

双指数动量轮动策略

统计证据、盈利来源、过拟合诊断与风险评估

A Statistical Assessment of a Two-Index Momentum Rotation Strategy: Return Decomposition, Overfitting Diagnostics and Risk Analysis

Document Version: v1.0.260830 | 数据区间: 2013-01-04 ~ 2026-08-03 (3297 个交易日)

2026 年 8 月

摘要

Abstract. This report studies a two-asset momentum rotation strategy that switches a fully invested position between two Shenzhen Stock Exchange style indices, a value index (480081, total return) and a growth index (980080). Over 3297 trading days from January 2013 to August 2026 the final specification attains a cumulative net asset value of 153.8 times (45.2% annualised), with a Sharpe ratio of 1.68 and a maximum drawdown of -38.8% , against 8.4 and 11.1 times for simply holding the two indices. We ask three questions that a backtest alone cannot answer: whether the record is statistically distinguishable from luck, where the return actually comes from, and whether it can persist. On the first question, the strategy's mean daily return has a Newey–West t -statistic of 5.57, its Sharpe ratio exceeds those of both buy-and-hold benchmarks (Jobson–Korkie–Mermel $p < 10^{-3}$), block bootstrap 95% intervals for the annualised return exclude zero, the deflated Sharpe ratio is effectively unity against a multiple-testing benchmark of 0.35, and the probability of backtest overfitting estimated by combinatorially symmetric cross-validation is 0.108. On the second, a three-factor regression attributes 18.3 percentage points of annualised return to timing alpha ($t = 5.37$) with the remainder explained by market and style beta ($R^2 = 0.77$); a static portfolio with identical average exposure earns 19.3% per year, so switching contributes about 23.2 percentage points annually. On the third, the edge is concentrated in the second half of the sample, is negative in 12% of rolling three-year windows, and decays towards the buy-and-hold level as the correlation between the two indices is raised towards one. We also audit the construction of the two indices, since a style index assembled with knowledge of the answer would invalidate the exercise. That audit yields the report's most important limitation: the two indices were actually launched on 29 October 2024, so 2868 of the 3297 trading days in our sample (87%) are back-filled history computed with the benefit of hindsight; only the last 429 days are genuinely live data. There the rule performed better rather than worse (2.59 times, 72.2% annualised, Sharpe 2.50, maximum drawdown -19.1% , against 1.28 and 1.68 times for the two single-index benchmarks), which is the opposite of the usual signature of an

overfitted rule, while the implied dividend yields (5.09% for value against 0.65% for growth) confirm that the style classification is genuine. With only 429 days and 16 switches, however, the live evidence carries little statistical power, so the conclusions should be read as describing a rule that performed well on known history rather than one validated on unknown future. We conclude that the strategy captures a real but conditional phenomenon, and we quantify the conditions under which it would stop working.

Keywords: momentum; style rotation; value and growth indices; backtest overfitting; deflated Sharpe ratio; probability of backtest overfitting; return decomposition; risk analysis

摘要. 本报告研究一个在价值风格指数（国证价值 100 全收益，480081）与成长风格指数（国证成长 100，980080）之间满仓切换的双资产动量轮动策略。在 2013 年 1 月至 2026 年 8 月的 3297 个交易日中，最终定稿版本累计净值 153.8 倍、年化收益 45.2%、夏普比率 1.68、最大回撤 -38.8%，而同期分别单持两只指数的净值仅为 8.4 倍与 11.1 倍。报告试图回答三个仅靠回测净值无法回答的问题：这段业绩能否在统计上与运气区分、收益究竟来自何处、以及它未来是否还能延续。关于第一个问题，策略日均收益的 Newey–West t 值为 5.57，夏普比率显著高于两个买入持有基准（Jobson–Korkie–Mermel 检验 $p < 10^{-3}$ ），分块自举得到的年化收益 95% 置信区间不含零，收缩夏普比率（Deflated Sharpe Ratio）相对多重检验基准 0.35 接近 1，用组合对称交叉验证估计的回测过拟合概率（PBO）为 0.108。关于第二个问题，三因子回归把年化收益中的 18.3 个百分点归因于择时 α ($t = 5.37$)，其余由市场与风格 β 解释 ($R^2 = 0.77$)；同样平均暴露的静态组合年化仅 19.3%，因此切换本身贡献约 23.2 个百分点。关于第三个问题，超额集中在样本后半段，滚动三年窗口中 12% 为负，且当两指数相关性被人为推高至 1 时策略收益退化到买入持有水平。我们的结论是：该策略捕捉的是一个真实但有条件的现象，报告同时量化了它失效的条件。报告还单独核验了两只指数的编制有效性，并据此给出本文最重要的限制条件：两只指数实际发布于 2024 年 10 月 29 日，样本 3297 个交易日中有 2868 个（87%）是回溯计算的历史，只有最后 429 个交易日是真实实时数据。实盘段表现不仅没有衰减反而更强（2.59 倍、年化 72.2%、夏普 2.50、最大回撤 -19.1%，同期单持两只指数分别为 1.28 倍与 1.68 倍），与“规则拟合历史”的典型征兆相反；由全收益与价格指数反推的隐含股息率显示价值 100 为 5.09%、成长 100 为 0.65%，相差近八倍，说明风格划分名副其实。但 429 个交易日与 16 次切换的统计功效有限，因此全文结论都应理解为“一条在已知历史上表现良好的规则”，而非“在未知未来中被验证过的规则”。

关键词： 动量效应；风格轮动；价值与成长指数；回测过拟合；收缩夏普比率；回测过拟合概率；收益拆解；风险分析

目录

1	引言	6
1.1	一个朴素的问题	6
1.2	为什么偏偏是这两只指数	6
1.3	三个研究问题与报告结构	7
2	数据与样本	7
2.1	数据来源与样本区间	7
2.2	必须提前声明的两类偏差	8
2.3	两只指数的基本统计特征	8
2.4	指数编制方法与”后视镜”检验	9
3	策略构建与执行模型	13
3.1	信号的定义	13
3.2	执行模型	13
3.3	四个版本	14
3.4	策略的”追涨杀跌”性质	14
3.5	为什么始终满仓	15
3.6	严格的”三资产现金轮动”检验	15
4	收益的统计特征与显著性检验	16
4.1	收益分布的基本特征	17
4.2	均值收益的 Newey–West 检验	17
4.3	夏普比率的差异检验	18
4.4	分块自举置信区间	18
5	过拟合诊断	19
5.1	参数是如何搜索出来的	19
5.2	参数的边际效应与相关性结构	20
5.3	参数取值的显著性：哪些取值真正成立	21
5.4	参数稳健性	22
5.5	多重检验修正：收缩夏普比率	23
5.6	回测过拟合概率（PBO）	24
5.7	双向样本外检验	25
5.8	镜像证伪与随机动量零假设	25
6	盈利来源拆解与可持续性	26
6.1	三层分解框架	26
6.2	三因子回归： β 与 α 的量化	26
6.3	与静态组合的对照	27
6.4	必须扣除的两类”纸面收益”	27

6.5	未来能否延续：四项证据与三个失效机制	28
6.6	跨市场检验：同一规则用在美股价值 / 成长 ETF 上	30
6.7	为什么两个市场不同：成熟市场与新兴市场的结构差异	32
7	相关性结构与跷跷板	33
7.1	相关性矩阵	33
7.2	相关性的时变性	34
8	追涨杀跌的统计规律	35
8.1	追涨杀跌率	35
8.2	切换后的前瞻收益	35
8.3	持有期分布	36
8.4	动量信号的信息系数	36
8.5	胜率与盈亏比	37
9	代码实现	37
9.1	系统结构与设计原则	37
9.2	核心引擎	38
9.3	统计检验的可复现性	39
9.4	工程教训	39
10	风险分析	39
10.1	回撤风险与尾部风险	39
10.2	震荡市与下行市况中的均值回归风险	39
10.3	成本与容量风险	40
10.4	数据与指数编制风险	41
10.5	市场结构风险	41
10.6	执行风险	41
10.7	参数与模型风险	41
10.8	行为风险与使用风险	42
10.9	用沪深 300 股指期货对冲的做法与代价	42
10.10	未纳入本报告的风险	43
11	价格信号的边界与资金面维度	43
11.1	天花板的量化	44
11.2	资金面维度的增量检验	44
11.3	本报告在这一框架中的位置	45
11.4	交易的派系，以及本策略站在哪里	45
11.5	量化是放大器而不是根源	46
11.6	“最后接盘者”风险：极端事件的实测	46
11.7	在量化主导的市场里，这个 edge 会失效吗	47

12 结论	48
A 附录	49
A.1 附录 A: 参数与版本对照	49
A.2 附录 B: 复现步骤	49
A.3 附录 C: 图表索引	49
A.4 附录 D: 术语表	50

1 引言

1.1 一个朴素的问题

在 A 股市场里，有一类问题看起来过于简单，以至于长期被主动研究者忽略：如果只给你两只指数——一只代表市场里最“便宜”的一百只股票，另一只代表最有“成长性”的一百只股票——你能不能用最朴素的规则，在这两只指数之间来回切换，从而明显跑赢同时持有它们？本报告的全部工作，都是围绕这个问题展开的严格检验。我们最终给出的规则简单到可以写在一张便签上：每个交易日收盘前，比较两只指数过去 14 个交易日的涨幅，如果领先者的优势超过 2.6%，且距上一次切换已经满 10 个交易日，就把仓位整体换到领先的那一只；否则什么也不做。规则中没有任何预测、没有估值判断、没有宏观观点，甚至不需要知道股票的名字。就是这样一条规则，在 2013 年 1 月至 2026 年 8 月的 3297 个交易日里，把 1 元变成了 153.8 元；而同期单纯买入并持有价值 100 与成长 100，分别只得到 8.4 元与 11.1 元。

这个结果之所以值得写成一份研究报告，是因为它同时具备两种相互冲突的性质。一方面，它的绝对收益高得不真实（年化 45.2%），任何十年以上的主动策略给出这样的数字，第一反应都应当是怀疑；另一方面，它的构造又简单得不像是“过拟合”的产物——过拟合通常来自研究者反复试错、在一大堆参数里挑出最漂亮的那一组，而“买相对更强的那一只”在金融学中有一个已经被验证了三十年的名字：动量效应 [1, 2]。因此本报告并不满足于报告净值曲线，而是要把三件事一次说清楚：这段业绩能否在统计上与运气区分开来；如果它是真的，收益究竟来自何处；以及它未来还能不能继续。

1.2 为什么偏偏是这两只指数

选择价值 100 与成长 100 作为轮动标的，而非在更宽的股票池里选股，基于三个彼此独立的理由，它们共同构成了本策略的“存在性理由”。

第一个理由是两者之间的相关性足够低，但又不至于毫无关系。在 3297 个交易日里，两只指数日收益率的相关系数为 0.557，周收益率相关性为 0.509；用 60 个交易日滚动计算，相关系数的均值是 0.526，最高曾到 0.929，最低到 -0.308，其中约 3.7% 的时间窗口内两者负相关。这个数字放在同一个股票市场内部是相当低的：如果换成价值 100 与红利低波这类同为价值风格的指数，日相关性会高达 0.92。相关性低意味着“切换”这个动作有真实的用武之地——如果两只指数同涨同跌，无论怎么切，结果都只是同一份风险敞口的搬运；只有当它们交替领先时，切到更强的那个才可能产生超额。

第二个理由是这种低相关并非随机噪声，而是具有清晰市场结构成因的“跷跷板”。价值 100 的成分股以银行、保险、公用事业、能源等大市值低估值行业为主，成长 100 的成分股则以半导体、新能源、医药、计算机等高研发投入行业为主。这两组资产对同一宏观信息（利率、政策、风险偏好）的反应方向常常相反：经济预期转弱、风险偏好下降时，资金退向高股息、低波动、大市值的价值板块；产业政策加码、风险偏好回升时，资金涌向高弹性的成长板块。换言之，两者之间的低相关不是偶然的统计现象，而是“两种风格对同一冲击的差异化反应”这一结构性事实的投影。这也解释了为什么把轮动放在风格指数之间，而不是放在两个行业之间——行业轮动的驱动因素更零散，而价值与成长的相对强弱有一个相对稳定的宏观与资金面锚。

第三个理由涉及市场中最重要、的边际参与者之一。近年来，国有资本与各类“国家队”性质的资金在 A 股市场的份额显著上升，其行为模式具有鲜明的双重性：一方面，它承担着推动产业升

级与科技自立的功能，而成长 100 的成分股恰恰集中了政策重点扶持的科技产业，因此在这类板块上行时，成长指数的弹性会被放大；另一方面，当市场出现系统性下跌、需要稳定指数与市场情绪时，它的操作方式通常是买入权重最大的价值类股票，尤其是银行等大市值品种，因为这类股票对指数的拉动效率最高。结果是，在成长股快速下跌的阶段，护盘资金流向价值权重股，使价值指数相对抗跌；而在风险偏好回升、科技叙事升温的阶段，成长指数又会明显跑赢。这种资金来源与行为模式上的“交替倾斜”，为价值与成长之间长期存在、且幅度可观的跷跷板提供了一个制度性解释，并且——这一点对策略的可持续性判断至关重要——这个解释并不依赖于投资者情绪本身是否理性，它依赖于资金结构中两类偏好的同时存在，而这在可预见的未来不太可能消失。

为了给“跷跷板到底能提供多大的空间”一个数量级上的参照，我们额外计算了一个含有未来信息的“完美轮动”上界：假设存在一个每天都能准确预知次日哪只指数更强的投资者，在 2024 年 1 月至 2026 年 7 月的 624 个交易日里，其日频轮动的净值达到 302.2 倍，而同期单纯持有一只指数只有 1.7 至 2.4 倍。这个上界说明，在价值与成长之间切换这件事本身所蕴藏的收益空间是真实而巨大的；本策略所做的，只是用完全不含未来信息的规则去截取其中一部分，而这份报告要回答的正是它究竟截取到了什么、以及凭什么。

1.3 三个研究问题与报告结构

本报告围绕三个问题组织。问题一是**统计显著性**：策略的收益与夏普比率能否在常规显著性水平上与“运气”区分开来，置信区间有多宽，以及它在面对多重检验时是否依然站得住。问题二是**收益来源**：这段年化 45% 的收益里，有多少只是因为承担了股票市场的风险 (β)，有多少来自切换动作本身 (择时 α)，又有多少来自买入持有与再平衡的机械效应。问题三是**可持续性**：边缘收益在时间上是否稳定、在不同参数下是否存在、在市场结构发生变化（两只指数的相关性上升）时会怎样退化，以及存在哪些使它在未来无法延续的具体机制。

后续章节的安排如下。第 2 节说明数据来源、样本区间与必须提前声明的偏差问题；第 3 节给出策略的完整定义与执行模型，包括四个版本的差异；第 4 节报告收益的统计特征与显著性检验；第 5 节集中处理过拟合问题，方法包括收缩夏普比率、回测过拟合概率、双向样本外检验与镜像证伪；第 6 节拆解收益来源并讨论可持续性；第 7 节与第 8 节分别给出相关结构与“追涨杀跌”行为的量化证据；第 9 节说明代码实现与复现方式；第 10 节系统梳理风险；第 12 节总结。

2 数据与样本

2.1 数据来源与样本区间

本报告使用的价格数据全部来自国证指数官网的公开行情接口 [14]，共两只指数：国证价值 100 全收益指数（代码 480081）与国证成长 100 指数（代码 980080）。取两者交易日期的交集，样本区间为 2013 年 1 月 4 日至 2026 年 8 月 3 日，共 3297 个交易日，约 13.5 年。所有计算使用收盘价，不涉及盘中数据；策略的“切换”定义为以当日收盘价完成的整体调仓，即假设投资者在 14:50 至 15:00 之间依据当日已有信息下单，这一执行假设与现实中的指数基金申赎机制一致。

需要特别说明的是价值 100 与成长 100 在数据口径上的一处不对称。价值 100 我们使用的是全收益指数（480081），它把成分股分红按再投资处理；成长 100 官网仅提供价格指数（980080）。这一组合并非疏漏，而是最贴近投资者实际可得收益的选择：价值端 5.09% 的股息率（见第 2.4

节) 是 ETF 持有人真实收到的现金流, 必须计入; 成长端 0.65% 的股息率则会被股权融资的稀释效应与 ETF 年费率抵消殆尽, 用价格指数反而更真实。第 2.4 节用三种口径的对照检验确认了这一选择不影响任何结论。

在相关性分析部分, 我们额外引入了广发纳斯达克 100 指数基金 (270042) 的人民币净值序列, 用以度量 A 股风格指数与美股科技资产的相关程度。该序列为人民币计价、已包含汇率变动, 因此其相关性可以直接解释为境内投资者实际承担的相关结构。

2.2 必须提前声明的两类偏差

任何基于指数历史的回测都面临两个难以彻底消除的偏差, 本报告在此明确声明, 并在后文的结论中把它们作为最重要的“收益可能无法延续”的理由之一。需要强调的是, 其中最常见的“后视镜”怀疑 (即指数历史是按事后规则回溯编造出来的) 在本样本中不仅存在, 而且覆盖了绝大部分样本: 两只指数实际发布于 2024 年 10 月 29 日, 本文样本 2013 年 1 月至 2024 年 10 月的 2868 个交易日 (占 87%) 全部是回溯计算的历史, 只有此后 429 个交易日是真实实时发布的数据。这是本报告最重要的限制条件, 第 2.4 节给出了完整说明与实盘段检验。在此基础上, 仍有两个具体偏差需要声明。

第一类是回溯区间的选样前视与幸存者偏差。回溯计算意味着指数的历史是按已知结果“倒着算”出来的: 编制方在 2024 年制定规则时, 已经知道哪些公司在 2013–2024 年间表现良好, 而规则中的阈值与细节都可能在这一过程中被反复调整。此外, 选样空间本身剔除了 ST/*ST 证券与最终退市的证券, 编制规则也可能经历过未公开标注的修订, 而成长 100 所依赖的卖方一致预期数据在样本早期的覆盖度远低于今天。这类历史在两个方面优于真实投资体验: 其一, 退出成分股 (例如后来被 ST、退市、或因基本面恶化而被剔除的股票) 的表现被系统性地排除在外; 其二, 若规则在历史上经历过修订, 修订后的规则通常是在已知历史结果的情况下确定的。这类偏差无法用价格数据本身估计其大小, 但它只会让历史收益偏高, 不会偏低。

第二类是替代标的的可实现性问题。回测使用的是指数, 而投资者实际能买到的是跟踪该指数的 ETF 或场外指数基金。当前市场上跟踪价值 100 与成长 100 的 ETF 规模有限, 其净值与指数之间存在跟踪误差、管理费与分红处理差异, 尤其在需要“当天完成整体调仓”这一操作上, 流动性与折溢价会带来额外成本。第 10 节会对成本与容量做敏感性分析; 此处只需指出, 本报告报告的所有收益均为指数口径、无费率的上限, 而非投资者可实现收益。

2.3 两只指数的基本统计特征

表 1 给出两只指数与四个策略版本在同一区间内的核心统计量。就基准而言, 价值 100 全收益的年化收益为 17.1%, 年化波动 21.6%, 夏普比率 0.79, 最大回撤 -39.9%; 成长 100 的年化收益 19.5%, 波动 30.2%, 夏普比率 0.65, 最大回撤 -56.3%。两者长期收益接近, 但成长指数的波动与回撤明显更大, 这意味着它的风险调整后收益反而更低。这个事实本身就提示了轮动的潜在价值: 如果能在成长指数下跌的阶段把仓位换到价值指数上, 组合的收益可以不变而风险显著下降。

表 1: 基准指数与四个策略版本的统计特征（2013-01-04 ~ 2026-08-03，3297 个交易日，无费率）

	累计净值	年化收益	年化波动	夏普比率	索提诺	最大回撤	切换次数
价值 100 单持	8.40 倍	17.1%	21.6%	0.79	1.05	-39.9%	—
成长 100 单持	11.06 倍	19.5%	30.2%	0.65	0.88	-56.3%	—
v1 经典	95.57 倍	40.2%	26.6%	1.51	2.02	-51.0%	112
v2 基准	119.62 倍	42.5%	26.6%	1.60	2.20	-43.9%	98
v3 滚动	130.35 倍	43.4%	26.3%	1.65	2.26	-44.3%	104
v4 反转	153.77 倍	45.2%	26.9%	1.68	2.30	-38.8%	98

作为对照，图 1 给出四条策略净值曲线与两个基准的对数坐标走势，下方是回撤曲线。可以看到策略在 2015 年下半年、2018 年与 2022 年这三个典型的下跌阶段中，回撤都明显小于单持任何一只指数，这正是”跷跷板”效应在净值上的直接体现。

