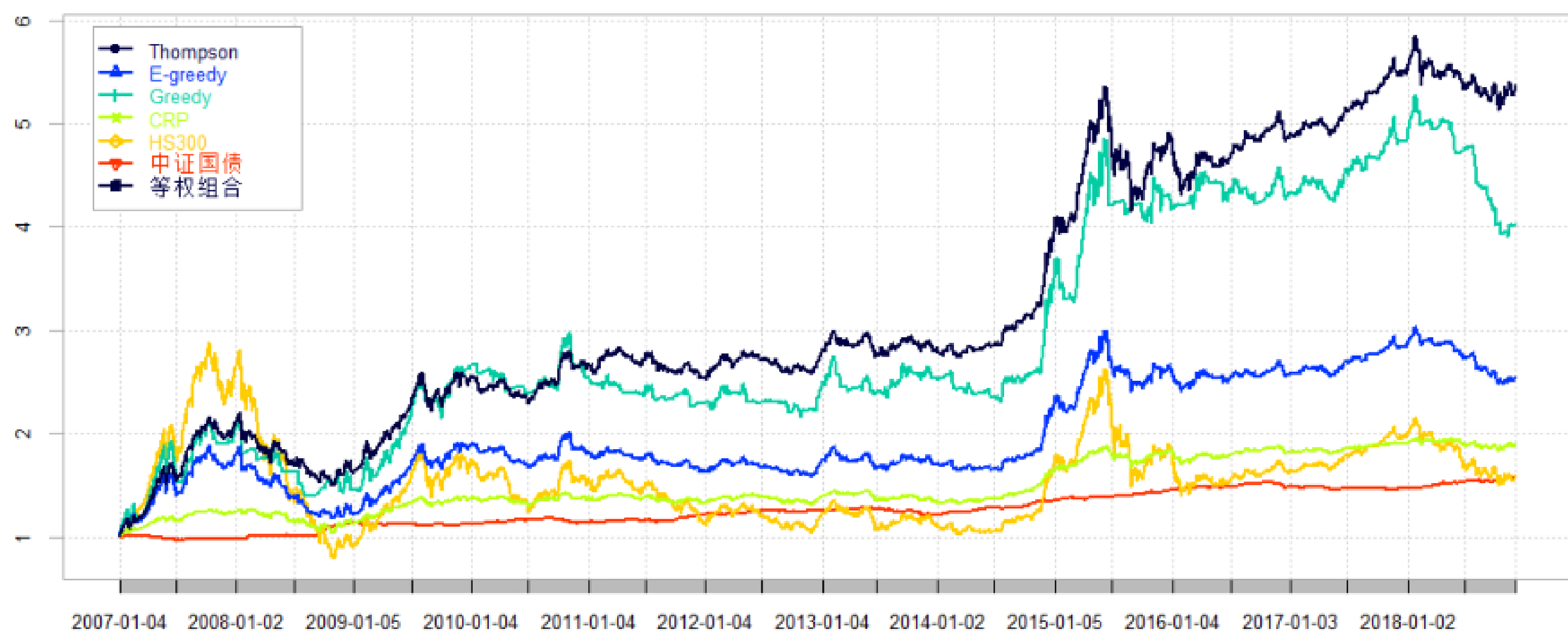
观察模型相对于股债等权组合的超额收益曲线,可以看出,两种资产的收益率呈现明显周期轮动的特性,而在每次周期轮动时, Thompson Sampling 的曲线就会得到一次提升,即每次轮动发生时, Thompson Sampling 方法都可以比其他算法更快的适应这种变化。

图 8：股债混合模型超额收益曲线



资料来源：Wind，渤海证券研究所

图 9：股债混合模型回测收益曲线



资料来源：Wind，渤海证券研究所

5. 总结与未来研究方向展望

本篇报告中，我们介绍了在线学习理论中的 Thompson Sampling 模型。并构建相关选股策略，应用于沪深 300 指数、中证 500 指数的市值因子择时以及股债混合模型。选股结果与传统多因子模型、以及之前介绍的 Bandit Learning 算法相比，表现出了更好的适应性，在风格转换的时候，可以最迅速的适应市场。我们

发现，在线学习算法尤其适用于具有周期性的资产投资中。对于其应用可以与市场判断相结合，当我们判断市场未来更多体现震荡行情时，可以使用这种方法来平滑风险；当我们判断市场未来更多体现趋势行情时，可以使用其他更偏重动量的策略，以取得更高的短期投资收益。在长期投资中，总的来说 Thompson Sampling 方法会优于其他方法，因为没有任何趋势可以一直持续，市场的周期性已经被很多研究论证过了。

2017 年以来，市场环境的剧变，使传统多因子模型面临挑战，对于未来的预测难度加大，Thompson Sampling 模型是一个可以被考虑的改进方法。

未来，我们会继续研究更多的在线学习模型，探索其适用范围、运行机制、收益来源以及成果的延续能力，并将其应用扩展到更多领域，如行业轮动、资产配置等。

风险提示：随着市场环境变化，模型存在失效风险。

附：报告参考

Shen W, Wang J, Jiang Y G, and Zha H. 2015. Portfolio choices with orthogonal bandit learning. *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, 974-980.