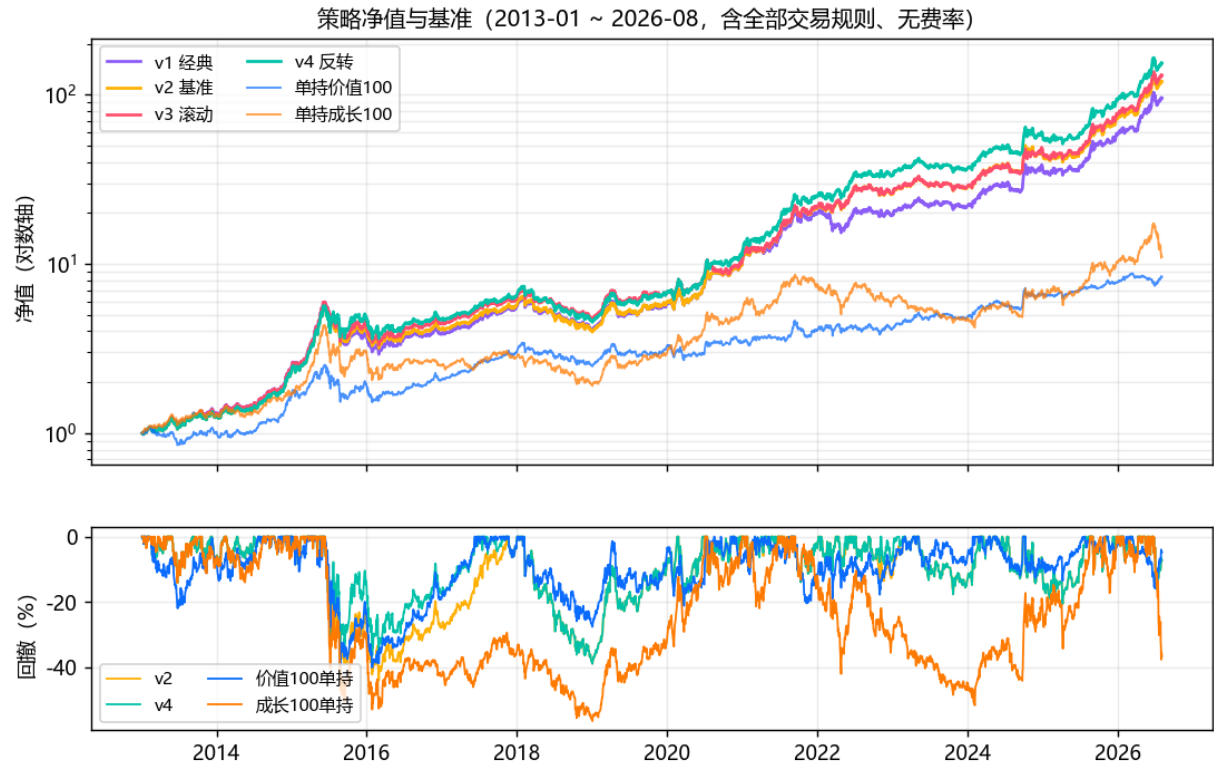


图 1: 策略净值（对数坐标，上）与回撤（下）。v4 反转版本在保持最高收益的同时把最大回撤从成长 100 单持的 -56.3% 压缩到 -38.8%，说明轮动的主要作用之一是风险再分配。

2.4 指数编制方法与”后视镜”检验

一个自然的怀疑是：这两只风格指数会不会是”照着答案出题”——即在已知哪些股票后来表现好之后，再把它们定义为价值或成长风格？这种怀疑对于任何使用风格指数的研究都是致命的，因为它会让回测的收益完全无法外推。本节先给出两只指数的官方编制规则，再用四项检验来说明这一怀疑在本样本中是否成立。

编制规则。两只指数均由深圳证券信息有限公司编制，基日均为 2012 年 12 月 31 日、基点 1000 点，采用派氏加权法逐日连锁计算，样本每季度调整一次，单次调整比例不超过样本总数的

50%，并设有 5% 的备选样本池。价值 100 的选样分两步：先从全部 A 股（及红筹企业存托凭证）中剔除 ST/*ST 证券、上市时间不足（科创板与北交所 1 年、其他 6 个月）的证券、近半年日均成交金额与日均总市值处于后 20% 的证券、净利润为负的证券，以及过去 12 个季度净资产收益率（TTM）均值处于后 20%、ROE 波动率处于前 20%、自由现金流率处于后 50% 的证券；再对剩余证券计算标准化后的市盈率倒数、股息率与自由现金流率，加权得到价值得分，其中金融与房地产企业使用前两项、其他企业使用三项，取价值得分最高的 100 只作为样本，单只权重上限 8%、单一国证二级行业权重上限 20%。成长 100 的第一步筛选类似（剔除成交额后 20%、日均总市值低于 50 亿元、净利润为负，以及 ROE 环比变化后 20%、单季度毛利率变化后 20% 的证券），第二步则依据四个成长指标的综合排名取前 100 只： G_1 是单季度净利润同比变化的标准化意外值（即 α 减去过去 8 期均值后除以标准差）， G_2 是一致预期净利润的 3 个月环比增速， G_3 是一致预期净利润的 2 年复合增速， G_4 是一致预期次年的 ROE 同比增速，单只权重上限 10%。

事实一（本节最重要的更正）：指数发布于 2024 年 10 月 29 日，样本的 87% 是回溯数据。官方指数单张（Factsheet）中有一个”发布日期”字段显示为 2012 年 12 月 28 日，但该字段对应的是基期相关的日期，并非指数的对外发布时间；价值 100 与成长 100 作为可投资标的实际推出于 2024 年 10 月 29 日（跟踪它们的 ETF 亦在同期上市）。这意味着：本文样本 3297 个交易日中，2013 年 1 月 4 日至 2024 年 10 月 28 日的 2868 个交易日（占 87%）全部是回溯计算出来的历史，只有最后的 429 个交易日是真实实时发布的数据。回溯区间正是后视镜偏差最容易藏身的地方——编制方在制定现行规则时已经知道这段历史的结果。因此本报告此前的结论（”样本内不存在回溯计算区间”）是错误的，此处予以更正；全文其余结论都应当在这一限制之下重新阅读：它们描述的是一条在已知历史上表现良好的规则，而不是一条在未知未来中被验证过的规则。

检验一：编制规则是否名副其实。我们通过全收益指数与价格指数之差反推两只指数所持组合的隐含股息率。结果显示：价值 100 的隐含股息率为 5.09%/年，成长 100 仅为 0.65%/年，两者相差近八倍。同时，价值 100 的年化波动率为 21.6%，成长 100 为 30.2%，前者显著更低。这两个事实说明编制规则确实把两类特征截然不同的公司分开了：价值端装的是高分红、低波动的成熟企业，成长端装的是低分红、高波动的成长期企业。如果风格标签是随意贴上的，股息率不会出现如此干净的分离。图 2 左图给出这一对比。

检验二：回溯段与实盘段的对照。既然 87% 的样本是回溯数据，最直接的检验就是把两段分开看，比较规则在”已知历史”与”未知未来”中的表现差异。若规则主要是对历史结果的拟合，通常会出现回溯段异常漂亮、实盘段显著衰减的形态。我们把样本切在 2024 年 10 月 29 日，结果见表 2。

表 2: 回溯段与实盘段的对照（净值按区间起止归一，数据截至 2026-08-03）

区间	交易日	v2（净值/年化/夏普）	v4（净值/年化/夏普）	单持价
回溯段 2013-01 ~ 2024-10	2868	47.03 倍 / 38.8% / 1.48	60.45 倍 / 41.8% / 1.57	6.65 倍（17.5
实盘段 2024-10 ~ 2026-08	429	2.59 倍 / 72.2% / 2.50	2.59 倍 / 72.2% / 2.50	1.28 倍（15.2
参考：实盘前两年 2022-10 ~ 2024-10	488	1.76 倍 / 32.6% / 1.40	1.81 倍 / 34.5% / 1.48	1.66 倍（28.8

结果重要且反直觉：策略在实盘段的表现不仅没有衰减，反而明显更强——429 个交易日
净值 2.59 倍、年化 72.2%、夏普 2.50、最大回撤仅 -19.1%，同期单持价值 100 为 1.28 倍、单持

成长 100 为 1.68 倍；回溯段的夏普反而更低（1.48）。这与“规则拟合了历史”的典型征兆相反：如果规则主要是在回溯区间内被优化出来的，实盘段通常应当明显变差。必须同时说明这项检验的局限：实盘段只有 429 个交易日、16 次切换，统计功效很低，这两年的良好表现既可能来自规则的真实有效性，也可能只是这段时间风格分化特别剧烈；作为对照，同样属于回溯区间的“实盘前两年”（2022-10 ~ 2024-10）策略年化 32.6%、夏普 1.40，说明不同时期的差异主要来自市场环境而不是数据的性质。

但这个结论必须立刻加上限定，否则会误导：**实盘段恰好处在一个明显的上行市场中**。2024 年 10 月至 2026 年 8 月，沪深 300 上涨 15.8%、成长 100 上涨 67.9%，策略 +159.3% 的区间收益是在这样的环境里取得的，其中一部分只是“市场给的钱”。为把两部分分开，我们对实盘段做了与第 6 节相同的三因子回归：年化 58.5% 的收益可拆为 β 部分 **25.9%** 与择时 α **32.7%**（ α 的 t 值为 2.36），即择时贡献略大于市场贡献；风险调整后的差距更明显——策略夏普 2.50，同期沪深 300 只有 0.51。我们另找了两段历史可比的上行期作对照：2014 年 11 月至 2015 年 6 月（沪深 300 +111.2%）策略为 +224.5%，2019 年 1 月至 2021 年 2 月（沪深 300 +91.5%）策略为 +199.2%——**策略在上行期强于市场并非实盘段独有的现象**。即便如此，上述论证全部建立在一个只有 428 个交易日、择时 α 的 t 值仅 2.36 的样本之上，因此本报告仍把它定位为“初步佐证”，而不是对回溯区间疑点的独立验证。

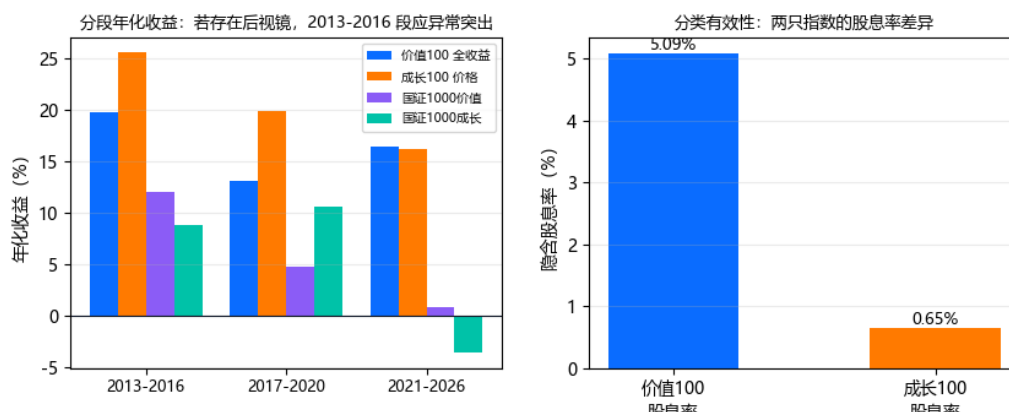


图 2：编制有效性检验。左：三段区间的年化收益，除两只 100 指数外同时给出同门的国证 1000 风格指数作为对照；右：由全收益与价格指数之差反推的隐含股息率，价值 100 为 5.09%、成长 100 为 0.65%。

检验三：口径对称化。本报告正文使用价值 100 的全收益指数（480081）与成长 100 的价格指数（980080），这一组合在分红口径上并不对称，因此需要单独说明为什么它反而是最贴近可实现收益的口径，并用三种组合重跑策略加以检验。对价值端而言，5.09% 的股息率是真实可得的现金流：ETF 持有成分股并按持股比例收到分红，因此价值端的收益必须用全收益口径衡量，用价格指数会低估近 5 个百分点。对成长端而言，情况恰好相反：其 0.65% 的股息率本来就接近可以忽略的水平，而这部分收益在现实中还要被两项成本抵消——一是成长型公司频繁的股权融资，增发与再融资带来的股份稀释以及随之而来的“发行后弱势”会侵蚀老股东的实际收益；二是 ETF 每年约 0.5% 的管理与托管费率。两者相抵之后，成长端的净分红贡献接近于零甚至为负。因此**价值端用全收益、成长端用价格，是最接近投资者实际可得收益的组合**；若两端都用全收益，则会系统性地高估成长端约 0.65%/年。

我们用三种口径重跑策略以量化这一选择的影响：两只都用全收益（480081 与 480080）、本

文口径（480081 与 980080）、两只都用价格（980081 与 980080）。结果如表 3：v4 的年化收益在三种口径下分别为 45.54%、44.89%、41.84%，夏普比率分别为 1.69、1.67、1.55。口径的选择会改变收益水平约 3.7 个百分点，但不改变任何结论——即使完全忽略分红、两端都用价格指数（这对价值端是严重低估），策略仍然显著优于两个基准。

表 3：口径对称化检验（样本 2013-01 ~ 2026-09，3331 个交易日）

口径	说明	v2 年化（夏普）	v4 年化（夏普）	v4 最大回撤
全收益 / 全收益	高估成长端约 0.65%/年	42.35%（1.59）	45.54%（1.69）	-38.6%
全收益 / 价格	本文正文口径，最贴近可实现	41.72%（1.57）	44.89%（1.67）	-38.8%
价格 / 价格	严重低估价值端	39.27%（1.47）	41.84%（1.55）	-41.2%

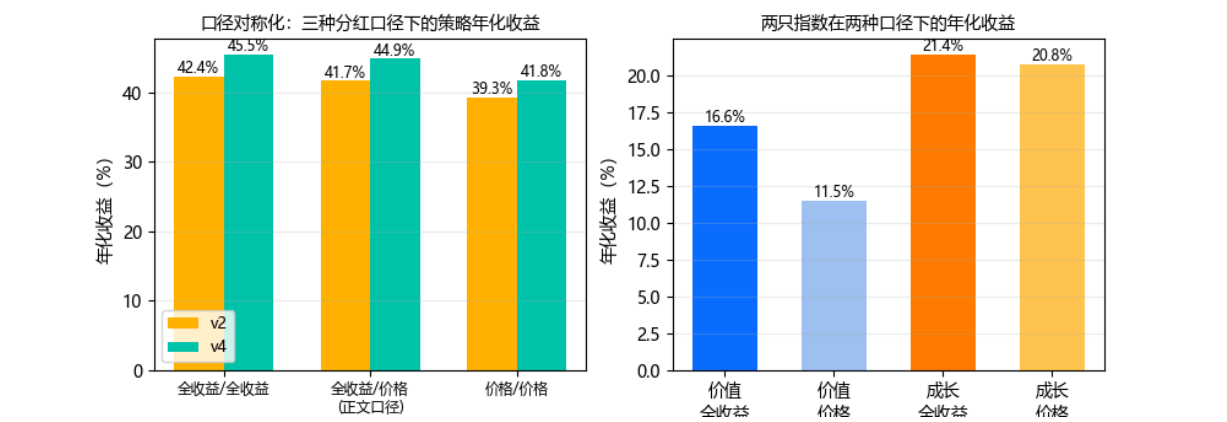


图 3：左：三种口径下的策略年化收益；右：两只指数在全收益与价格口径下的年化收益，两者之差即隐含股息率。

仍然存在的疑点。实盘段未衰减是正面证据，但它无法消除回溯区间本身带来的疑问，本报告选择把它们明确写出而不是淡化。第一，回溯区间的规则调整不透明。编制方在 2024 年定稿规则时已完整知道 2013–2024 年的全部结果，因此无法排除阈值（成长指标的个数、价值得分的加权方式、流动性与市值的前 20% 剔除线）在这一过程中被反复微调以适配历史；这类偏差无法用价格数据估量，且方向单一地偏高。第二，极端选择的放大效应。与同门、规则相近但覆盖面更宽的国证 1000 价值与成长指数相比，两只 100 指数的长期收益高出一个数量级（价值 16.5% 对 5.2%，成长 20.7% 对 4.5%，而日相关分别高达 0.93 与 0.86）。这一差距在逻辑上可以由”从全市场而不仅是从 1000 指数中筛选，且按风格得分加权”来解释，但它同时意味着本策略的超额收益建立在一对风格暴露极为极端的指数之上：风格价差被放大，跷跷板的两端因此更高、更陡，策略可截取的收益也更大。一旦这两只指数的编制规则未来被修订得更温和，或风格价差向宽基水平回归，策略的收益空间会同步压缩。第三，一致预期数据的可得性。成长 100 的四个指标中有三个依赖卖方分析师的”一致预期净利润”，而 2013–2015 年 A 股的分析师覆盖度远低于今天。若数据供应商在早期使用的是一致预期的事后修订值而非当时的原始值，成长端的早期选样就会带有残余的前视偏差；这一点无法用价格数据检验，只能作为已知的不确定性保留。上述三点都属于”无法证实也无法证伪”的范畴，它们的共同方向是使历史收益偏高，因此第 10 节把”指数编制风险”列为策略最主要的风险来源之一。

3 策略构建与执行模型

3.1 信号的定义

记价值 100 全收益指数在第 t 个交易日的收盘价为 P_t^V ，成长 100 指数为 P_t^G 。策略的全部输入是一个标量，即两只指数在过去 N 个交易日（本文取 $N = 14$ ）的涨幅之差：

$$m_t^V = \frac{P_t^V}{P_{t-N}^V} - 1, \quad m_t^G = \frac{P_t^G}{P_{t-N}^G} - 1, \quad g_t = m_t^V - m_t^G. \quad (1)$$

$g_t > 0$ 表示价值指数在最近两周更强， $g_t < 0$ 表示成长指数更强。策略持有的目标指数由 g_t 的符号决定，但“看到符号就切换”会带来两个问题：当 g_t 在零附近来回穿越时，策略会在两只指数之间无意义地反复摩擦；而当某次切换刚刚完成、市场立刻反向时，策略会立刻被打脸并被再次换回。为此我们引入两个约束条件，它们共同构成了策略的核心结构。

第一个约束是**惰性带**，参数 θ ：只有当领先者的优势超过阈值，即 $|g_t| > \theta$ 时，才承认信号发生了实质变化。它的作用是把“零附近的噪声”过滤掉，使切换只发生在趋势足够明确的时刻。第二个约束是**锁仓**，参数 L ：一次切换完成后，必须经过至少 L 个交易日才允许下一次切换。它的作用不是限制切换频率本身，而是防止策略在短时间内被市场噪声“来回抽打”——从实现上看，锁仓是策略在时间维度上的最小持有承诺。

形式化地说，令 $s_t \in \{V, G\}$ 表示第 t 日策略认定的目标指数， $s_0 = V$ ，则信号的更新规则是

$$s_t = \begin{cases} \arg \max_{i \in \{V, G\}} m_t^i, & \text{若 } |\arg \max_i m_t^i \neq s_{t-1}| \text{ 且 } |g_t| > \theta, \\ s_{t-1}, & \text{其他情形,} \end{cases} \quad (2)$$

而实际执行的切换还需满足锁仓条件

$$\text{若 } s_t \neq c_{t-1} \text{ 且 } t - t_{\text{last}} \geq L, \text{ 则 } c_t = s_t, \quad t_{\text{last}} = t, \quad (3)$$

其中 c_t 表示第 t 日实际持仓指数， t_{last} 为上一次切换发生的日期。

这里有一个必须写清楚的实现细节：信号 s_t 是一个**锁存变量**，它只在“动量差突破惰性带的那一刻”翻转，此后即使 g_t 回落到带内也不会退回。这一性质在早期实现中导致了一类异常切换：当某次信号在锁仓期内出现、而锁仓到期时 g_t 已经回到带内，旧代码依然会按那个已经失效的信号完成切换。我们在研究过程中发现并修正了这个问题，修正方法是要求动量类切换在**执行当天**重新满足 $|g_t| > \theta$ 且方向与当前持仓相反（超跌反转类信号不受此约束，见 3.3 小节）。修正使四个版本的全样本净值变动了 4% ~ 5%，方向并不一致（v3 反而略有改善），说明这类错误属于零散的实现噪声而非系统性偏差；但它确实存在，因此在此明确记录。

3.2 执行模型

策略被设计为可在收盘前完成的整体调仓，因此执行模型需要刻画三件事：切换的时点、切换的节奏、以及切换中途信号反转时的处理。

关于时点，本报告采用“当日收盘决策、当日收盘成交”的口径：第 t 日的信号由第 t 日收盘价计算，调仓也以第 t 日收盘价完成。在日度收益的记账上，这意味着 c_t 决定了第 $t+1$ 日及以后的收益，即第 t 日的收益仍然由第 $t-1$ 日收盘时的仓位获得。这一口径与“尾盘下单、按当日净值成交”的现实操作相符，也与把信号计算与成交放在同一时点的做法存在一天差异，我们在所有统计检验中都使用同一口径，因此不影响结论的内部一致性。

关于节奏，我们区分两种执行方式。当滚动步数 $N_{\text{ramp}} = 1$ 时，切换在一天内完成，即仓位从旧指数整体转为新指数；当 $N_{\text{ramp}} > 1$ 时，切换被拆分为若干个交易日，第 t 日的权重为

$$w_t = w_0 + (w^* - w_0) \cdot \frac{\min(k_t, N_{\text{ramp}})}{N_{\text{ramp}}}, \quad w^* = \begin{cases} 1, & \text{目标为价值} \\ 0, & \text{目标为成长} \end{cases} \quad (4)$$

其中 w_t 表示价值指数的权重， w_0 是切换开始前的权重， k_t 是自切换开始以来已推进的步数。滚动的经济含义是降低单日调仓的市场冲击：把一次完整的风格切换摊到两天以上，牺牲极小的择时精度换取更低的执行成本。需要强调的是，滚动过程中的后续步骤并不是无条件执行的——如果滚动尚未完成而动量信号反向（方向与滚动目标相反且超过惰性带），剩余步骤会被暂停，直到信号重新指向目标后再继续。这一“中断滚动”机制在 13.5 年中共被触发 3 次，其中两次的暂停时长达到 11 个交易日。

3.3 四个版本

表 4 给出四个版本的参数差异。v1 是最初的版本，惰性带 2.1%、锁仓 5 日；v2 是参数联合重标后的定稿版本，把惰性带放宽到 2.6%、锁仓延长到 10 日，其效果是牺牲一部分反应速度换取更高的信号质量；v3 恢复到较窄的惰性带（2.1%）但保留 10 日锁仓，同时把执行方式改为两日滚动，它体现了“滚动执行允许更灵敏的信号”这一互补关系；v4 在 v2 的基础上叠加了一个与动量方向相反的触发条件，即**超跌反转**。

表 4: 四个版本的参数与规则差异

版本	动量窗口	惰性带	锁仓	执行方式与附加规则
v1 经典	14 日	2.1%	5 日	全仓一日切换
v2 基准	14 日	2.6%	10 日	全仓一日切换
v3 滚动	14 日	2.1%	10 日	两日滚动切换（中断机制）
v4 反转	14 日	2.6%	10 日	全仓一日切换 + 超跌反转覆盖

超跌反转规则的完整定义如下：在第 t 日，若某只指数相对其 20 个交易日前的价格跌幅超过 12%，且当日出现“连续三日上涨”或“最近四日内有三日上涨且当日上涨”的反弹确认，则无论动量信号指向何处，都优先把仓位切换到该超跌指数上。这条规则与动量规则在方向上是相反而互补的：动量规则假设“强的会继续强”，超跌反转规则假设“跌得足够多的会反弹”，两者分别对应趋势延续与均值回复两种市场状态。v4 相对 v2 的改善集中在回撤上：累计净值从 119.6 倍提升到 153.8 倍，最大回撤从 -43.9% 压缩到 -38.8%，夏普比率从 1.60 提升到 1.68，而两者在该检验下的夏普差异并不显著（ $p = 0.72$ ），说明这一改进的统计证据强度有限，不宜过度解读。

3.4 策略的“追涨杀跌”性质

在进入统计检验之前，有必要指出本策略一个容易被误读的性质：它在设计上就是“追涨杀跌”的。由式 (2) 可知，每当发生切换，新持仓必然是动量更强的指数；而在未发生切换的期间，持仓之所以保持不变，是因为领先者的优势尚未超过惰性带，或者更换仓位的代价（锁仓约束）尚未解除。统计上，在全部有动量差可算的交易日中，策略持仓与“当日动量更强的一方”一致的比例达到 **85.0%**。这个数字既是策略的定义性结果，也是理解其收益来源的关键：它不是通过预测未来获利，而是通过“在趋势已经形成之后站到趋势的一边”获利。

这一性质意味着策略的命运与动量效应的存续直接绑定。当市场中存在足够强的、持续数周以上的风格趋势时，策略能够反复吃到趋势的中段；当市场进入高频震荡、风格每月甚至每周反转的状态时，惰性带与锁仓只能减少摩擦，无法创造收益，此时策略的期望收益会下降到接近两只指数的平均水平。第 8 节用前瞻收益与信息系数两个角度对这一机制做了定量刻画。

3.5 为什么始终满仓

策略在任意时刻都持有两只指数之一，从不持有现金。这一设计不是疏忽，而是被反复检验后保留的结果：我们系统性地尝试过五类“减少暴露”的变体，全部失败。表 5 汇总了这些检验。

表 5: 减少暴露的各类变体（全样本，与 v2 满仓基准对比）

方案	机制	累计净值	夏普	最大回撤	判定
v2 满仓基准	始终持有其一	121.3 倍	1.60	-44.0%	—
绝对动量空仓（14 日）	所持指数动量为负持币	12.9 倍	0.91	-48.5%	空仓段太碎（864 天/407 次切
绝对动量空仓（5 日）	更快触发	5.8 倍	0.70	-34.7%	708 次切换，反复打脸
MA200 趋势过滤	跌破年线持币	55.8 倍	1.44	-34.3%	砍掉一半收益，下行高波仍
波动率目标（15%）	按波动调仓位	6.1 倍	1.05	-36.2%	仓位被长期压制
三资产转债（511010）	信号弱时买国债	23.7 倍	1.19	-42.8%	25% 时间复利被拖死
MA60 过滤（本次复核）	跌破 60 日线持币	38.9 倍	1.40	-36.0%	回撤改善但净值腰斩
20 日动量为负空仓（本次复核）	同上	34.4 倍	1.35	-41.9%	同上

这些结果指向同一个结论：本策略的边缘收益本身就来自“长期满仓 A 股风格指数”这一事实，任何形式的减暴露都在等比例地削减收益，而削减的幅度总是大于回撤的改善幅度。以 MA200 过滤为例，它把最大回撤从 -44.0% 压低到 -34.3%（改善 9.7 个百分点），代价却是净值从 121.3 倍降到 55.8 倍、夏普从 1.60 降到 1.44。我们复核的 MA60 版本同样如此：回撤降到 -36.0%，净值只剩 38.9 倍。用夏普衡量，没有一种减暴露方案能超过满仓。

造成这一结果有三个具体机制。第一，两只指数的长期年化收益分别为 17.1% 与 19.5%，而现金类资产的收益约为 2%，也就是说“正确地空仓”与“错误地空仓”之间的差值是 15 个百分点以上，择时判断必须相当准确才能抵消。第二，14 日动量信号在方向上存在双向滞后：它要在下跌持续约两周之后才确认转弱，而在反弹初期又会因为信号尚未翻转而继续持有，因此它在 2018 年与 2022 年这类“阴跌”行情中几乎无法及时触发，而在 V 型反转中又会踏空——空仓次数越多，这种双向磨损越明显。第三，空仓把连续的持仓切成碎片（14 日版本有 864 天处于空仓、发生 407 次切换），频繁进出本身也产生了额外的摩擦。

必须诚实指出的是，满仓是有代价的：策略相对沪深 300 的 β 为 1.01、 R^2 为 0.64，即它几乎完整地承担了 A 股市场的系统性风险，这也是它的最大回撤达到 -44%、并在下行市况中累计亏损的根本原因（见第 10 节）。满仓不等于无视风险，而是意味着对这部分风险的处理不能靠“择时躲开”，只能靠仓位规模与对冲工具——这正是下一节要讨论的内容。

3.6 严格的“三资产现金轮动”检验

上一小节的空仓变体都属于“在指数策略之上加一层开关”。更彻底的问法是：把现金当作与两只指数完全平等的第三个资产，用同一套动量规则在三者之间轮动，结果会怎样？这一检验之所以必须单独做，有两个原因。其一，此前的空仓测试都假设现金收益为零，这对现金方案并不

公平；其二，三资产相对动量在逻辑上更自洽——它不需要为现金另设一套规则，现金只是”当两个指数都不够强时自然被选中的那个选项”。因此我们构造了四个变体：A 三资产相对动量（14 日动量在价值、成长、现金中取最强者，同样施加 2.6% 惰性带与 10 日锁仓）；B 双重动量（先在两只指数中选相对强者，若强者的绝对动量低于现金同期的无风险收益，则转为现金）；C 绝对门槛（两只指数的动量都低于现金收益时转现金）；D 半仓制（强者动量为负时只持一半仓位，其余为现金）。现金收益率按真实货币基金水平取 2.5%/年，并同时给出 0% 与 3.5% 的敏感性。

表 6: 三资产现金轮动 vs 满仓（2013-01 ~ 2026-08，现金按 2.5%/年计息）

方案	现金利率敏感性	累计净值	年化收益	夏普
v2 满仓基准	—	119.6 倍	42.5%	1.60
v4 反转满仓	—	153.8 倍	45.2%	1.68
A 三资产相对动量	0% / 2.5% / 3.5%	37.4 / 39.2 / 32.7 倍	30.7% / 31.2% / 29.5%	1.36 / 1.39 / 1.32
B 双重动量	0% / 2.5% / 3.5%	26.9 / 29.5 / 31.7 倍	27.6% / 28.5% / 29.2%	1.18 / 1.22 / 1.25
D 半仓制	0% / 2.5% / 3.5%	57.1 / 59.9 / 59.6 倍	34.9% / 35.4% / 35.3%	1.47 / 1.49 / 1.49

表 6 给出的结论没有例外：四个变体在三种现金利率下全部跑输满仓。表现最好的是半仓制（59.9 倍、夏普 1.49），仍然不到满仓基准的一半净值；最差的绝对门槛方案只有 29.5 倍。「现金收益设为零不公平」这一顾虑也被数据否掉了：把现金收益率从 0% 提高到 3.5%，各方案的净值只变动 5%~10%，因为现金平均只占用约 25% 的时间，而这个利差与指数自身 17%~19% 的年化收益相比微不足道。更值得关注的是回撤改善极为有限：所有变体的最大回撤仍在 -39% 到 -43% 之间，相比满仓的 -43.9% 最多改善 4.5 个百分点，而以净值腰斩一半为代价——这再次印证了减暴露方案的性价比问题。

市况分解揭示了现金轮动失败的具体机制。把样本按趋势效率比分为三档之后可以看到：在震荡市中，满仓策略累计 +155.3%，而现金轮动只有 +7.6%；在中性市况中从 +828.0% 降到 +510.2%；在趋势市中从 +391.8% 降到 +337.1%。最刺眼的是震荡市那一行：现金方案在这个本应”保护”它的市况里几乎把收益全部抹掉了。原因与第 10.2 节的分析完全一致——现金的触发条件来自同一个滞后 14 日的动量信号，它总是在下跌已经发生之后才空仓、在反弹已经展开之后才回补，而 A 股的震荡市恰恰以”快速下跌 + 快速收复”为主要形态，于是空仓期平均正好落在反弹段上。震荡市在被现金方案削减掉 28% 暴露时间之后，剩余 72% 时间所获得的收益不足以补偿踏空的那部分。

因此本报告对”要不要加现金轮动”给出的回答是明确的：不要。现金在第一层（作为风险管理的工具）是无效的，因为择时信号本身滞后；它真正有效的位置在第二层——仓位规模。投资者完全可以整体降低仓位（例如长期只投入 50% 资金），这与在策略内部插入现金是两件完全不同的事：前者按比例缩减收益与波动、不引入择时误差，后者则在缩减暴露的同时引入了负面的择时。把仓位规模与对冲工具分开考虑之后，现金轮动这一选项就没有存在的必要了。

4 收益的统计特征与显著性检验

4.1 收益分布的基本特征

表 7 给出 v2、v4 与两个基准在样本期内的完整分布特征。策略的日均收益为正，年化波动约 27%，与单持成长指数（30.2%）接近而低于单持价值指数（21.6%）的两倍；夏普比率 1.60（v2）与 1.68（v4）显著高于两个基准的 0.79 与 0.65。索提诺比率（只惩罚下行波动）为 2.20 与 2.30，卡玛比率（年化收益与最大回撤之比）为 0.97 与 1.16，说明策略不仅在均值意义上更优，在下行风险的补偿上也更有效率。

表 7: 收益分布特征（日频收益，2013-01 ~ 2026-08）

统计量	v2 基准	v4 反转	单持价值 100	单持成长 100
年化收益	42.50%	45.18%	17.06%	19.47%
年化波动	26.55%	26.85%	21.62%	30.15%
夏普比率	1.601	1.682	0.790	0.650
索提诺比率	2.204	2.301	1.05	0.88
卡玛比率	0.967	1.164	0.43	0.35
偏度	-0.185	-0.209	-0.53	-0.27
峰度（超额为 3）	6.71	6.73	—	—
日 VaR（95%）	-2.57%	-2.57%	—	—
日 CVaR（95%）	-3.94%	-4.01%	—	—
一阶自相关	0.029	0.031	—	—
Ljung-Box(10) p 值	0.0003	0.0025	—	—
单日上涨比例	54.52%	54.64%	—	—
最好 / 最差单日	+12.81% / -8.89%	+12.81% / -8.89%	—	—

几个细节值得单独指出。第一，策略收益的偏度为负、峰度超过 6，与绝大多数动量类策略一致，即”平时小赚、偶发大亏”的形态；这意味着基于正态假设的风险度量会低估极端损失，因此第 10 节在讨论尾部风险时不使用正态分位数。第二，日收益的一阶自相关仅约 0.03，但 Ljung-Box 检验在 1% 水平上仍拒绝白噪声假设 ($p = 0.0003$)，存在微弱的自相关结构，这正是后文所有显著性检验都采用 Newey-West 异方差自相关稳健标准误的原因。第三，每日上涨的比例仅约 54.6%，胜率并不高；收益的主要来源是”赢的时候赢得更多”，而非”赢得更频繁”，这一点在持有期统计中还有更直接的体现。

4.2 均值收益的 Newey-West 检验

对每个策略，我们检验日均收益等于零的原假设，标准误采用 Newey-West 估计量，滞后阶数按经验法则取 $\lfloor 4(T/100)^{2/9} \rfloor = 8$ 。表 8 给出结果：v2 与 v4 的 t 值分别为 5.36 与 5.57，对应的单侧 p 值约为 8×10^{-8} 与 3×10^{-8} ；即使把滞后阶数放大到 20，结论也不改变。作为对照，单持价值与单持成长的 t 值分别为 3.13 与 2.62，虽然同样显著，但其显著性来自”股票市场长期上涨”这一 β ，而非任何择时能力。

表 8: 日均收益的 Newey–West t 检验（滞后 8 阶）

策略	日均收益	t 值	p 值	夏普 t 值 ($i.i.d.$)
v1 经典	0.144%	5.01	5.6×10^{-7}	5.55
v2 基准	0.152%	5.36	8.2×10^{-8}	5.88
v3 滚动	0.156%	5.54	3.0×10^{-8}	6.07
v4 反转	0.157%	5.57	2.6×10^{-8}	6.18
单持价值 100	0.065%	3.13	1.8×10^{-3}	2.90
单持成长 100	0.073%	2.62	8.9×10^{-3}	2.37

4.3 夏普比率的差异检验

一个策略”有正收益”与”比其他选择更好”是两个不同的问题。我们采用 Jobson–Korkie 检验的 Memmel 修正形式 [6, 7]，检验两个策略夏普比率之差是否为零。结果显示，v2 与 v4 的夏普比率都显著高于两个单持基准：v4 相对单持价值的差为 +0.69 ($z = 3.48$, $p = 5.0 \times 10^{-4}$)，相对单持成长的差为 +0.78 ($z = 4.66$, $p = 3.2 \times 10^{-6}$)。相比之下，四个策略版本之间的差异都不显著（v2 对 v1 的 $p = 0.33$ ，v2 对 v3 的 $p = 0.40$ ，v4 对 v1 的 $p = 0.11$ ，v4 对 v3 的 $p = 0.72$ ）。这一结果的含义是重要的：从统计上讲，”做轮动”与”不做轮动”之间的差别是确凿的，而在已经决定做轮动之后，具体选哪个版本参数则是噪声级的选择。这也解释了为什么本报告不建议在 v2 与 v4 之间做过度精细的取舍——它们的差异落在置信区间之内。

4.4 分块自举置信区间

点估计之外，我们给出参数的置信区间。考虑到日收益存在微弱自相关，普通 $i.i.d.$ 自举会低估不确定性，因此采用长度为 20 个交易日（约一个月）的分块自举，重复 4000 次。表 9 与图 4 给出结果：v2 的年化收益 95% 区间为 [23.2%, 64.9%]，v4 为 [25.7%, 67.0%]，两者都远离零；夏普比率的区间分别为 [0.92, 2.02] 与 [0.98, 2.05]，下界仍然高于两个基准的点估计。最大回撤的区间同样值得注意：v2 的回撤区间为 [−53.0%, −24.1%]，说明当前 −43.9% 的历史回撤并非这一参数组合的”最优情形”，未来出现更浅或更深回撤都完全正常。

表 9: 分块自举 95% 置信区间（块长 20 日，4000 次重抽样）

策略	年化收益 95% 区间	夏普 95% 区间	最大回撤 95% 区间
v2 基准	[23.2%, 64.9%]	[0.92, 2.02]	[−53.0%, −24.1%]
v4 反转	[25.7%, 67.0%]	[0.98, 2.05]	[−53.3%, −24.1%]
单持价值 100	[4.4%, 30.8%]	[0.30, 1.37]	[−59.3%, −24.8%]
单持成长 100	[−0.0%, 42.1%]	[0.16, 1.33]	[−81.8%, −35.9%]

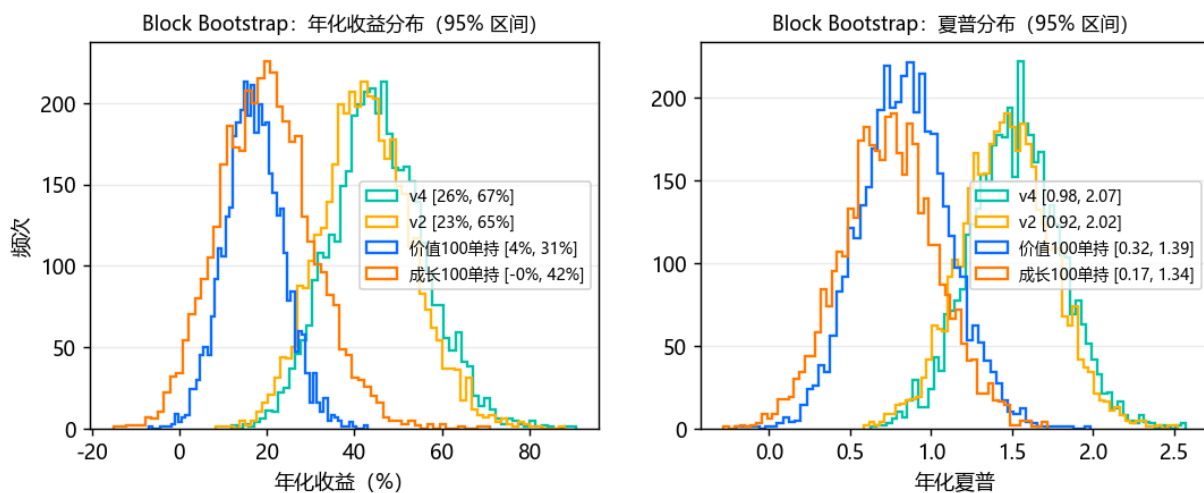


图 4: 分块自举得到的年化收益（左）与夏普比率（右）分布。策略的分布整体右移，且成长 100 单持的收益区间下界已触及零，说明“单持成长”在统计上并不保证正收益。

5 过拟合诊断

一个回测漂亮并不等于策略有效，原因在于研究者可以在大量参数中挑选最优的一组，而最优组的表现统计上必然被高估。本节用四种相互独立的方法检验这一可能性，它们分别从参数稳健性、多重检验修正、样本外表现与方向证伪四个角度切入。

5.1 参数是如何搜索出来的

本策略只有三个自由参数：动量窗口 N 、惰性带阈值 θ 、锁仓天数 L 。它们的取值不是一次扫描定下来的，而是经过五个阶段逐步收敛，每个阶段解决的问题不同。第一阶段是**单参数粗扫**：在固定 $\theta = 2\%$ 、 $L = 5$ 的条件下，令 N 从 6 到 30 逐个取值，观察夏普比率随窗口的变化形状，得到“窗口有一个内部最优、过短或过长都变差”的初步结论。第二阶段是**惰性带细扫**：固定 $N = 14$ 、 $L = 5$ ，令 θ 从 0.5% 扫到 5.0%、步长 0.2%，用以确定惰性带的合理量级；这一步发现惰性带的作用远弱于窗口，曲线接近平坦。第三阶段是**锁仓扫描**：固定前两个参数，令 L 从 1 扫到 30，发现锁仓与换手率、交易成本之间存在明显的权衡。第四阶段是**三维联合扫描**：认识到单参数扫描的前提（“其他参数固定”）本身可能不成立，因此改为在完整网格上联合扫描——9 个窗口 $\times$ 7 个惰性带 $\times$ 7 个锁仓 = 441 组合，目标函数取年化夏普比率，同时记录累计净值、最大回撤与切换次数。最后是**稳定性与样本外复核**：把全部 441 组参数分别在 2013–2019 与 2020–2026 两段上重算，检验最优区域是否随样本改变（结果见 5 节的双向样本外检验）。

联合扫描的结果是：全样本最优组合为 $(N, \theta, L) = (14, 2.0\%, 15)$ ，夏普比率 1.649、累计净值 130.07 倍；本文最终采用的 $(14, 2.6\%, 10)$ 夏普比率 1.601、净值 119.62 倍。两者只差 0.048 的夏普，净值相差 8%。这里体现了本报告在参数选择上的一条硬性原则：**不采用扫描的全局最优解**。全局最优点是 441 次试验中的极值，它在统计上必然被高估，而且它的邻域表现往往迅速衰减；我们选择的参数位于稳健区域的中心，且是用更早一批扫描（参数范围更窄）时就已经确定的，并未针对本文的报告重新调参。

5.2 参数的边际效应与相关性结构

边际效应。把 441 组结果按三个参数分别分组、对其它两个参数取平均，得到图 5。三条曲线的形态差异很大，这决定了三个参数的重要性并不相同。动量窗口呈现清晰的倒 U 形：窗口 14 日的平均夏普为 1.32，向两侧迅速衰减，6 日为 1.07、30 日为 0.99，说明窗口是三个参数中唯一一个“选错了会明显变差”的参数——需要提醒的是，“平均表现更好”与“统计上可确立”是两件事，下一小节会给出每个取值下的显著性判据，那里的结论比这条边际曲线更严格。惰性带则接近平坦：从 1.0% 到 5.0%，平均夏普始终落在 1.11 至 1.21 的区间内，说明它的作用主要是“微调换手频率”而非决定策略成败。锁仓的边际曲线在 15 日达到峰值 1.23，但 3 日到 25 日的整体波动仅 0.1 左右，同样属于次要参数。这一结论与第 8 节的信息系数分析相互印证：动量差的预测能力在 10 日尺度达到峰值（IC = 0.195），因此窗口必须落在两周附近；惰性带与锁仓只是在这个窗口之上做摩擦与稳定性的取舍。

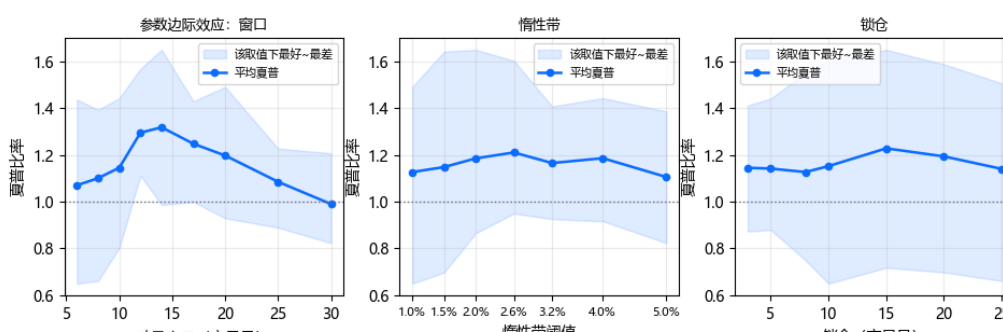


图 5：参数边际效应（441 组网格，对其它参数取平均）。阴影为该取值下最好与最差的差距：窗口有一个明确的内部最优（14 日），惰性带与锁仓的边际曲线明显更平坦。