Shen W, Wang J. 2016. Portfolio Blending via Thompson Sampling. *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, 1983-1989.

Ran El-Yaniv 1998. Competitive Solutions for Online Financial Problems. *ACM Computing Surveys*, Vol.30, No.1

Li B, Hoi S C. 2014. Online portfolio selection: A survey. *ACM Computing Survey*, 46(3):35.

Bai J, Ng S. 2002. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191-221

投资评级说明

项目名称	投资评级	评级说明
公司评级标准	买入	未来 6 个月内相对沪深 300 指数涨幅超过 20%
	增持	未来 6 个月内相对沪深 300 指数涨幅介于 10%~20%之间
	中性	未来 6 个月内相对沪深 300 指数涨幅介于-10%~10%之间
	减持	未来 6 个月内相对沪深 300 指数跌幅超过 10%
行业评级标准	看好	未来 12 个月内相对于沪深 300 指数涨幅超过 10%
	中性	未来 12 个月内相对于沪深 300 指数涨幅介于-10%-10%之间
	看淡	未来 12 个月内相对于沪深 300 指数跌幅超过 10%

免责声明：本报告中的信息均来源于已公开的资料，我公司对这些信息的准确性和完整性不作任何保证，不保证该信息未经任何更新，也不保证本公司做出的任何建议不会发生任何变更。在任何情况下，报告中的信息或所表达的意见并不构成所述证券买卖的出价或询价。在任何情况下，我公司不就本报告中的任何内容对任何投资做出任何形式的担保，投资者自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失书面或口头承诺均为无效。我公司及其关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行或财务顾问服务。我公司的关联机构或个人可能在本报告公开发表之前已经使用或了解其中的信息。本报告的版权归渤海证券股份有限公司所有，未获得渤海证券股份有限公司事先书面授权，任何人不得对本报告进行任何形式的发布、复制。如引用、刊发，需注明出处为“渤海证券股份有限公司”，也不得对本报告进行有悖原意的删节和修改。

渤海证券股份有限公司研究所

所长&金融行业研究

张继袖

+86 22 2845 1845

副所长&产品研发部经理

崔健

+86 22 2845 1618

计算机行业研究小组

王洪磊（部门副经理）

+86 22 2845 1975

王磊

+86 22 2845 1802

汽车行业研究小组

郑连声

+86 22 2845 1904

张冬明

+86 22 2845 1857

陈兰芳

新材料行业研究

张敬华

+86 10 6810 4651

电力设备与新能源行业研究

刘瑀

+86 22 2386 1670

刘秀峰

+86 10 6810 4658

滕飞

+86 10 6810 4686

医药行业研究小组

赵波

+86 22 2845 1632

甘英健

陈晨

通信行业研究小组

徐勇

+86 10 6810 4602

节能环保行业研究

张敬华

+86 10 6810 4651

刘蕾

+86 10 6810 4662

餐饮旅游行业研究

刘瑀

+86 22 2386 1670

杨旭

+86 22 2845 1879

非银金融行业研究

洪程程

+86 10 6810 4609

中小盘行业研究

徐中华

+86 10 6810 4898

金融科技行业研究

王洪磊（部门副经理）

+86 22 2845 1975

张源

传媒行业研究

姚磊

固定收益研究

冯振

+86 22 2845 1605

夏捷

+86 22 2386 1355

金融工程研究

宋旻

+86 22 2845 1131

刘洋

+86 22 2386 1563

李莘泰

+86 22 2387 3122

张世良

金融工程研究

祝涛

+86 22 2845 1653

李元玮

+86 22 2387 3121

郝惊

+86 22 2386 1600

流动性、战略研究&部门经理

周喜

+86 22 2845 1972

策略研究

宋亦威

+86 22 2386 1608

严佩佩

宏观研究

宋亦威

+86 22 2386 1608

张扬

+86 22 28451945

孟凡迪

博士后工作站

朱林宇 资产配置

+86 22 2387 3123

张佳佳 资产配置

张一帆 公用事业、信用评级

综合质控&部门经理

齐艳莉

+86 22 2845 1625

机构销售•投资顾问

朱艳君

+86 22 2845 1995

刘璐

合规管理&部门经理

任宪功

+86 10 6810 4615

风控专员

白骐玮

+86 22 2845 1659

渤海证券研究所

天津

天津市南开区宾水西道 8 号

邮政编码: 300381

电话: (022) 28451888

传真: (022) 28451615

北京

北京市西城区西直门外大街甲 143 号 凯旋大厦 A 座 2 层

邮政编码: 100086

电话: (010) 68104192

传真: (010) 68104192

渤海证券研究所网址: www.ewww.com.cn