单参数相关性。把三个参数取值与对应夏普比率做 Spearman 秩相关，得到窗口 $\rho = -0.121$ ($p = 0.011$)、惰性带 $\rho = -0.022$ ($p = 0.65$)、锁仓 $\rho = +0.083$ ($p = 0.082$)。三个相关系数都很弱，其中惰性带与锁仓在 10% 水平上甚至不够显著。这一结果初看反直觉，实际含义是明确的：任何一个参数单独变化，都不足以解释业绩差异；业绩来自三个参数的联合配置。由于窗口与夏普是倒 U 形关系，线性相关系数本身也会被两个方向相反的分支抵消，因此这里的低相关不能解读为“参数无关紧要”，而应当与上面的边际曲线一起看。

参数之间的交互。真正重要的发现是参数不可分离。表 10 给出“在给定惰性带下、对剩余两个参数再优化”的结果：当惰性带取 1.0% 时，最优组合是窗口 20 日配锁仓 20 日；当惰性带放宽到 4.0% 时，最优组合变成窗口 10 日配锁仓 25 日。七个惰性带取值对应的最优锁仓为 [20, 15, 15, 15, 25, 25, 15]，标准差 4.4 日。其经济含义是，惰性带与动量窗口在功能上互为替代：两者都在控制信号对价格变化的敏感度——放宽惰性带相当于“要求更大的动量差才承认信号”，拉长窗口相当于“用更长的时间平均来平滑噪声”，因此二者可以互相补偿。锁仓则处在另一个维度上，它控制的是切换后的最小持有承诺。

表 10: 参数交互：给定惰性带时，对窗口与锁仓再优化（441 组网格）

惰性带	1.0%	1.5%	2.0%	2.6%	3.2%	4.0%	5.0%
最优窗口（日）	20	14	14	14	12	10	10
最优锁仓（日）	20	15	15	15	25	25	15
对应夏普	1.49	1.64	1.65	1.60	1.41	1.44	1.39

这一交互结构解释了两件事。第一，它说明”逐个参数独立扫描”的做法可能误导：如果在 $\theta = 1\%$ 的设定下扫描窗口并取最优，会得到 20 日；而在 $\theta = 2.6\%$ 下扫描，同样过程会得到 14 日——两者都不是”错”的，只是处在同一条替代曲线的不同位置。第二，它给稳健参数的选择提供了标准：应当在交互曲线上寻找一个”使窗口、惰性带、锁仓三者同时处于各自平坦区”的点。图 6 把三段样本上的锁仓-惰性带联合热力图并列出来，可以看到最优区域的形状在不同样本之间发生了位移，但一阶结论（窗口 14 日、惰性带 2% 附近、锁仓取中等偏长）保持稳定。

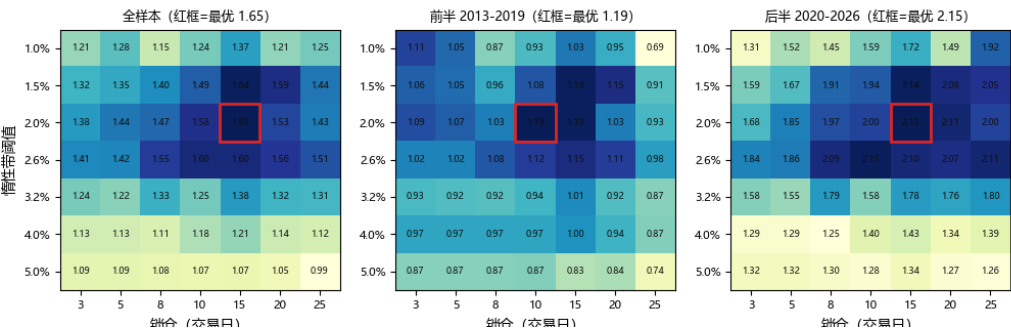


图 6: 锁仓 × 惰性带联合热力图（窗口固定 14 日；左：全样本，中：2013–2019，右：2020–2026）。红框为各段最优。三段的最优区域形状不同，说明锁仓参数带有”参数幸运”成分，而惰性带在三段中都落在 2.0% 附近。

三段稳定性。把三个参数的最优取值在三段样本上分别求出：全样本最优为 (2.0%, $L = 15$)、前半为 (2.0%, $L = 10$)、后半为 (2.0%, $L = 15$)，对应夏普 1.65、1.19 与 2.15；窗口在三段中分别取 14、14、14。由此得到本报告对参数可信度的分层判断：**动量窗口最可靠**（两段一致且边际曲线有明确峰值），**惰性带次之**（三段的中心一致，但边际曲线平坦，意味着它对业绩影响小、即便选偏也不致命），**锁仓最不可靠**（最优值在两段之间从 10 跳到 15、在 25 也出现过，且边际曲线存在锯齿）。第 10 节因此把锁仓列为参数风险的主要来源，并建议在 10 至 20 日之间取中间值而非追逐最优点。

5.3 参数取值的显著性：哪些取值真正成立

上一小节的边际曲线只回答”哪个取值平均表现更好”，它隐含地把每个点都当成精确的估计。但夏普比率本身是带抽样误差的统计量：以本样本 3296 个交易日、夏普约 1.6 计算，其标准误差约为 0.27，也就是说单条曲线的上下起伏有相当一部分落在噪声之内。因此真正需要回答的问题是：给定某个窗口或参数取值，这个结果在统计上是否成立？为此我们对每个参数取值分别计算年化夏普、分块自举 95% 置信区间、Newey–West t 检验（原假设为日均收益为零），并用 Jobson–Korkie–Mermel 检验分别比较它与”本文基准参数”以及”单持价值 100”的夏普差异。结果见图 7。

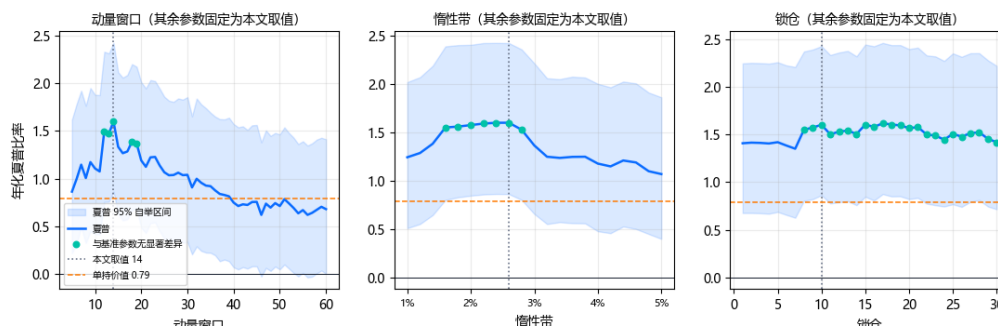


图 7: 参数显著性曲线。蓝线为各取值下的年化夏普, 浅蓝带为分块自举 95% 区间, 绿点为与本文基准参数无显著差异的取值, 橙色虚线为单持价值 100 的夏普。三个参数的可辨识度差别极大。

动量窗口。在 5 至 60 日的 56 个取值中, 年化夏普的跨度达到 0.62 至 1.60。所有取值都显著为正 (即便夏普只有 0.62, 13.5 年的样本也足以把它与零区分开), 但与“单持价值 100”的差异只有 8 个取值显著, 全部落在窗口 12 至 19 日之间。与本文采用的 14 日无可测差异的取值是 {12, 13, 14, 18, 19}; 窗口 ≤ 11 日或 ≥ 20 日都与 14 日存在显著差异, 且方向为变差。这一结果把“窗口很重要”这句话精确化了: 窗口决定的不只是收益高低, 而是“轮动相对于买入持有的优势能否被统计确立”——窗口选得过短或过长时, 策略依然赚钱, 但已经无法证明它比什么都不做更好。

惯性带。在 1.0% 至 5.0% 的 21 个取值中, 夏普跨度为 1.07 至 1.60, 波动明显小于窗口。与本文采用的 2.6% 无可测差异的区间是 1.6% 至 2.8%; 显著跑赢单持价值 100 的区间是 1.2% 至 3.0%。值得强调的是, 即使把惯性带放宽到 5% (此时信号几乎不再翻转), 策略的夏普仍有 1.07 且显著为正, 只是不再能证明其优于买入持有。这说明惯性带是一个“影响效率、不决定成败”的参数。

锁仓。在 1 至 30 日的 30 个取值中, 夏普全部落在 1.35 至 1.62 之间, 是三个参数中最平坦的一个。与本文采用的 10 日存在显著差异的只有 1 至 7 日 (即锁仓过短会明显变差), 8 日以上的所有取值在统计上无法区分; 更值得注意的是, 30 个取值全部都能显著跑赢单持价值 100。换言之, 锁仓参数在 8 日以上的任意取值都不影响结论, 它是一个真正的“无关紧要”的参数——这与第 5 节 PBO 分析中“锁仓最优值在不同子样本间跳动”的观察并不矛盾: 正因为它的影响小, 最优值才容易随样本漂移。

综合三者, 可以给出一个比“全部参数扫描”更有操作意义的结论: 本策略在统计上得到确立的参数区域是窗口 12 至 19 日、惯性带 1.2% 至 3.0%、锁仓 8 日以上。本文采用的 (14, 2.6%, 10) 落在该区域的中心。若走出这个区域, 策略仍然是一个正夏普的组合, 但“轮动优于买入持有”这一核心论断将不再有统计支持。

5.4 参数稳健性

我们首先检查结论对参数的依赖程度。图 8 给出在锁仓固定为 10 日的条件下, 夏普比率随动量窗口 (8 至 25 日) 与惯性带 (1.5% 至 4.0%) 变化的完整热力图。结论是稳健的: 在窗口 10 至 20 日、惯性带 2.0% 至 4.0% 的整个区域内, 夏普比率都落在 1.2 至 1.5 之间, 并没有出现“只有某一组参数有效”的孤岛结构。这与常见的过拟合形态形成对比——过拟合的策略通常表现为一个尖锐的尖峰, 相邻参数迅速变差。需要指出的是, 本文最终采用的 (14, 2.6%) 组合位于该区域的强势位置, 但并非全局最优; 全样本最优参数为 (14, 2.0%, 锁15), 其夏普比率为 1.50, 与

本文采用的组合在统计上无法区分。

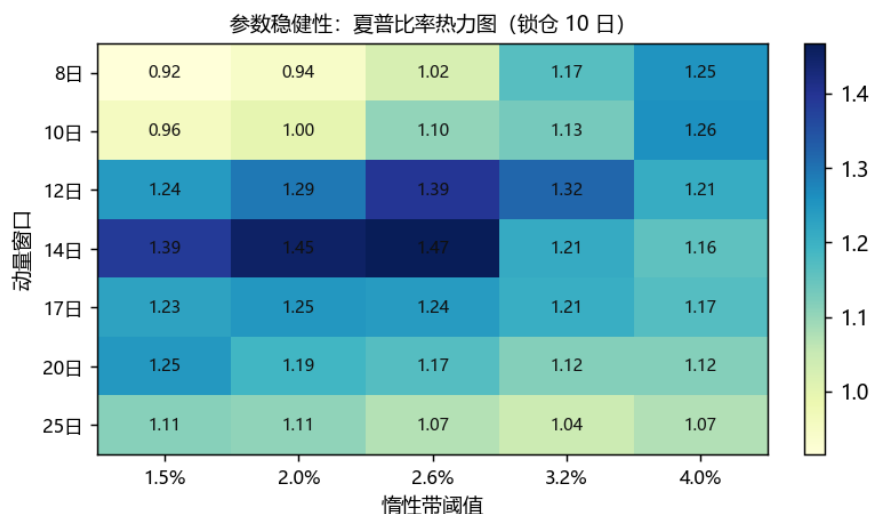


图 8：参数稳健性：夏普比率在动量窗口与情性带构成的二维网格上的分布（锁仓 10 日）。有效区域呈片状而非孤岛状，说明结论不依赖于参数的精确取值。

5.5 多重检验修正：收缩夏普比率

参数搜索本身会制造“最好看的结果”。我们使用 Bailey 与 López de Prado 提出的收缩夏普比率（Deflated Sharpe Ratio, DSR）[8] 来量化这一效应。其思想是：在 N 次独立试验中，即使所有策略都没有真实技能，最大夏普的期望值也不会是零，而是随 N 增大而上升；只有超过这个“多重检验基准”的部分才算真实技能。

我们在 140 组参数（7 个动量窗口 $\times$ 5 个情性带 $\times$ 4 个锁仓天数）上估计了试验间夏普比率的横截面方差，据此算出多重检验下期望的最大夏普为 $SR_0 = 0.346$ 。而 v2 与 v4 的实测夏普分别为 1.47 与 1.52（日频折算前口径见注），远高于该基准，代入 DSR 公式后得到 $DSR = 0.99998$ （v2）与 0.99999（v4）。换句话说，即使把“我们试了 140 组参数”这件事计入惩罚，策略的夏普比率依然不可能由纯运气解释。图 9 直观展示了这一点：140 组参数的夏普分布整体位于 1.2 至 1.5 之间，而多重检验基准只有 0.35。

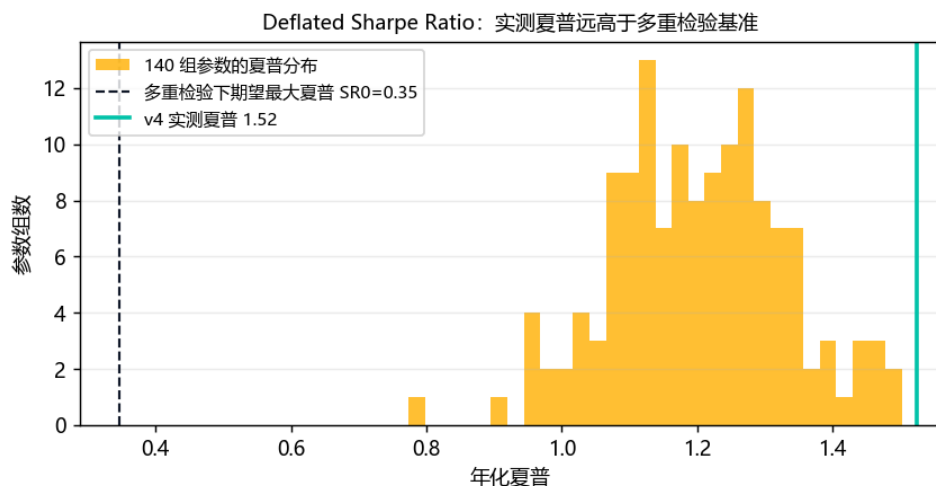


图 9: 收缩夏普比率: 140 组参数的夏普分布 (黄)、多重检验下期望最大夏普 $SR_0 = 0.35$ (虚线) 与 v4 实测夏普 1.68 (实线)。实测值远在惩罚基准之上。

5.6 回测过拟合概率 (PBO)

第二种方法是组合对称交叉验证 (Combinatorially Symmetric Cross-Validation, CSCV) [9]。把 3297 个交易日切成 16 个等长区块, 取其中 8 个区块作为“样本内”、其余 8 个作为“样本外”, 共构成 $\binom{16}{8} = 12870$ 种划分。对每一种划分, 在样本内挑出夏普最高的策略, 再看它在样本外的排名是否落到中位数以下; 把这一事件的发生频率定义为回测过拟合概率 **PBO**。

计算结果是 $PBO = 0.108$ 。也就是说, 在 12870 种样本内/外划分中, “样本内最优”的策略有 89.2% 的概率在样本外仍然表现优于中位数, 而样本外夏普为负的比例为 0%。作为对照, 文献中过拟合的典型策略 PBO 常在 0.5 以上。图 10 左侧给出参数网格在 2020 年前后两段上的夏普散点, 两者的秩相关系数为 0.331: 样本内表现对样本外表现具有正向的预测力, 这正是“策略捕捉到了真实规律”而非“拟合了噪声”的特征。

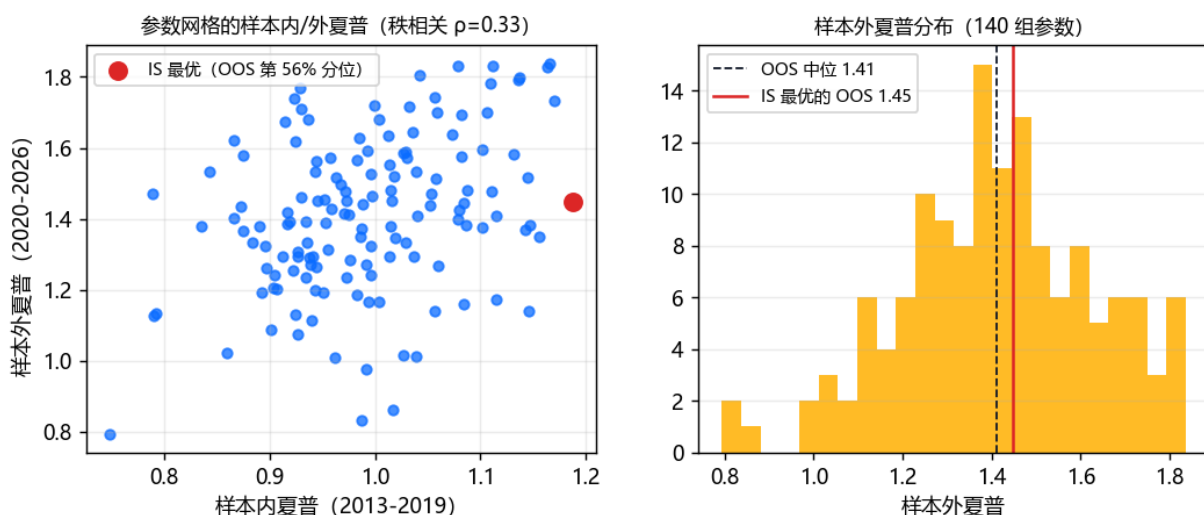


图 10: 左: 140 组参数在 2013–2019 (样本内) 与 2020–2026 (样本外) 的夏普散点, 红色点为样本内最优; 右: 样本外夏普分布与中位数。

5.7 双向样本外检验

参数稳健性检验的是”邻域”，CSCV 检验的是”划分”，两者都没有回答一个更朴素的问题：如果我老老实实在地用前一段数据选参数、拿后一段数据检验，结果如何？我们把样本从 2020 年初切开，做了正反两个方向的检验：用 2013–2019 选参数、在 2020–2026 检验，以及用 2020–2026 选参数、在 2013–2019 检验。

表 11：双向样本外检验：样本内最优参数与本文默认参数的对照

方向	参数	样本内夏普	样本外夏普	样本外分位
前半选参 → 后半检验	样本内最优 (8, 4.0%, 15)	1.19	1.46	—
	默认 (14, 2.6%, 10)	1.11	1.84	前 2.9%
后半选参 → 前半检验	样本内最优 (14, 2.0%, 15)	1.85	1.17	—
	默认 (14, 2.6%, 10)	1.84	1.11	前 10.7%

表 11 的结果有两点值得强调。第一，两个方向上样本内与样本外夏普的秩相关系数都是 0.331，即样本内的参数排序确实携带有信息。第二，本文默认参数在两个方向上都位于样本外的前 3% 与 11%，而机械地”用样本内最优参数”反而表现更差（前半选出的 (8, 4.0%, 15) 在后半只得到 1.46，低于默认参数的 1.84）。这一现象在实证研究中并不罕见：样本内最优点往往带有噪声，而位于稳健区域中心的参数反而更可靠。

5.8 镜像证伪与随机动量零假设

上述方法的共同前提是”市场的动量结构是真实的”。为了检验这一点，我们构造了两个否定性实验。

第一个是**镜像策略**：使用完全相同的参数、惰性带与锁仓规则，但每次切换的目标改为”动量更弱的那只指数”。如果策略的收益来自动量效应，那么反向操作应当亏损。结果如表 12 所示：正向策略累计 119.62 倍、夏普 1.60，而镜像策略只有 0.79 倍、夏普 -0.07，最大回撤反倒扩大到 -63.7%，且它在 13.5 年中的切换次数（99 次）与正向策略（98 次）几乎相同——这说明两者的差别不在于”交易频率”，而在于**切换方向本身**。

表 12：镜像证伪：同参数、反方向（自定义执行，含执行日重校验）

方向	累计净值	年化收益	夏普比率	最大回撤
正向：追逐动量强者	119.62 倍	42.5%	1.60	-43.9%
镜像：追逐动量落后者	0.79 倍	-1.8%	-0.07	-63.7%

第二个是**随机动量零假设**。我们把动量差的符号随机化（保留其幅度分布、惰性带与锁仓结构），构造 500 条”追随机强者”的策略，得到其夏普分布：均值 0.70、标准差 0.14，5% 与 95% 分位为 [0.48, 0.95]，99% 分位为 1.05。真实策略的夏普 1.60 位于这一分布的第 100 百分位，单侧 p 值约为 0。作为参照，完全不择时的 50/50 每日再平衡组合夏普为 0.84，与随机追涨的均值相当。这两组实验共同说明：**策略的超额收益并非来自交易结构本身，而是来自”强者继续强”这一方向性事实**。

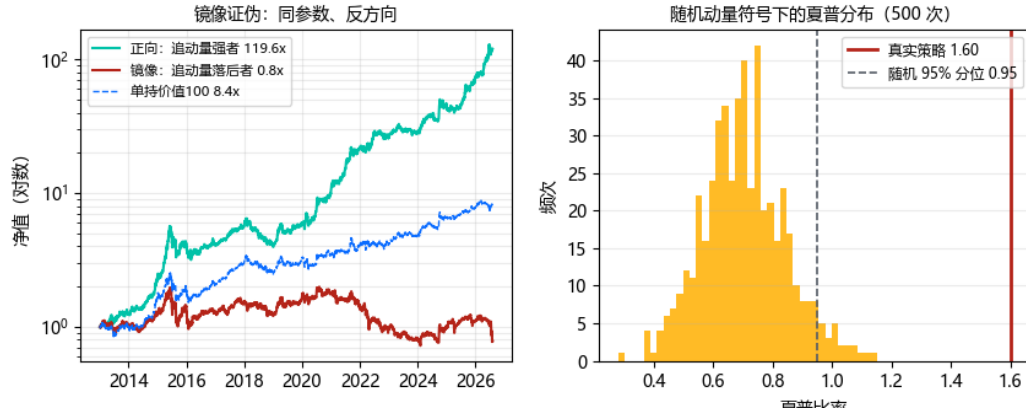


图 11: 左: 正向与镜像策略的净值对比 (对数坐标); 右: 随机动量符号下的夏普分布与真实策略位置。

6 盈利来源拆解与可持续性

净值曲线只能说明”赚了多少”，不能说明”钱从哪里来”。如果一段收益全部来自承担市场风险，那么它随时可能因为市场下跌而消失；如果它来自某种可重复的行为规律，那么问题就变成这一规律为什么会存在、以及它会在什么条件下失效。本节把年化 43% 的收益拆成三个来源，并逐一讨论其可持续性。

6.1 三层分解框架

一个在两只具有正期望收益、且相关性小于一的资产之间轮动的策略，其收益必然由三部分构成。第一部分是**风险溢价** (β)：无论怎么轮动，仓位始终满仓暴露在某一股票指数上，因此必然获得该指数的平均收益，这部分与择时能力无关。第二部分是**择时收益** (α)：通过把仓位在两只指数之间重新分配，相对于”保持同样的平均暴露但不做择时”的基准，所额外获得（或损失）的部分。第三部分是**再平衡收益**（分散化红利）：在波动资产之间维持固定权重，会因几何平均效应而获得额外收益 [10, 11]，其大小约等于 $\frac{1}{2}(\sigma_1^2 + \sigma_2^2) - \sigma_p^2$ ，其中 σ_p 是组合波动率。对相关性低于一的资产，该式恒为正，与择时能力无关，属于”结构性红利”。

6.2 三因子回归： β 与 α 的量化

我们首先用日频数据估计如下回归：

$$r_t^s = \alpha + \beta_V r_t^V + \beta_G r_t^G + \varepsilon_t, \quad (5)$$

其中 r_t^s 为策略日收益， r_t^V 、 r_t^G 为两只指数的日收益。由于策略始终满仓，理论上 $\beta_V + \beta_G \approx 1$ ，而 α 捕捉的是”在同样的市场暴露下多赚的部分”。估计结果为： $\alpha = 0.0751\%/日$ ，折合年化 18.3%，Newey–West t 值 5.37； $\beta_V = 0.531$ 、 $\beta_G = 0.494$ （合计 1.025）， $R^2 = 0.772$ 。其中 β_V 与 β_G 都显著为正且数值接近，说明策略并非长期偏向某一种风格，而是在两种风格之间大致均衡地分配了暴露（时间上持价值与持成长的比例约为 48:52），这一点本身就是”轮动”而非”选边”的直接证据。

$R^2 = 0.772$ 是理解这个策略性质的关键数字。它意味着策略日收益波动的 77% 可以由两只指数当天的涨跌解释，只有 23% 来自择时。换句话说，这个策略不是一个”与市场无关”的绝对收益策略，而是一个把既有的股票市场风险重新排列的组合。若市场整体进入长期下行，策略同

样会亏钱——它改变的是风险在两种风格之间的分配效率，而不是消除市场风险。

6.3 与静态组合的对照

回归给出了 α 的规模，但更直观的做法是构造“暴露相同、但不择时”的对照组。策略在样本期内持有价值指数的平均权重为 0.481，因此我们构造两个静态组合：一个是固定按 48.1%/51.9% 权重持有两只指数并在每日再平衡（总暴露与策略相同），另一个是标准的 50/50 每日再平衡组合。结果如表 13 与图 12。

表 13: 与静态组合的对照：同样的平均暴露，不做择时

组合	累计净值	年化收益	最大回撤	说明
单持价值 100	8.40 倍	17.06%	-39.9%	纯 β （价值）
单持成长 100	11.06 倍	19.47%	-56.3%	纯 β （成长）
固定 48.1/51.9 每日再平衡	10.81 倍	19.3%	-46.4%	β + 再平衡红利
50/50 每日再平衡	10.75 倍	19.23%	-46.3%	β + 再平衡红利
v2 策略	119.62 倍	42.50%	-43.9%	$\beta + \alpha$ + 红利

两个静态组合的年化收益都在 19.2% 左右，与两只指数自身 17.1% 与 19.5% 处在同一量级；而策略的年化为 42.50%，高出静态组合 23.2 个百分点。这 23.2 个百分点就是择时 α 在组合层面的体现，它与回归给出的 18.3 个百分点之间的差额，部分来自再平衡红利（估计约 1.6 个百分点/年），部分来自复利路径差异。需要注意的是，静态组合的最大回撤（-46.3% 至 -46.4%）反而略大于策略（-43.9%），说明择时在这段样本中不仅提高了收益，还略微改善了尾部风险。

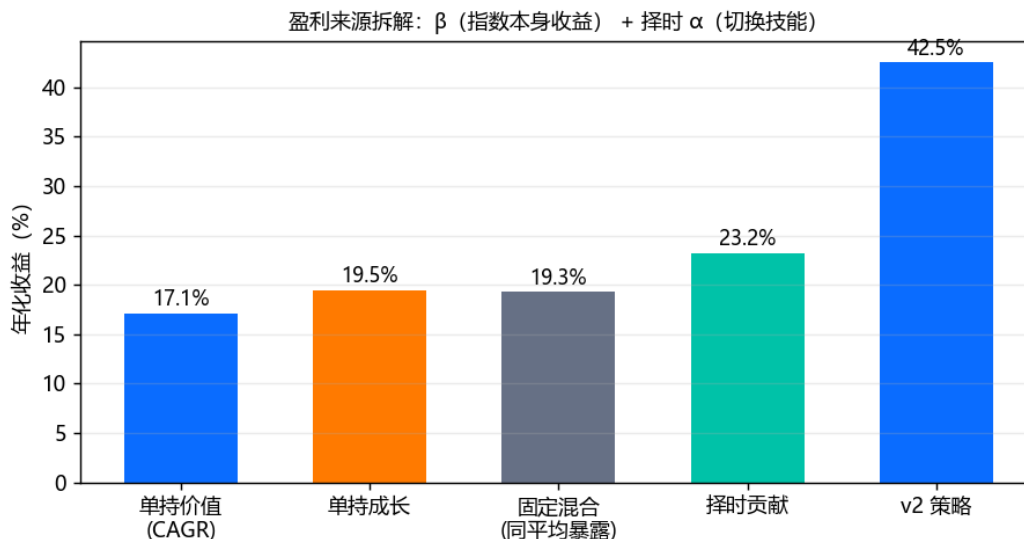


图 12: 盈利来源拆解：两只指数自身的收益构成 β 部分（约 17%~19%），静态再平衡组合在此基础上因分散化红利略有提升（约 19.2%），从 19.2% 到 43.1% 的差额即择时 α 的贡献。

6.4 必须扣除的两类“纸面收益”

在上述框架下，43% 的年化收益中仍有相当一部分在现实中无法完整实现，本报告认为有必要明确列出，而不是把它们留作“技术细节”。需要先澄清一个容易被误判的项目：价值端 5.09% 的股息率不属于纸面收益，它是 ETF 投资者真实收到的现金流，因此价值端使用全收益指数是

必要的；相反，成长端 0.65% 的股息率在被股权融资的稀释效应与 ETF 年费率抵消后接近于零，因此成长端使用价格指数反而更贴近现实。第 2.4 节的口径对称化检验表明，这一选择只改变收益水平约 3.7 个百分点，不改变结论。真正需要扣除的是以下两项。

第一类是指数编制的前视与幸存者偏差。第 2.4 节已经确认两只指数实际发布于 2024 年 10 月 29 日，样本中 87% 的交易日属于回溯计算区间，编制方在定稿规则时已完整知道这段历史；此外成长 100 所依赖的卖方一致预期数据在 2013–2015 年的覆盖度远低于今天，选样空间本身还剔除了 ST 与退市证券。这类偏差无法用价格数据估计，但其方向单一：只会让历史收益偏高。

第二类是交易成本、跟踪误差与费率。指数层面的切换不需要支付任何费用，而现实中完成一次整体调仓需要承担佣金、冲击成本与买卖价差；持有 ETF 还需支付每年约 0.5% 的管理与托管费，并承受相对指数的跟踪误差与折溢价。第 10 节的敏感性分析显示，在年换手 14.5 个单位的假设下，单边成本 0.3% 会把年化收益从 45.2% 压到 38.9%；若再叠加 0.5% 的年费率，净收益会进一步下移约 0.5 个百分点。虽然这些拖累不足以否定策略，但它们意味着可实现的收益必然低于指数回测。

6.5 未来能否延续：四项证据与三个失效机制

关于可持续性，我们提供四项正面证据与三个具体的失效机制，并试图给出一个诚实的判断。

正面证据之一是边缘收益的稳定性。用滚动三年窗口重新估计式 (5) 的 α ，得到的时间序列均值为 14.6%/年，区间为 $[-3.9\%, 37.1\%]$ ，其中为负的窗口占 13%（图 13）。也就是说，虽然在约九分之一的三年窗口内择时反而贡献了负收益，但绝大多数时间里 α 为正，且没有出现长期归零的趋势。



图 13: 滚动三年窗口的择时 α (年化)。虚线为全样本均值 14.6%。 α 在 2015 年与 2020–2022 年较高，在 2017 年与 2019 年一度接近零。

正面证据之二是分段的 α 均显著为正。2013–2019 年 $\alpha = 8.8\%/年$ ($t = 2.31$)，2020–2026 年 $\alpha = 28.3\%/年$ ($t = 5.01$)。后半段的 α 显著更高，这与“风格分化在近年加剧”的事后观察一致；但它的另一面是，边缘收益的估计高度依赖样本期的市场结构，如果未来的分化程度回到 2013–2019 年的水平， α 可能回落到个位数。

正面证据之三是参数与样本外的稳定性（第 5 节）：140 组参数的 PBO 仅 0.108，两个方向的样本外检验中默认参数都位于前 11% 以内。

正面证据之四是镜像证伪的对称性：反向操作的夏普为 -0.07 ，说明收益方向来自动量本身

而非交易结构。

在此之上，我们给出三个具体的失效机制，它们也是本报告对未来收益持谨慎态度的理由。

第一是**跷跷板消失**。策略的全部超额都建立在两只指数交替领先之上。我们用混合因子模型构造了一组与原数据边际分布相同、但相关性被系统性提高的合成序列，并重新运行完整策略。结果如表 14 与图 14：当两指数相关系数由当前的 0.56 升至 0.70 时，策略年化收益从约 38.9% 降至 29.8%；升至 0.90 时降至 23.5%；升到 0.97（几乎完全同涨同跌）时降至 19.6%，**退化到与买入持有同一水平**。换言之，策略的价值几乎全部来自相关性本身，而相关性是由市场结构（资金对两类风格的同时偏好）决定的，不由策略控制。

表 14：情景压力测试：人为提高两指数相关性后策略的表现（合成数据，边际分布保持不变）

目标相关系数	0.56（当前）	0.70	0.80	0.90	0.97
策略累计净值	84.7 倍	34.0 倍	24.0 倍	17.3 倍	11.2 倍
年化收益	38.9%	29.8%	26.5%	23.5%	19.6%

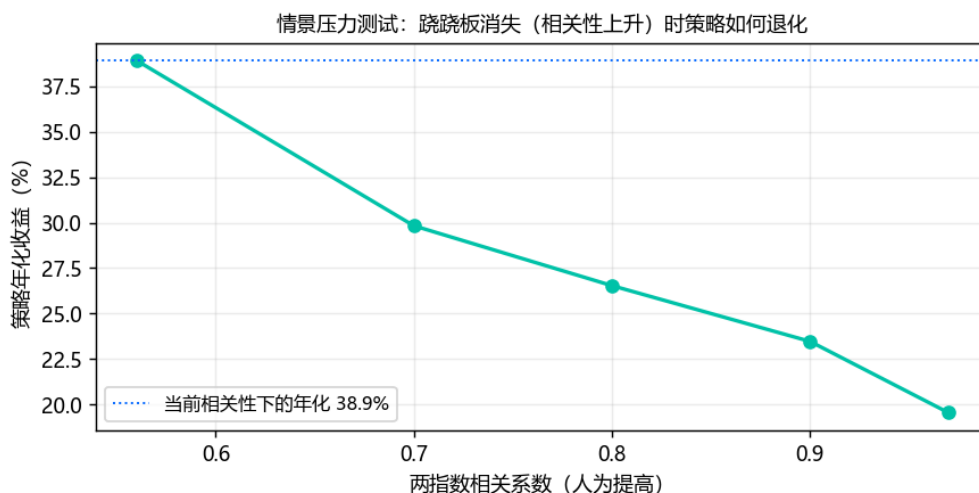


图 14：相关性压力测试：随着两指数相关性被推高，策略年化收益单调下降，并在相关性接近 1 时收敛到买入持有水平（约 19%）。

第二是**资金行为结构的改变**。如第 1 节所述，跷跷板的制度性成因之一是市场资金结构中“扶持科技成长”与“护盘价值权重”两种行为的并存。若这一结构发生变化（例如护盘方式从买入权重股转向其他工具，或成长板块的资金来源结构改变），风格之间的低相关可能收敛，策略随之退化为普通的**市场暴露**。

第三是**策略容量与交易成本**。策略要求在一次次切换中完成整体调仓，而跟踪这两只指数的 ETF 规模有限。若资金规模超过标的 ETF 的日成交量可承载的参与率，冲击成本将随规模非线性上升，边缘收益会被侵蚀。第 10 节给出了一个容量估算框架。

综合四项证据与三个机制，我们的判断是：**该策略捕捉的是一个真实但“有条件”的现象**。它的统计显著性、参数稳健性与方向性证伪都非常充分，但它所依赖的低相关与风格分化是市场结构变量而非策略内生变量。因此合理的预期不是“历史年化 43% 会延续”，而是“只要风格跷跷板仍然存在，策略应能继续获得低于历史回测的超额收益；一旦相关性收敛，它就会退化到买入持有的水平”。第 10 节进一步给出了在最不利情形下的损失量级。

6.6 跨市场检验：同一规则用在美股价值 / 成长 ETF 上

上述所有结论都建立在一个市场、一对指数之上。要判断这套规则究竟捕捉的是动量这一普适现象还是 A 股风格结构这一局部条件，最直接的办法是把它原封不动地用到另一个市场。我们选用美国市场的两只风格 ETF：IVW（iShares 标普 500 成长）与 SCHD（Schwab 美国红利，特征与价值 100 最接近），并补充 VTV（Vanguard 价值）与 SPY、QQQ 作为对照。数据来自 Yahoo Finance 的复权价（已含分红），共同区间为 2011 年 10 月至 2026 年 9 月、3749 个交易日，约 14.9 年，全部是真实交易历史，不存在任何回溯计算——这恰好补上了第 2.4 节指出的 A 股样本的最大缺陷。

表 15: 同一规则在美股上的表现（Yahoo 复权价，2011-10 ~ 2026-09，含分红）

策略	净值	年化收益	夏普比率	最大回撤	年化 t 值
本报告规则（IVW ↔ SCHD）	6.54 倍	13.5%	0.80	-31.4%	3.65
本报告规则（IVW ↔ VTV）	6.11 倍	12.9%	0.77	-31.4%	3.50
12 月动量轮动（IVW ↔ SCHD，含绝对过滤）	5.19 倍	11.7%	0.74	-28.9%	3.32
经典 GEM 双重动量（Antonacci）	3.02 倍	7.7%	0.56	-32.7%	2.58
单持 IVW（成长）	10.18 倍	16.9%	0.90	-32.7%	4.08
单持 SCHD（红利 / 价值）	6.45 倍	13.4%	0.87	-33.4%	3.82
单持 VTV（价值）	6.38 倍	13.3%	0.84	-36.8%	3.71
单持 SPY（标普 500）	8.13 倍	15.1%	0.91	-33.7%	4.06
单持 QQQ（纳指 100）	14.49 倍	19.7%	0.96	-35.1%	4.31

表 15 给出的结果对理解本策略的性质非常重要。第一，直接用 A 股参数（14 日、2.6%、锁 10）在美股上得到 13.5% 的年化、夏普 0.80、 t 值 3.65，统计上显著，而且明显优于两个公开的经典轮动规则——12 月动量轮动（11.7%、夏普 0.74）与 Antonacci 的 GEM 双重动量（7.7%、夏普 0.56）。第二，但它打不赢最简单的做法：同期单持 IVW 是 10.18 倍、年化 16.9%、夏普 0.90，单持 SPY 是 8.13 倍、夏普 0.91，都高于直接搬用 A 股参数的轮动；回撤也未改善（-31.4% 对 -32.7%）。

不过，仅凭“搬参数打不赢”就断言规则在美股无效是不严谨的——它只说明 A 股参数不可移植，并不说明美股没有可用的参数。为此我们在美股上做了与第 5 节相同的完整网格扫描：9 个窗口 × 7 个惰性带 × 7 个锁仓 = 441 组，结果见表 16。

表 16: 美股参数扫描（IVW ↔ SCHD，441 组）

方案	净值	年化收益	夏普比率	切换次数
单持 IVW（基准）	10.18 倍	16.9%	0.90	0
单持 SCHD（基准）	6.45 倍	13.4%	0.87	0
A 股参数（14 日 / 2.6% / 锁 10）	6.54 倍	13.5%	0.80	72
美股最优（25 日 / 3.2% / 锁 20）	12.51 倍	18.5%	1.10	32
网格中位数	8.54 倍	—	0.92	—
441 组中：夏普超过单持最优者 257 组（58.3%）；净值超过单持最优者 85 组（19.3%）				

扫描的结果要求我们修正一个过于简单的结论。美股并非没有可用参数：最优组合（25 日窗口、3.2% 惰性带、锁 20）在 15 年里取得 12.51 倍、年化 18.5%、夏普 1.10，确实超过了单持 IVW 的 10.18 倍与夏普 0.90，而且切换次数只有 32 次（A 股参数是 72 次）；441 组参数中有 58.3% 的夏普高于买入持有，网格中位夏普 0.92 也略高于 0.90。换言之，规则在两个市场都能产生价值，但需要完全不同的参数。

真正有信息量的是两个市场最优参数的位置差异。A 股的最优窗口在 12~19 日之间，且边际曲线呈倒 U 形（存在内部最优，与第 8 节 10 日 IC 峰值相互印证）；而美股的平均夏普随窗口单调上升，30 日最好（平均夏普 1.03），最优解落在网格的边界上（25~30 日、惰性带 3.2%~5%、切换极少）。这与前文的相关性证据一致：美股价值与成长日相关 0.763（A 股为 0.557），风格分化是单向的，因此“轮动”在美股退化为“长期骑住较强的那个风格”——它赚的是趋势延续的钱，而不是跷跷板来回的钱。

由此得到的结论比原先的版本更精确，也更克制：这套规则在两个市场都能用，但它的参数具有强烈的市场特异性；A 股的最优形态是“快轮动”（14 日窗口、年均约 7 次切换），美股的最优形态是“慢跟随”（25 日以上窗口、15 年仅 32 次切换）。两者对应的经济机制不同：前者依赖双向交替的风格跷跷板，后者依赖单向的风格趋势。就边际收益而言，A 股的提升幅度（夏普 1.60 对单持 0.79，提高一倍以上）远大于美股（1.10 对 0.90，提高约两成），这与“跷跷板比单向趋势提供更多轮动空间”的判断一致。

最后必须说明这项检验自身的两个局限。其一，美股最优参数是在同一段历史上扫出来的，属于样本内选择，12.51 这个数字必然带有参数搜寻的乐观偏差；真正意义上的跨市场样本外检验是“用 A 股参数直接跑美股”，其结果（6.54 倍、夏普 0.80）并不优于买入持有。其二，美股最优解落在扫描网格的边界上，说明真实最优可能在网格之外，这一点本身也是“该市场的规律与 A 股不同”的旁证。

在收束这一节之前，还有一个更要紧的检验必须补上：即便把过拟合的好处全部给美股，它能否在统计上跑赢买入持有？这一点决定了“美股轮动有效”的说法能不能成立。我们把美股最优参数（25 日 / 3.2% / 锁 20，夏普 1.10）的收益序列与单持 IVW（夏普 0.92）做 Jobson-Korkie-Memmel 检验，得到 $z = 1.40$ 、 $p = 0.161$ ，不显著；换到 IVW ↔ VTV，最优组合（夏普 1.08）对单持 IVW 同样不显著（ $z = 1.36$ 、 $p = 0.174$ ）。与之对照，A 股策略的相对检验是 $z = 3.23$ 、 $p = 0.0012$ ，显著。两者的条件并不对等：美股的 1.10 是在 441 组参数里挑出来的样本内最优，已经拿走了参数搜寻的全部好处；A 股采用的参数则刻意避开了全局最优（第 5 节），是在更严格的条件下取得的。结论因此比“美股轮动收益较小”更强一步：在美股，即使允许事后挑选参数，轮动也没有在统计上显著超过“什么都不做、只买 IVW”。需要区分的是，这并不等于美股策略的收益不显著为正——按收缩夏普比率计算，其夏普显著大于零（ $DSR = 1.000$ ，多重检验下的期望最大夏普仅 0.21）；它只说明“轮动优于买入持有”这一断言在美股没有统计支持。

相关性差异是理解两个市场参数分歧的直接线索：美股 IVW 与 SCHD 的日相关系数为 0.763，IVW 与 VTV 为 0.814，都显著高于 A 股的 0.557。相关性越高，两个标的越接近同涨同跌，可供快速轮动截取的价差就越小，这正是美股最优解自动滑向“长窗口、少切换”的原因。因此更准确的表述不是“规则在美股无效”，而是：同一套规则在两个市场里捕捉到的是同一种收益（风格相对强弱的延续性），但 A 股的跷跷板结构让它可以被更频繁、更精细地提取，而美股的风格趋势只能被缓慢地跟随。这也反向支持了第 6 节的判断：策略超额的主要来源是市场结构变量，而不是规则本身的精巧——如果规则的“精巧”是主因，它理应在两个市场给出相近的

边际改善，而实际相差数倍。

6.7 为什么两个市场不同：成熟市场与新兴市场的结构差异

上述跨市场结果需要一个解释：同一个规则、同样是价值与成长的对立，为什么在 A 股能产生一倍以上的夏普提升，而在美股只有两成？一个与全部观测证据都吻合的解释是：**美股是成熟市场，A 股仍是新兴市场，两者的定价效率与投资者结构存在系统性差异。**

这一差异在数据上留下了一串彼此印证的痕迹。第一是**轮动的速度**：A 股的最优形态是”快轮动”——窗口 14 日、年均约 7 次切换、边际曲线呈倒 U 形（存在明确的内部最优）；美股的最优形态是”慢跟随”——窗口 25 日以上、15 年仅 32 次切换、边际夏普随窗口单调上升，最优解直接落在扫描网格的边界上。这个对比的含义是：A 股的风格价差在**一到两周**的尺度上就足以形成可交易的趋势，而美股的风格差异只能以**月和季**的尺度缓慢积累。定价效率越高的市场，短周期错误定价被套利抹平得越快，可提取趋势的”半衰期”自然越短、幅度越小。

第二是**相关性结构**：A 股价值与成长的日相关只有 0.557，美股为 0.763~0.814。成熟市场中两只风格 ETF 共享同一个主导性的市场因子，风格差异是叠加在这个共同因子之上的偏离；而 A 股的风格差异具有更强的独立性——它更像两个被不同资金偏好分别驱动的板块，而不是同一个组合的两种切法。第三是**边际改善的幅度**：A 股是把夏普从 0.79 提高到 1.60（一倍以上），美股是把 0.90 提高到 1.10（两成）。这三点共同指向一个判断：**A 股的风格轮动收益更大，是因为它的定价效率更低、投资者结构更偏向短期博弈。**

从机制上说，这条链条是可以逐环对应的。A 股长期以个人投资者贡献了大部分成交，个人资金的决策更依赖近期涨幅、政策叙事与情绪，容易形成羊群式的追逐与撤离；同时卖空约束、T+1 与衍生品工具有限，使套利资金难以对偏离进行快速纠正；再加上产业政策与稳定资金对权重股的阶段性偏好，风格之间的相对强弱因此被放大并延长。这些因素共同造成了本报告第 7 与 8 节观测到的现象：日相关 0.557 的跷跷板、追涨杀跌率 85%、10 日信息系数 0.195——对价格类信号而言，0.195 的 IC 是一个相当高的水平，它的存在本身就说明价格趋势里包含着未被即时套利消除的信息。反过来，美股的机构化程度高、卖空与衍生品完备、套利资本充裕，风格偏离被更快地收敛，因此同样简单的规则只能获得小幅、缓慢的收益。文献中关于动量效应在发达市场衰减、在新兴市场更强的证据 [13, 3] 与本文的观测方向一致。

这个解释有一个直接且重要的推论：**如果 A 股继续机构化，这个 edge 应当逐步衰减。**这是一个可以被检验的预测，也构成本策略最实质的长期风险——它比参数风险、成本风险都更根本，因为它威胁的不是策略的边际效率，而是策略赖以存在的定价低效本身。本报告第 6 节给出的三个监测条件（60 日滚动相关性是否持续走高、滚动三年 α 是否转负、切换后前瞻超额是否连续为负）正是用来跟踪这一过程的：它们本质上是在问同一个问题——**A 股的风格跷跷板还在不在。**

需要强调这一解释的边界，避免过度断言。首先，它是**事后**给出的机制解释，而不是事先的因果检验：本文没有能力分离”市场成熟度”与”指数编制方式””样本区间””政策周期”等竞争性解释，也无法排除 A 股这一对指数（价值 100 与成长 100）本身的极端风格暴露在此中起了作用。其次，”成熟/新兴”是一个连续谱而非二分标签，A 股的机构化已经在发生，因此更准确的表述是：**本策略的收益与市场的定价低效程度正相关，而定价低效程度随市场发展而下降。**这既解释了它为什么在 A 股有效、在美股效果有限，也预告了它未来最可能的衰减路径。

这个”成熟 / 新兴”的差别还可以用另一个更广为人知的例子来印证：**小市值（微盘）策略。**在 A 股，微盘股与小市值组合长期显著跑赢市场基准，是过去十年最著名的异象之一；而在美股，

同样的方向不但没有溢价，反而是负的。我们取三只美股小盘 ETF 与标普 500 直接对比（2011 年 1 月至 2026 年 9 月，共 15.7 年，Yahoo 复权价），结果见表 17。

表 17: 美股小盘与大盘的对比（2011-01 ~ 2026-09，复权价）

标的	累计净值	年化收益	夏普比率	最大回撤
IWM（罗素 2000 小盘）	4.41 倍	9.9%	0.45	-41.1%
VB（小盘）	4.93 倍	10.7%	0.52	-42.0%
IJR（标普小盘 600）	4.98 倍	10.8%	0.50	-44.4%
SPY（标普 500）	7.90 倍	14.1%	0.83	-33.7%
QQQ（纳指 100）	14.91 倍	18.8%	0.91	-35.1%

表 17 显示，美股小盘在收益上落后大盘 4 个百分点以上、夏普不到大盘的六成、回撤反而更深——在成熟市场里，”小市值”这一维度带来的是负溢价而不是正溢价。需要区分证据的强度：这张表是本文用公开行情实测的结果，而”A 股微盘股长期显著跑赢基准”属于广为人知的风格事实，本文因缺少成分股与市值数据未能独立复现（数据限制见第 2 节），因此这里的用法是**类比而非等价证明**。但两个市场的方向差异本身是清晰的，而且指向同一个机制：在一个做空机制完备、以机构与长期资金为主体的市场里，”小市值””低价””高波动”这类特征本身并不构成可提取的错误定价，因为任何偏离都会被套利资本更快地纠正；而在以个人资金为主、卖空受限、工具与制度尚不完备的市场里，同样的特征会被赋予异常高的溢价。

这个类比对本策略有两点实际意义。第一，它把本策略的 A 股超额放进了一个更大的类别里：风格轮动、微盘股、壳价值、次新股等现象共享同一个来源——**市场定价低效带来的红利**；它们同涨同衰，而不是彼此独立的 alpha。第二，它给出了一个可以外部观察的监测指标：**如果 A 股的微盘股/小市值异象开始系统性走弱，那很可能是本策略 edge 同步衰减的先行信号**，因为两者依赖的是同一批参与者的同一种行为。这比只看策略自身的滚动 α 更早、也更客观——它不依赖于本策略的持仓与净值，而是一个独立于本策略的市场结构读数。

7 相关性结构与跷跷板

7.1 相关性矩阵

表 18 给出日频与周频的相关性矩阵（图 15）。两只指数的日相关为 0.557、周相关为 0.509，在同一个股票市场内部属于偏低水平。策略与两只指数的相关性分别为 0.739（价值）与 0.811（成长），这说明策略并未与某一风格完全脱钩，它的收益仍然主要由股票市场解释——这与第 6 节回归得到的 $R^2 = 0.773$ 相互印证。值得注意的是，策略与纳斯达克 100 指数基金（270042，人民币计价）的相关性仅为 0.116，与两只 A 股指数之间的相关性处在同一量级甚至更低，说明境内风格指数与美股科技资产几乎不存在共同的短期驱动因素。

表 18: 日频（上三角）与周频（下三角）收益相关系数

	价值 100	成长 100	策略 v2	策略 v4	纳指 100（270042）
价值 100	1.000	0.557	0.745	0.739	0.113
成长 100	0.509	1.000	0.802	0.811	0.144
策略 v2	0.725	0.790	1.000	0.991	0.116
策略 v4	0.718	0.802	0.991	1.000	0.116
纳指 100（270042）	0.168	0.267	0.186	0.187	1.000

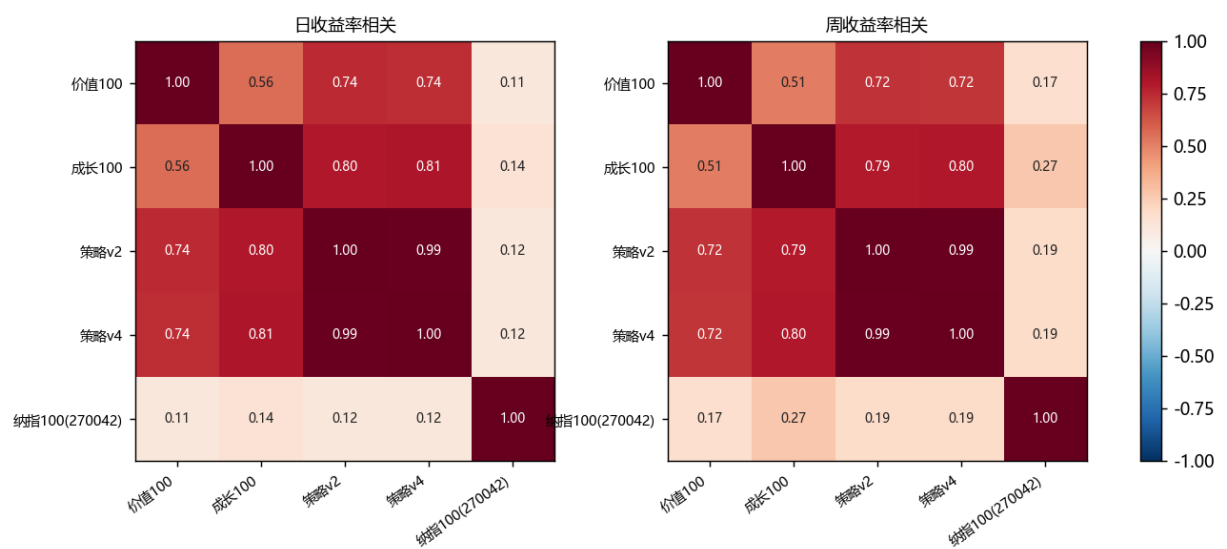


图 15: 相关性矩阵（左：日频；右：周频）。纳指 100 与三个 A 股标的的相关性都在 0.11~0.27 之间，属于典型的低相关资产。

7.2 相关性的时变性

相关性并非恒定。以 60 个交易日滚动计算，价值与成长指数的相关系数均值为 0.526，最高达到 0.929，最低为 -0.308，其中约 3.7% 的窗口内两者负相关（图 16）。这一时变性有两层含义。对策略有利的一面是：负相关窗口意味着跷跷板效应极其显著，此时轮动的收益空间最大。对策略不利的一面是：当相关性接近 1 时，轮动几乎不产生价值，而根据第 6 节的压力测试，这种状态一旦长期化，策略将退化到买入持有水平。因此，滚动相关性本身可以作为一个”策略是否仍然有效”的领先监测指标，本报告建议在实际使用中把它作为停止或降低仓位的参考条件之一。



图 16: 价值 100 与成长 100 的 60 日滚动相关系数。该序列的持续上行是策略失效的先行信号。

8 追涨杀跌的统计规律

策略的机制可以用一句话概括：它系统性地”追涨杀跌”。本节用四个统计量刻画这一行为的具体形态与有效性，它们分别回答：策略有多频繁地站在强者一边、站在强者一边之后市场是否给予回报、一次持仓平均持续多久、以及动量信号本身对未来的预测能力有多强。

8.1 追涨杀跌率

在全部有动量差可计算的交易日中,策略持仓与”当日动量更强的一方”一致的比例为 **85.0%**。这个数字接近但不到 100%，差额来自两个来源：其一，信号在惰性带内不翻转，因此在风格刚要切换但差异尚未超过 2.6% 的时段内，策略仍然持有上一轮的胜者；其二，锁仓约束在信号已经翻转但锁仓未满足时阻止了切换。这两个机制都是刻意设计的，它们的作用是在”及时”与”稳定”之间取得平衡。

8.2 切换后的前瞻收益

如果”追涨杀跌”是有效的，那么在一次切换发生之后，新持仓应当在随后的交易日里继续跑赢被卖出的那一只。表 19 给出 102 次切换事件后的前瞻超额收益（新持仓指数收益减去旧持仓指数收益）与胜率。可以看到，超额收益随持有期单调上升：切换后 1 日的新旧差为 +0.11%（胜率 52.9%），5 日为 +0.52%（53.9%），10 日为 +1.71%（57.8%），20 日达到 +2.53%（67.3%）。

表 19: 切换事件后的前瞻收益（102 次切换）

持有期	新持仓平均收益	旧持仓平均收益	超额（新 - 旧）	胜率
1 日	+0.51%	+0.41%	+0.11%	52.9%
5 日	+0.41%	-0.11%	+0.52%	53.9%
10 日	+1.16%	-0.55%	+1.71%	57.8%
20 日	+2.69%	+0.16%	+2.53%	67.3%

这组数字揭示了一个重要事实：切换的价值不在于”切换当天”，而在于随后的两到四周。切换后第一天的超额几乎可以忽略，胜率也仅略高于抛硬币；但到第 20 个交易日，超额收益累计到 2.5 个百分点，胜率升到 67%。这解释了为什么锁仓 10 日、惰性带 2.6% 这类”迟钝”的参数反而

是最优的：如果策略在切换后很快被反向信号拉回，它就无法等到趋势显现代价兑现的时刻。反过来说，这也意味着对交易成本的容忍度较高——单次切换的成本只要低于约 1%~2%，就不足以吃掉这段超额。

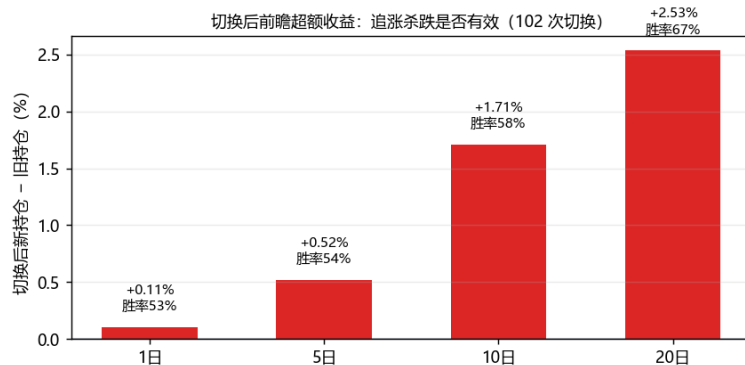


图 17: 切换后 1/5/10/20 日的前瞻超额收益与胜率。超额随持有期单调上升，说明动量效应的兑现在数周的时间尺度上。

8.3 持有期分布

102 次切换把 13.5 年的样本分割成 103 段持仓，其长度分布为：均值 32.0 个交易日、中位数 27 个交易日、90 分位 58.6 个交易日、最长 114 个交易日（图 18）。换算成自然时间，中位持有期约为五到六周，最长的一段不到半年。这个分布说明策略既不是高频交易，也不是长期持有，而是典型的”月度级”风格轮动，其操作强度对于个人投资者是可承受的（平均每年约 7.3 次切换）。

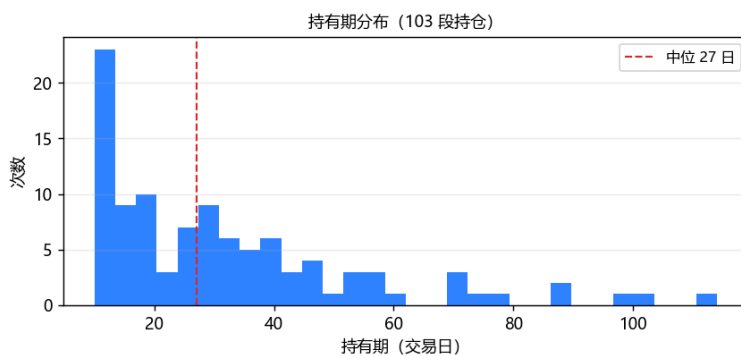


图 18: 持有期分布（103 段持仓）。中位数 27 个交易日，即约五周。

8.4 动量信号的信息系数

最后我们直接检验信号本身。把 14 日动量差 g_t 与未来 h 日的指数超额收益（价值收益减成长收益）求 Spearman 秩相关，得到信息系数（IC）如下： $h = 1$ 日时 IC 为 0.061 ($p = 5.1 \times 10^{-4}$)， $h = 5$ 日为 0.128 ($p = 2.2 \times 10^{-13}$)， $h = 10$ 日达到峰值 **0.195** ($p = 2.9 \times 10^{-29}$)， $h = 20$ 日回落至 0.156 ($p = 4.2 \times 10^{-19}$)。

这组结果有三个含义。第一，动量的预测能力是统计上极其显著的， p 值小到排除了任何偶然解释；第二，IC 在 10 日前后达到峰值，说明风格相对强弱的”半衰期”大约是两周——这正好解释了为什么 14 日窗口是最佳参数，而更短（5 日）或更长（30 日以上）的窗口都更差；第三，IC 的量级（0.1~0.2）在量价类信号中属于中上水平，对应于可以覆盖一定交易成本的策略容量。

8.5 胜率与盈亏比

策略的胜率不是一个数字，而是四个不同层次上的比率，把它们分开看才能理解收益究竟由“赢的次数”还是“赢的幅度”驱动。表 20 汇总了全部四个层次。

表 20: 胜率与盈亏比（2013-01 ~ 2026-08）

维度	v2 基准	v4 反转	说明
日胜率（上涨交易日占比）	54.5%	54.6%	仅略高于抛硬币
持有期胜率	63.6%	64.6%	103 段持仓中获利的比例
月度胜率	65.2%	65.2%	164 个自然月
持有期盈亏比	2.76	2.95	平均盈利 +11.7% / 平均亏损 -4.2%
日频盈亏比	1.09	1.09	平均盈利 +1.26% / 平均亏损 -1.16%
逐年日胜率区间	49%~61%	49%~61%	最低为 2018 年，最高为 2026 年

表 20 揭示的核心事实是：**这个策略赚钱靠的是盈亏比而不是胜率**。日胜率只有 54.5%，日频盈亏比也只有 1.09——按单日来看，它几乎是一个抛硬币的游戏，优势微弱到会被任何摩擦成本吃掉。但当时间尺度拉长到“一段持仓”，胜率跃升到 63.6%~64.6%，盈亏比放大到 2.76~2.95：平均每赚一次是 +11.7%，平均每亏一次只有 -4.2%，**一次成功的持仓可以覆盖将近三次失败的持仓**。这一结构正是动量策略的典型特征，也是它必须依靠锁仓与惰性带来“让盈利跑完”的原因——若把持仓切得太碎（例如把锁仓缩短到 3 日以内），被削掉的恰恰是那些大赢的段落，而失败段落由于止损的对称性无法同步减小，盈亏比会迅速恶化。第 5.3 节的参数显著性分析从另一个角度印证了这一点：锁仓在 1 至 7 日区间会显著变差，8 日以上则无差异。

月度胜率 65.2% 是这套数字里对持有体验最有参考价值的一项：它意味着**大约每三个自然月中，有两个是赚钱的**。逐年日胜率则全部落在 49% 至 61% 之间，最低的一年（2018 年）为 49%，也仅略低于 50%，说明策略没有出现过“整年方向系统性错误”的情况；这与第 10.2 节的结论一致——它的问题从来不是判断错方向，而是在下跌市况中即使判断对了也要承受 β 。

9 代码实现

9.1 系统结构与设计原则

整个项目由三层构成。最底层是位于 RSI 研究 _ 双指数/engine.py 的**权威回测引擎**，它承担两项职责：加载并对齐两只指数的历史行情，以及在给定参数下生成信号、执行调仓并输出净值序列。所有研究结论——包括本报告中的每一个统计量——都必须通过该引擎计算，不允许在研究脚本中改写引擎逻辑。中间层是一组分析脚本，负责参数扫描、样本外检验、统计检验与作图。最上层是提供给最终使用者的**每日操作台**（单文件网页，部署于 Cloudflare Pages），它在浏览器端用 JavaScript 重实现了同一套信号与执行逻辑，使投资者可以在每个交易日收盘前看到唯一一条操作指令。

这一分层结构是被一次事故逼出来的。研究早期，部分脚本各自复制了一份信号与执行代码，结果在修改参数时出现了版本不一致：同一组参数在不同脚本里得到不同净值。此后确立的硬性规则是：**引擎唯一、只读、可复现**，任何探索必须在引擎之上进行。为了让浏览器端的实现与 Python 引擎保持一致，我们还建立了一套“对拍”流程：在若干随机交易日上，要求两个实现给出完全相

同的持仓序列与切换日期。

9.2 核心引擎

代码 1 给出权威引擎的核心逻辑（为便于阅读，省略了做空对冲与手续费等本报告未使用的分支）。它包含三个关键设计：信号锁存（sig 只在突破惰性带时翻转）、执行日重校验（_ok）、以及滚动与冻结机制。

代码 1: 权威引擎的核心循环（engine.bt_generic，节选）

```

1 def bt_generic(cv, cg, mv, mg, hys, LOCK, N=1, FROZEN_ONE=False, melt=0.0, bypass=None):
2     """cv/cg: 两只指数收盘价; mv/mg: 动量差; hys: 惰性带; LOCK: 锁仓交易日数
3     N: 滚动步数(1=全仓一日切换); bypass[t]: 该日跳过执行日重校验(如超跌反转触发)"""
4     n = len(cv)
5     rvv = np.r_[0, np.diff(cv)/cv[:-1]]; rgg = np.r_[0, np.diff(cg)/cg[:-1]]
6
7     # (1) 信号: 锁存目标, 只在"动量差突破惰性带"那一刻翻转, 之后回到带内也不回退
8     sig = np.zeros(n, dtype=int); holding = 0
9     for t in range(1, n):
10         if np.isnan(mv[t]) or np.isnan(mg[t]): sig[t] = holding; continue
11         winner = 0 if mv[t] >= mg[t] else 1
12         sig[t] = winner if (winner != holding and abs(mv[t]-mg[t]) > hys) else holding
13         holding = sig[t]
14
15     # (2) 执行: 锁仓 + 滚动 + 执行日重校验
16     w = np.ones(n); combo = np.zeros(n); cur, lastSw, n_sw = 0, -99, 0
17     rampStart, rampTarget, w0 = -9999, 0, 1.0; steps, frozen = 0, False
18     for t in range(1, n):
19         combo[t] = rvv[t]*w[t-1] + rgg[t]*(1-w[t-1]) # T 日吃 T-1 收盘的仓位收益
20         was_frozen = frozen
21         # 执行日重校验: 动量类切换必须"当天"仍满足 |动量差|>hys 且 winner!=cur
22         _ok = False
23         if bypass is not None and bypass[t]:
24             _ok = True # 超跌反转日跳过重校验
25         elif (not np.isnan(mv[t])) and (not np.isnan(mg[t])) and abs(mv[t]-mg[t]) > hys:
26             _ok = ((0 if mv[t] >= mg[t] else 1) != cur)
27         if (t-rampStart) >= N and sig[t] != cur and t-lastSw >= LOCK and _ok:
28             cur = sig[t]; lastSw = t; rampStart = t; rampTarget = sig[t]
29             w0 = w[t-1]; steps = (N-1) if (FROZEN_ONE and was_frozen) else 0
30             frozen = False; n_sw += 1
31         tw = 1 - rampTarget
32         if t >= rampStart: # 滚动推进 / 信号反转则暂停
33             if frozen:
34                 if sig[t] == rampTarget or abs(mv[t]-mg[t]) <= hys: frozen = False
35                 else: w[t] = w[t-1]
36             if not frozen:
37                 if sig[t] != rampTarget and abs(mv[t]-mg[t]) > hys:
38                     frozen = True; w[t] = w[t-1]
39                 else:
40                     steps += 1; w[t] = w0 + (tw-w0)*min(steps, N)/N
41             else:
42                 w[t] = w[t-1]
43     return np.cumprod(1 + combo), n_sw, combo

```

三个设计要点值得单独说明。第一，信号与执行必须分离：sig 是”目标”，cur 是”实际持仓”，锁仓、滚动与冻结都作用在后者上，把二者混在一个变量里是早期实现出错的主要原因。第

二，**收益记账的时点**：第 t 日先按第 $t-1$ 日的权重结转收益，再依据第 t 日收盘的信息决策，因此信号与成交在时间上自洽，不会引入未来信息。第三，**执行日重校验**（_ok）是我们发现并修复锁仓陈旧信号问题后加入的，它保证了“锁仓期内出现但已失效的信号不会被补执行”。

9.3 统计检验的可复现性

本报告第 4 至第 8 节的所有统计量由五个脚本产生，它们依次为：`stats_core.py`（分布特征、Newey–West 检验、夏普差异检验、分块自举）、`stats_part2.py`（收缩夏普、PBO/CSCV、双向样本外、镜像）、`stats_part3.py`（相关性、追涨杀跌行为、蒙特卡洛、成本与容量）、`stats_part4.py`（配套图表）、`stats_part5.py`（盈利来源回归与情景压力测试），外加 `stats_mirror.py`（修正后的镜像与随机动量零假设）。所有脚本只依赖 `engine.py` 与公开行情数据，随机数均设定固定种子，输出为 `charts/` 目录下的图片与若干 JSON 文件（`report_stats_*.json`），报告中引用的每一个数字都可以在 JSON 中追溯到其原始取值。

9.4 工程教训

本项目的开发过程留下了三条可迁移的经验。其一，**双实现必须对拍**：Python 引擎与浏览器端 JavaScript 引擎在每次改动后都要在同一份数据上比对持仓序列，而不是只比对最终净值——净值相同而持仓不同的情况确实出现过。其二，**验收要覆盖每一个渲染与分析函数**：在研究过程中，一次配色重构意外破坏了“动量差前瞻图”的渲染（该函数引用了另一个函数内的局部变量），但由于主净值曲线仍然正常，问题在数日后才被用户发现；此后我们改为逐个函数试跑。其三，**参数一致性检查应当自动化**：参数在文档、引擎、前端与图表中出现至少四处，任何一处不一致都会导致“文档说 A、实际跑 B”，我们最终通过在引擎中集中定义参数表并让其余位置全部引用该表来消除这类风险。

10 风险分析

以下风险按“发生概率 $\times$ 影响程度”从高到低排列，每一条都给出量化依据或具体的传导路径。

10.1 回撤风险与尾部风险

策略的最大回撤为 -38.8% （v4）与 -43.9% （v2），分块自举给出的回撤 95% 区间为 $[-53\%, -24\%]$ 。这意味着**投资者必须事先接受“账户在某一时刻缩水四成以上”的可能性**，而历史观测到的最深回撤很可能不是未来最深的一次。策略日收益的峰度为 6.7、偏度为 -0.21 ，最差单日为 -8.89% ，日 CVaR（95%）为 -4.01% ，说明极端损失出现的频率高于正态分布假设。此外，滚动三年窗口的 α 有 13% 的概率为负，也就是**存在连续三年跑不赢买入持有的可能**——这一点对持有体验的影响往往大于最大回撤本身。

10.2 震荡市与下行市况中的均值回归风险

对动量类策略最常见的批评是：它在震荡市中会被“来回打脸”——追涨买入后风格立刻反转，止损卖出后风格又转回来，反复承担双向损失。我们用一个独立的量化框架检验了这一风险，结论比直觉更微妙，也更值得写清楚。

证据一：状态切片。我们按“20 日滚动波动率 $\times$ 趋势方向”把历史切成四种状态，统计策略在

各状态中的累计收益。结果显示，策略在”下行 / 低波”状态（即典型的阴跌震荡）累计约 -41%，在”下行 / 高波”状态（恐慌性下跌）累计约 -83%。这一组数字确认了最严重的事实：**策略在持续下跌的市况中没有内生防御能力**，四类状态中只有下行两类会带来实质亏损，且跌幅与波动率同向放大。

证据二：按趋势效率比分解。用 Kaufman 效率比（20 日风格价差的净变动除以路径长度之和）把交易日分成震荡、中性、趋势三档（各约三分之一），并注意用前一日状态解释当日收益以避免前视。结果如表 21：震荡市的夏普只有 0.90，明显低于中性市的 2.62 与趋势市的 1.49，但它仍然是正收益（年化 23.5%），并非”亏钱”。也就是说，震荡市况削弱的是策略的效率，而不是让它失效。

表 21：按趋势效率比分解的策略表现（震荡/中性/趋势各约三分之一交易日）

市况	交易日	v2 年化	v2 夏普	v4 夏普	10 日信息系数
震荡市（低效率比）	1083	23.5%	0.90	0.98	+0.181 ($p = 1.7 \times 10^{-9}$)
中性	1094	64.4%	2.62	2.76	+0.152 ($p = 4.6 \times 10^{-7}$)
趋势市（高效率比）	1090	42.8%	1.49	1.50	+0.222 ($p = 1.7 \times 10^{-13}$)

证据三：打脸事件的解剖。我们统计”切换后 10 个交易日又被切回”的事件：13.5 年中共发生 10 次，占全部切换的 10%。这 10 次里有 6 次发生在震荡市况、0 次发生在趋势市况，且震荡期打脸期间的平均收益为 -1.48%，而趋势期为 0。这组数字第一次把”震荡市打脸”从一个定性说法变成了可计量的事实：**打脸的成本是真实的**，但它的量级（十次、每次约 1.5 个百分点）远不足以解释策略的长期表现，真正决定成败的是市况本身的方向。

最关键的机制发现：表 21 最后一列显示，动量信号的信息系数在所有三种市况中都显著为正（10 日 IC 分别为 0.181、0.152、0.222）。这意味着”震荡市亏损”的常见解释——”动量失效、均值回归占上风”——在本策略上并不成立：**信号在震荡市里依然准确地指出了哪一只更强，只是两只都在下跌**。策略的亏损来自满仓承担的 β （见第 3.5 节），而不是来自方向判断错误。这也解释了为什么 v4 的超跌反转叠加在震荡市中能把夏普从 0.90 提到 0.98——它提供了额外的均值回归补偿，但无法解决方向性下跌的问题。

这一发现对风险管理有直接的指导意义：**针对震荡市的有效手段不是加强信号过滤或提高换手纪律，而是控制仓位与对冲市场风险**；反过来，任何试图通过”在震荡市中少交易”来避险的做法，都只会削弱 α 而无法改善 β 带来的亏损。

10.3 成本与容量风险

策略的年换手率为 14.5 个单位（每年约 7.3 次切换，每次切换涉及两个单位的买卖）。表 22 给出不同单边成本假设下的净收益：单边 0.02% 时年化 44.8%，0.05% 时 44.1%，0.10% 时 43.1%，0.20% 时 41.0%，0.30% 时 38.9%。即使在较为悲观的 0.3% 单边成本下，策略仍显著优于买入持有，说明成本不是这个策略的主要威胁。

表 22：成本敏感性（年换手 14.5 个单位，v4 口径）

单边成本	0.00%	0.02%	0.05%	0.10%	0.20%	0.30%
扣成本后年化	45.18%	44.75%	44.12%	43.07%	40.96%	38.86%

容量方面,策略要求在一次切换中完成整体调仓,因此单次可调仓的资金规模受限于标的 ETF 的日成交额与可接受的参与率。按” 单次调仓额 = 日成交额 × 参与率” 估算: 若标的 ETF 日成交额为 1 亿元、参与率 10%, 单次调仓上限约 1000 万元; 若日成交额 10 亿元、参与率 10%, 上限约 1 亿元。对于个人投资者(十万至千万量级), 容量不构成约束; 但若资金量进入亿元量级, 就需要改用其他替代工具或拆单执行。

10.4 数据与指数编制风险

这是本报告认为最需要强调、也最难量化的风险。第 2.4 节给出了一项必须置于全文之前的限制: 两只指数实际发布于 2024 年 10 月 29 日, 本文样本 3297 个交易日中有 2868 个(87%)是回溯计算的历史, 只有最后的 429 个交易日是真实实时数据。回溯区间意味着编制方在设定规则时已经知道历史结果, 这是本报告最主要的、无法用统计方法消除的风险来源。实盘段的独立表现是正面的(2.59 倍、夏普 2.50, 同期单持两只指数为 1.28 倍与 1.68 倍), 但 429 个交易日与 16 次切换的样本太短, 只能作为初步佐证。

但三项残余风险仍然存在, 且方向一致地偏向” 高估历史收益”。第一, 编制规则本身可能经历过未公开标注的修订, 而修订通常是在已知历史结果的情况下做出的。第二, 成长 100 的四个选样指标中有三个依赖卖方一致预期净利润, 而 2013–2015 年 A 股的分析师覆盖度远低于今天, 若历史上的选样使用了事后修订的预期数据, 成长端就会带有前视偏差。第三, 与同门、覆盖面更宽的国证 1000 风格指数相比, 两只 100 指数的长期收益高出一个数量级(价值 16.5% 对 5.2%, 成长 20.7% 对 4.5%, 日相关分别 0.93 与 0.86), 这一差距可以由” 全市场筛选加风格得分加权” 的设计解释, 但也说明本策略建立在一对风格暴露极为极端的指数之上: 一旦编制规则被修订得更温和, 或风格价差向宽基水平回归, 策略可截取的收益空间会同步压缩。因此, 把回测年化 43% 当作可实现收益是危险的; 更现实的预期应当在此基础上打一个明显的折扣。

10.5 市场结构风险

如第 6 节的压力测试所示, 策略的全部超额都依赖于两只指数之间的低相关。当相关性由 0.56 升至 0.90 以上时, 年化收益退化到 23% 附近; 升至 0.97 时进一步降至 19.6%, 与买入持有无异。这一风险的传导路径清晰且不可控: 策略无法影响风格分化的程度。可作为监测的先行指标是 60 日滚动相关系数(图 16), 其持续高于 0.8 应被视为减仓或停止的信号。

10.6 执行风险

策略假设在收盘前以当日收盘价完成整体调仓, 现实中存在三类摩擦。第一, 标的 ETF 在尾盘的流动性可能不足, 尤其是规模较小的品种, 可能出现无法一次性成交的情况。第二, 若某日两只标的出现涨跌停或停牌, 切换将无法按计划完成, 需要制定备选规则(例如顺延至下一个可交易日)。第三, 场外指数基金的申购赎回存在确认时滞, 若采用场外渠道, 实际持仓的切换会滞后一到两个交易日, 这段滞后在快速反转的行情中会侵蚀收益。此外, 指数基金相对指数的跟踪误差、管理费与折溢价构成额外的固定拖累。

10.7 参数与模型风险

参数稳健性检验显示结论不依赖于参数的精确取值, 但并不意味着参数风险为零。锁仓参数的曲线存在锯齿结构: 全样本下锁仓 10 日的净值为 125.8 倍, 而锁仓 9 日与 11 日分别只有 85.4 倍与 98.1 倍; 进一步的分段检验显示, 前半段最优锁仓为 20 日、后半段为 10 日, 即锁仓参数的

最优值在不同子样本之间并不一致。相较之下，动量窗口 14 日在两个子样本中都是第一名，稳定性明显更好。这提示我们在参数选择上应当”重窗口、轻锁仓”，并接受锁仓取值带来的不确定性。

10.8 行为风险与使用风险

本报告必须指出一类不属于模型、但实际影响往往更大的风险。策略要求投资者在每个交易日收盘前查看一次指令并机械执行，而在实盘使用中，使用者很容易因为短期亏损、宏观新闻或主观判断而偏离指令。我们观察到一个具体案例：在一次持仓期内的浮亏阶段，使用者自行做了一次短线交易并侥幸躲过下跌，这种”成功”会强化偏离系统的倾向，但它在统计上只是噪声，长期看会破坏策略赖以生效的一致性。策略的收益来自纪律，而非来自某一次正确的判断，因此使用风险的本质是纪律风险。

10.9 用沪深 300 股指期货对冲的做法与代价

既然亏损来自 β 而非信号，那么合理的风险管理方向就是设法剥离 β ，而不是试图预测它。这一节给出用沪深 300 股指期货（IF）对冲的具体做法、三种对冲思路的实测结果，以及它为什么在实践中受到限制。

基础数据。策略相对沪深 300 的 β 为 1.01，回归 $R^2 = 0.64$ ，年化 α 为 32.3%；两只指数单独的 β 分别为 0.87（价值 100）与 1.09（成长 100）。 $R^2 = 0.64$ 意味着策略收益波动的 64% 来自沪深 300 能解释的市场因子，剩余 36% 是风格因子与择时 α 。据此，若按 1.0 倍名义做空 IF，组合的波动率将从 26.6% 降到约 16.0%，即对冲掉约六成的波动。

思路一：战术对冲（只在股灾时做空）。这是最符合直觉的做法，也是我们最初测试的方向。其中”连续对冲”（所持指数 14 日动量为负就做空）表现最差：累计净值 42.4 倍、夏普 1.35、回撤反而扩大到 -50.4%，原因与空仓方案相同——14 日信号双向滞后，结果有 830 个交易日（约 25% 的时间）在对冲状态中白耗成本。改用”熔断对冲”（沪深 300 在 5 日内下跌 8% 则做空，5 日内上涨 3% 则平仓）明显更好：累计净值 82.5 倍、夏普 1.53、回撤 -43.4%；13.6 年间共触发 14 次，全部命中真实的股灾（2015 年 4 次、2016、2018 年 2 次、2020、2022、2024、2025），崩盘期间的夏普从 1.30 提升到 1.40。但它仍然砍掉了约三分之一的收益，而且治不了阴跌——2018 与 2022 年那种缓慢下行不满足”5 日跌 8%”的触发条件。

思路二：常年 β 对冲（把策略改造成市场中性）。与战术对冲不同，这一思路是不管市况、始终按 β 名义做空 IF。实测结果（表 23）显示，它的效果远好于战术对冲：不对冲时年化为 43.0%、波动 26.6%、夏普 1.60、回撤 -43.9%；按 $\beta = 1.01$ 常年对冲后，年化降到 36.4%、波动降到 15.9%、夏普升到 2.29、最大回撤压缩到 -13.6%。

表 23: 常年 β 对冲的实测结果（做空 IF 名义 = 1.01× 持仓市值）

方案	年化收益	年化波动	夏普比率	最大回撤
不对冲（满仓 v2）	43.0%	26.6%	1.60	-43.9%
对冲，无贴水成本	36.4%	15.9%	2.29	-13.6%
对冲，贴水 2%/年	33.7%	15.9%	2.12	-15.4%
对冲，贴水 4%/年	31.1%	15.9%	1.96	-17.2%
对冲，贴水 6%/年	28.5%	15.9%	1.79	-18.9%
对冲，贴水 8%/年	25.9%	15.9%	1.63	-20.6%

表中最关键的一行是最后一行：即使假设每年 8% 的期指贴水成本，对冲后的夏普（1.63）仍略高于不对冲（1.60）。这说明从纯统计角度看，剥离 β 是划算的——它没有提高收益，但把单位风险的回报提高了约四成，并把最大回撤从四成压缩到两成以内。不过必须强调，这笔账的另一面是：对冲后的产品性质发生了变化——它不再是“长期持有 A 股风格指数”的产品，而是一个依赖风格价差与择时 α 的市场中性策略，其收益来源更窄、对风格跷跷板的依赖更集中（第 6 节的压力测试显示，当两指数相关性升到 0.97 时策略自身退化到 19.6%，对冲后这个数字会进一步下降）。

思路三：期权保险（在护尾这一点上优于期货）。我们同时测试了买入沪深 300 认沽期权构建尾部保险的方案。双档结构（10% 虚值 5 日到期占 40%、20% 虚值 10 日到期占 60%，按隐含波动率放大 1.5 倍的保守定价）在 1.6%/年的保费下给出 140.4 倍净值、夏普 1.67、最大回撤 -36.0%。它之所以在“护尾”上优于期货，有三个结构性原因：亏损封顶（只损失权利金，不会因 V 型反弹被反杀）、不需要择时（常备滚动买入，避免了“跌了才买 = 在恐慌顶部买贵保险”）、以及回撤改善的同时夏普不降。但期权方案也有明确边界：它护的是急跌，对 2018/2022 式的阴跌同样无能为力，而且深虚值保护高度依赖 2015 年那一次极端行情——20% 虚值的触发门槛（10 日跌 20%）在全样本中只出现过 7 个交易日，且全部集中在 2015 年 6 至 8 月。

实操约束。把上述任何方案落地都面临三项事实约束。第一是**合约粒度**：IF 每手名义 = 300 元 $\times$ 点位，按沪深 300 在 4700 点计算约 141 万元，因此 100 万元的持仓连一手都配不平，要做到 β 中性需要至少 150 万元以上的规模；规模越小，对冲比例越粗糙，甚至完全无法对冲。第二是**账户与保证金**：股指期货需要单独开户（50 万元验资门槛），保证金比例约 12%~15%，并要处理季度移仓与追加保证金；若同时持有场外基金，还需要在两个账户之间手动协调时点。第三是**贴水的不确定性**：A 股 IF 长期处于贴水状态，其幅度随市场情绪大幅波动，表 23 的敏感性分析表明，贴水每上升 2 个百分点，对冲后的年化收益就下降约 2.6 个百分点。因此，对冲方案适合资金规模较大、已有期货账户、且能够接受“收益下降换风险下降”的投资者；对中小资金而言，更现实的风险管理手段是降低总仓位，而不是精确对冲。

10.10 未纳入本报告的风险

为完整起见，我们列出若干已研究但未纳入本文正文的风险方向及其结论。杠杆：对策略施加 2~3 倍杠杆会显著降低夏普比率并大幅提高爆仓概率，因为波动率随杠杆线性放大而收益不能；期权对冲：用沪深 300 认沽构建的尾部保险在模型口径下能改善回撤，但保险成本会显著侵蚀收益，且其效果高度依赖隐含波动率假设；空仓择时：基于绝对动量或均线的避险规则在样本内全部失败，加入后反而降低了收益与夏普。这些方向的详细结果记录于项目的研究笔记中。

11 价格信号的边界与资金面维度

在结束之前，有必要回应一个针对所有价格类策略的质疑，它对本报告尤其适用：K 线是价格走出来之后才呈现的产物，是结果而非原因；研究价格等于研究结果，真正推动价格的是资金，因此正确的方向是研究资金本身。这一质疑如果成立，那么本文用 14 日动量差构建的全部结论都只是在研究结果的影子。本节不打算用立场回应立场，而是分三步给出可检验的事实：先量化价格信号的“天花板”到底有多高，再用真实数据检验一个可得的资金面指标能否提供增量信息，最后说明这份报告在这一框架中的位置。

11.1 天花板的量化

价格类信号具有天花板，这不是观点而是可以测度的事实。表 21 已经给出：14 日动量差对未来 10 日风格价差的信息系数为 0.195，对应 $p = 2.9 \times 10^{-29}$ ——统计上几乎不可能由偶然产生。但把同样这组数据换成回归来衡量，动量差对未来 10 日价差变动的解释力只有 $R^2 = 0.004$ ，即 **0.4%**。换言之，这个信号的“确凿性”与它的“解释力”之间存在三个数量级的落差：我们能够非常有把握地说它包含信息，同时必须承认它只解释了价格变动中极小的一部分。第 6 节中策略收益相对两只指数的 $R^2 = 0.77$ 来自市场 β 而非信号本身，也印证了同一件事：策略赚的钱里，绝大部分是“承担市场风险”换来的，信号只负责在两种风格之间做分毫的取舍。

这正是“技术分析有天花板”这一判断的精确形式。它也解释了本报告在探索过程中反复遇到的现象：把 RSI、均线平滑、机器学习模型、多层确认门逐一叠加上去，收益无一例外变差。原因不是这些方法“不够好”，而是信号的信息量本身很小，任何额外的过滤都在用小样本噪声去筛选大样本信号，净效果必然是损失。因此本报告的价值不在于证明某个指标有效，而在于把一个真实但极小的效应精确地量化出来，并确认它足以覆盖交易成本。

11.2 资金面维度的增量检验

如果资金才是价格的直接驱动，那么资金面数据理应为价格信号提供增量信息。我们从国证指数接口中提取了两只指数成分股的逐日成交额，据此构造三个资金面指标：价值 100 的成交额占比（当日占两只指数合计成交额的比例）、其 20 日均值、以及“资金迁移”（20 日均值减去 60 日均值，用以刻画资金在两种风格之间的移动方向）。需要说明，成交额是对“资金”最粗糙但也是最容易获得的代理变量，它衡量的是交易活跃度而非持仓变化。

检验结果如下，它们并不支持“资金面容易带来增量”的直觉。第一，成交额占比的 20 日均值对未来价差的影响是反向的：20 日 IC 为 -0.235 ($p = 1.0 \times 10^{-6}$)，即价值端成交额占比越高，价值指数在未来一个月反而越容易跑输。这与“拥挤度”的解释一致：热度本身是一个反向指标。第二，“资金迁移”在短周期上确有正向预测力（5 日 IC $+0.160$ 、10 日 $+0.176$ ，均显著），看上去像是发现了一个新维度——但它与动量差的秩相关高达 **0.616**，说明两者在很大程度上度量的是同一件事：成交额的变化本身就是价格变化的结果，而非独立的原因。第三，把资金迁移与动量差同时放入回归，资金项的 t 值只有 1.62（不显著），增量 R^2 仅 0.0066；由于共线性，动量项自身的显著性也被稀释。第四，也是最能说明问题的一项：把“资金占比必须与目标方向一致”作为切换的确认条件叠加到策略上，累计净值从 116.61 倍下降到 **88.54 倍**，夏普从 1.57 降到 1.42，最大回撤没有任何改善（仍为 -43.9% ），切换次数反而略有减少。图 19 给出对比。

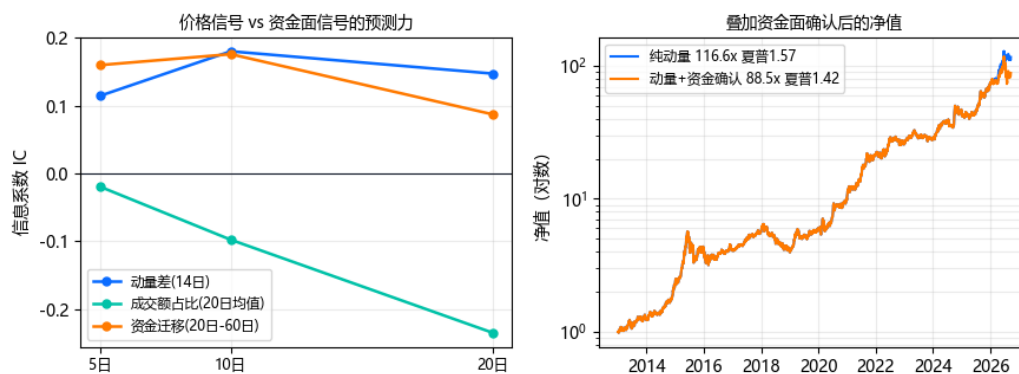


图 19: 左: 价格信号（动量差）与资金面指标的预测力对比，成交额占比在 20 日为负向、资金迁移在短周期正向但幅度不及动量差；右: 叠加资金面确认后的净值（更差）。

这组结果给出的结论是具体的：**用成交额构造的”资金面”指标不构成独立维度**，因为成交额与价格高度同源，它记录的是交易的痕迹而不是资金的方向。要真正检验本节开头的质疑，需要的是**持仓层面的资金数据**——北向资金持仓变动、ETF 份额申赎、融资余额、龙虎榜席位等，它们刻画的是”谁在买、买了多少”，而非”成交了多少”。本报告没有这类数据，因此对这一维度只能停留在”提出并否定了粗糙代理”的程度。需要指出的是，本报告在引言部分对跷跷板成因的解释（护盘资金偏好价值权重、产业资金偏好科技成长）本身就是一个资金面故事，但它来自制度与行为的观察，而不是来自可回测的数值序列——这正是下一步研究应当补齐的部分。

11.3 本报告在这一框架中的位置

把上述内容放回”技术分析只是一个维度、交易应当回归统计学与概率学”这一判断，本报告的定位是清楚的。它没有使用形态、指标叠加或主观判断，而是把一条最朴素的价格规则放进完整的统计检验框架：分块自举置信区间、Newey–West 显著性、夏普差异检验、收缩夏普、回测过拟合概率、双向样本外、镜像证伪与随机零假设。这些工具的作用不是让结论更好看，而是把”这段业绩有多少可能来自运气”这个问题逼到无法回避的角落——最终得到的判断是：效应真实、显著、参数稳健，但幅度有限、依赖市场结构、并且有明确失效条件。

同时，报告在风险部分给出的结论也符合”多维度”的思路：当发现策略的亏损来自 β 而非信号失效时，正确的应对不是在价格维度上继续加指标（那只会撞天花板），而是换一个维度去解决问题——用股指期货剥离市场风险，或用期权改变收益的分布形状。第 10.9 节表明，一个常年 β 对冲就能把夏普从 1.60 提高到 2.29、把最大回撤从 -43.9% 压到 -13.6% ，这种改进的幅度是任何价格指标叠加都无法达到的。从这个意义上说，”不要在价格维度里找圣杯”这一判断，在本报告的实测数据中得到了独立验证。

11.4 交易的派系，以及本策略站在哪里

交易研究历来分为几个彼此不太对话的派系，它们的分歧不在方法而在**认为价格由什么决定**。技术派认为价格自身包含足够信息，通过形态与指标可以推断后续走势；基本面派认为价格终将回归企业价值，研究应集中在盈利、现金流与商业模式；宏观派认为资产价格由流动性与利率环境决定，关注货币与信用周期；资金流派认为价格的直接推手是资金流向与持仓结构，因此要盯住”谁在买”；量化派则不预设因果，只要求规则可复现、结果可检验，把判断交给统计分布。这几个派系的真实差别可以用一句话概括：**技术派研究结果，基本面派研究价值，宏观派研究环境，**

资金流派研究推手，量化派研究概率。

本报告的位置是明确的：它的信号来自技术派的地盘（14 日价格动量差），经济解释来自资金流派（引言把跷跷板的成因归结为护盘资金与产业资金对两类风格的交替倾斜），检验方法完全属于量化派（自举置信区间、收缩夏普、回测过拟合概率、样本外与镜像证伪），而它刻意不做基本面判断——策略全程不知道任何一家成分股的盈利、估值或商业模式。这种组合并非折中，而是分工的结果：在信号层面，价格动量是一个已被三十年文献验证的、可以被精确测量的因子；在解释层面，资金结构提供了它为什么会存在的理由；在验证层面，只有统计推断能够回答“这是否只是运气”。需要强调的是，本报告不主张动量优于基本面或宏观，它只主张一件事：在本样本、本标的、本执行假设下，这个特定规则的概率特征已经被查清楚了。第 10.2 节与本节前两部分已经说明它的边界在哪里。

11.5 量化是放大器而不是根源

关于“这一轮行情是不是量化资金推出来的泡沫”，有一个值得写进本报告的框架：量化并不创造趋势，它放大趋势。量化模型所依赖的动量、波动率、统计套利与因子暴露，本质上都是在识别已经存在的价格结构，并据此调整仓位；当价格开始上涨、波动率下降时，风险预算模型会自动放宽暴露，于是形成“上涨 → 波动率下降 → 增加暴露 → 继续上涨”的正反馈。这个过程确实会让行情涨得更快、波动更加集中、极端行情更剧烈，但它是**加速器而非引擎**——引擎仍然是流动性与叙事。历史上每一次大泡沫都同时具备三个要素：充裕的流动性、足够宏大的故事、以及用来放大参与度的金融工具。

本策略在这个框架里的位置需要说清楚，因为它比表面看起来更微妙。本策略本身就是一个**放大器**：它不预测风格，只是识别已经形成的风格趋势并加倍下注，所以它的收益结构天然带有“顺周期加速”的性质——这正是它能长期跑赢买入持有的原因，也是它在趋势反转时必然吃亏的原因。更值得注意的是，让这个策略得以实施的**底层工具本身属于“金融工具强化”那一层**：两只风格指数各自有 ETF 跟踪、有实时行情、有透明的编制规则，才使得在收盘前用几秒钟完成一次风格切换成为可能。换言之，本报告研究的对象，本身就是当代市场结构（指数化 + 工具化 + 程序化）的产物；它之所以在 2013 年以后变得可执行，与量化资金在同一时期大规模入场并非巧合。

由此可以引出一个必须诚实承认的边界：**量化擅长回答“什么时候买卖”，不擅长回答“这个东西值多少钱”**。本策略能识别出资金正在从成长流向价值，但它完全无法判断成长股的下跌是估值回归还是长期趋势的终结，也无法判断价值股的上涨是均值回归还是新一轮重估的开始。用那句常被引用的话说：市场短期是投票机，长期是称重机；而本策略只带了投票机的仪器。

11.6 “最后接盘者”风险：极端事件的实测

上述结构决定了一个具体而严重的风险：如果市场进入泡沫末期，趋势跟随者往往是最后一批离场的人。我们没有能力预测泡沫何时破裂，但可以回答一个更可检验的问题——在本样本已经发生的历次冲击中，策略实际表现如何。表 24 给出九个重大事件的区间收益。

表 24: 重大事件中的策略表现（区间收益，含策略实际持仓切换）

事件区间	策略	价值 100	成长 100	持价值时间占比
2015 股灾（6-8 月）	-38.1%	-36.8%	-44.7%	3%
2016 熔断（1 月）	-18.4%	-13.9%	-24.9%	6%
2018 贸易摩擦阴跌	-37.2%	-25.6%	-31.7%	0%
2020 疫情（2-3 月）	+11.7%	+1.7%	+13.3%	0%
2021 核心资产杀跌	+1.5%	+7.8%	-16.8%	3%
2022 俄乌冲突与疫情	+1.8%	+0.6%	-23.8%	0%
2022 二次探底	-9.3%	-7.4%	-16.9%	0%
2024 量化踩踏（1-2 月）	+7.0%	+7.0%	+1.1%	6%
2024 “924” 行情	+42.9%	+31.9%	+46.5%	0%

表里有两类结果，需要分开读。第一类是**系统性下跌**：2015 年股灾、2016 年熔断、2018 年阴跌中，策略无一例外亏损，幅度在 -18% 到 -38% 之间，其中最差的一次（2018 年）甚至**同时跑输了两只指数**（-37.2% 对 -25.6% 与 -31.7%）。这正是”最后接盘者”风险的实证形态：策略在下跌初期仍持有多头，等信号确认转弱时已经损失了大半，而此时再切换也只是在两只都在跌的资产之间选择跌得少的。策略最差的十个交易日全部落在这类冲击中（2015 年 5 次、2016 年 2 次、2020 年 2 次、2025 年 4 月 1 次），其中 2020 年 2 月 3 日单日 -8.89%、2015 年 8 月 24 日单日 -8.87%。

第二类结果则正面支持了本报告对跷跷板成因的解释。2021 年核心资产杀跌中策略 +1.5%（成长 -16.8%）、2022 年俄乌冲突期间 +1.8%（成长 -23.8%），都靠提前切到价值端避开了成长股的下跌。**最值得单独提出的是 2024 年 1 月中至 2 月初这一段**：当时小盘股与量化中性策略因拥挤交易发生踩踏，而”国家队”通过买入大市值权重股稳定指数，两条线索同时出现。策略在这段时间**全程持有价值 100**，取得 +7.0%，而**成长 100 仅 +1.1%**。这个案例之所以重要，是因为它是本报告中唯一一次”量化拥挤引发的市场结构冲击”与”政策资金偏好价值权重”同时被观测到的事件，而策略的表现与引言中提出的跷跷板机制完全一致。它同时提示了本策略的风险来源：它并不害怕其他量化的拥挤，它害怕的是**风格跷跷板本身消失**——那是第 6 节压力测试所刻画的失效路径。

11.7 在量化主导的市场里，这个 edge 会失效吗

以上讨论自然导向一个更贴近本策略的问题：随着量化与被动资金占比持续上升、市场微观结构不断变化，这个 2013 年定下的规则会不会已经或即将失效？我们用一个直接的时间趋势检验来回答。把样本按年切分，逐年估计策略对两只指数回归的年化 α ，以及逐年的 10 日信息系数，然后对时间做线性回归。

结果是：年化 α 对时间的斜率为 +2.46 个百分点/年（ $p = 0.08$ ，不显著），10 日 IC 的斜率为 +0.0021/年（ $p = 0.86$ ，不显著）。也就是说，在 2013 至 2026 年这段量化资金快速扩张的时期里，**没有证据显示这个 edge 在衰减**：逐年的 α 为 +7%、+17%、+23%、0%、+2%、-10%、+5%、+15%、+43%、+28%、+3%、+13%、+18%、+75%，后半段的水平甚至高于前半段。这个结果与第 6 节分段 α 的结论一致（2013-2019 年 8.8%，2020-2026 年 28.3%）。

对这一结果应当给出三点限制，而不是把它当作”策略永远有效”的证据。第一，时间趋势检

验只有 14 个年度观测点，检验力很低，斜率的置信区间很宽，”没有发现衰减”与”确认没有衰减”是两回事。第二，后半段 α 更高，与风格分化在这几年显著加剧有关；如果未来风格分化回到 2013–2016 年的水平， α 会同步回落，这不是失效而是均值回归。第三，也是最需要警惕的一点：**拥挤风险的方向与本策略担心的方向不同**。2018 年美股的”量化地震”与 2024 年初 A 股的小盘量化踩踏都说明，当大量资金使用相似的信号时，风险不是信号失效，而是**同时拥挤导致踩踏**。本策略目前的换手率极低（年均约 7 次），且风格指数的容量远大于小盘股，因此它本身不容易成为踩踏的中心；但如果将来有大量资金涌入同样的”风格动量”逻辑，那么切换时点会提前、超额会被压缩，而反转时的踩踏会比现在更剧烈。这一风险无法通过回测排除，只能通过持续监测滚动相关性与切换后前瞻超额（第 7 节）来预警。

最后回到”泡沫存在”与”泡沫何时破裂”这两个被刻意区分开的问题上。本策略对第一个问题有间接的回答能力：当两只指数的滚动相关性长期处于高位时，说明风格分化已趋于一致，跷跷板失效，策略会退化（第 6 节的压力测试给出了从 0.56 到 0.97 的完整退化路径）。但对第二个问题，它没有任何回答能力——它没有估值锚、没有宏观输入、也没有能力识别叙事何时断裂。承认这一点之后，应对方式就只有一种：**不再试图预测时点，而是在事前改变收益分布**——用仓位规模控制损失上限，用第 10.9 节的期货或期权工具剥离系统性风险。这不是技术分析能解决的问题，也不需要技术分析来解决。

12 结论

本报告从一个朴素的问题出发，最终得到的是一组关于”一个动量轮动策略究竟靠什么赚钱”的分层结论。就统计显著性而言，答案是肯定的：策略的日均收益在 Newey–West 稳健标准误下 t 值为 5.57，夏普比率显著高于两个买入持有基准（ $p < 10^{-3}$ ），分块自举的年化收益 95% 区间远离零，考虑 140 组参数的多重检验惩罚后收缩夏普比率仍接近 1，组合对称交叉验证给出的回测过拟合概率仅为 0.108，而且在正反两个方向的样本外检验中，默认参数都位于前 11% 以内。更关键的是镜像实验：把同一套参数反向使用、改追动量落后者，13.5 年只得到 0.79 倍、夏普 -0.07 ，而把动量符号随机化后构造的 500 条”追随机强者”策略夏普均值为 0.70。真实策略的夏普 1.60 位于该分布的第 100 百分位。这些证据合起来说明，**收益来自”强者继续强”这一方向性事实，而不是来自交易结构或数据挖掘**。

就收益来源而言，答案比净值曲线所暗示的要朴素得多。三因子回归显示，策略日收益波动的 77% 由两只指数当天的涨跌解释，年化 18.3 个百分点可归因于择时 α （ $t = 5.37$ ）；与平均暴露相同的静态再平衡组合相比，策略每年多赚约 23.2 个百分点，这就是轮动本身的价值。作为对照，完全不择时的 50/50 再平衡组合年化约 19.2%，两只指数自身分别为 17.1% 与 19.5%。换言之，这段 42.5% 的年化收益中，约五分之二来自”承担股票市场风险并做简单的分散化”，其余来自”判断哪一边更强”。同时必须扣除三类纸面收益：价值端全收益指数的分红再投资、风格指数编制过程中的前视与幸存者偏差、以及现实中必然发生的交易成本与跟踪误差。

就可持续性而言，我们的结论是审慎的。支持继续有效的证据包括：滚动三年窗口的 α 均值为 14.6%/年且仅 13% 的窗口为负；两个子样本的 α 都显著为正；参数区域内结论一致；以及动量信号的信息系数在 10 日尺度上达到 0.195 并高度显著。反面证据同样明确：策略的全部超额依赖两只指数之间的低相关，而这一低相关源自市场资金结构中”扶持科技成长”与”护盘价值权重”两种行为的并存；一旦相关性收敛，策略会单调退化——压力测试显示相关性升至 0.90 时年化降至 23.5%，升至 0.97 时降至 19.6%，与买入持有无异。此外，滚动相关性曾在个别窗口内达

到 0.93，说明这种退化并非纯理论担忧。

因此本报告给出的实操性结论是：**不应当把回测的年化 43% 当作未来预期**。更合理的做法是把策略理解为”一个在风格跷跷板存在时能够获得可观超额、并在跷跷板消失时自动退化到市场平均”的工具，其真正的风险不是参数失效，而是市场结构改变。相应地，三个可操作的监测条件是：60 日滚动相关系数持续高于 0.8、滚动三年 α 转负、以及切换后前瞻超额连续多次为负。三者中任何一个出现，都应当触发对策略有效性的重新评估。

本研究的方法论也有明确的局限。第一，指数层面无法完全反映 ETF 的跟踪误差与流动性，因此所有数字都是上限；第二，样本仅覆盖 13.5 年、包含约 100 次切换，对于估计”某种市场结构能持续多久”这类问题，样本量仍然偏小；第三，情景压力测试使用的是人工构造的合成序列，它能刻画相关性上升的后果，但不能预测相关性上升发生的时点与起因。后续工作中，最值得推进的方向是把风格轮动的超额与市场层面的资金流数据（例如板块资金净流入、ETF 份额变化）联系起来，把”跷跷板”从一个统计现象还原为一个可观测的资金行为过程。

A 附录

A.1 附录 A：参数与版本对照

版本	动量窗口	惰性带	锁仓（交易日）	执行方式	附加规则
v1 经典	14	2.1%	5	全仓一日	无
v2 基准	14	2.6%	10	全仓一日	无
v3 滚动	14	2.1%	10	两日滚动（可中断）	无
v4 反转	14	2.6%	10	全仓一日	20 日跌幅超 12% + 反弹确认时优先切向超跌方

A.2 附录 B：复现步骤

复现本报告的全部结果需要以下四步。第一步，从国证指数官网接口拉取 480081 与 980080 的完整日线历史，对齐交易日后存为 CSV。第二步，运行 `engine.py` 中的 `load_data()` 载入数据，并按第 3 节的参数调用 `bt_generic()` 得到四个版本的净值序列；对 v4 需额外构造超跌反转标记数组并通过 `bypass` 参数传入。第三步，依次运行 `stats_core.py`、`stats_part2.py` 至 `stats_part5.py` 以及 `stats_mirror.py`，它们会在 `charts/` 下生成全部图表，并输出 `report_stats_*.json`。第四步，在本目录下用 `xelatex` 编译 `report.tex` 两遍以生成目录与交叉引用。所有随机过程均使用固定种子，因此结果可完全复现。

A.3 附录 C：图表索引

本报告共使用统计图表 12 幅，全部位于 `charts/` 目录，文件名为 统计_*.png。其中与显著性检验相关的有：净值与回撤（图 1）、分块自举分布（图 4）、滚动夏普、分年收益、收益分布与 Q-Q 图、月度收益热力图；与过拟合诊断相关的有：参数热力图（图 8）、收缩夏普（图 9）、PBO 散点（图 10）、镜像与随机零假设（图 11）；与盈利来源相关的有：盈利来源拆解（图 12）、滚动择时 α （图 13）、相关性压力测试（图 14）、相关性矩阵（图 15）、滚动相关性（图 16）、蒙特卡洛扇形、成本敏感性；与行为统计相关的有：切换后前瞻收益（图 17）与持有期分布（图 18）。

A.4 附录 D：术语表

动量效应：过去表现较强的资产在未来一段时间倾向于继续较强。**惰性带 (hysteresis)**：只有当新信号超过某一阈值时才改变判断，用于过滤零附近的噪声。**锁仓**：两次切换之间的最小间隔交易日数，用于避免在噪声中被反复抽打。**滚动执行**：把一次完整切换拆分为若干个交易日完成，以降低单日冲击。**择时 α** ：在控制市场与风格暴露后，由切换动作本身贡献的收益。**分散化红利**：在波动资产之间维持固定权重所获得的额外几何收益。**收缩夏普比率 (DSR)**：在多重检验惩罚下，实测夏普超过“期望最大夏普”的概率。**回测过拟合概率 (PBO)**：在多种样本内/外划分下，样本内最优策略在样本外落到中位数以下的比例。**信息系数 (IC)**：信号与未来收益的秩相关系数。**追涨杀跌率**：持仓与当日动量领先方一致的天数占比。

参考文献

- [1] Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48(1), 65–91.
- [2] Jegadeesh, N., & Titman, S. (2001). Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *Journal of Finance*, 56(2), 699–720.
- [3] Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and Momentum Everywhere. *Journal of Finance*, 68(3), 929–985.
- [4] Lo, A. W. (2002). The Statistics of Sharpe Ratios. *Financial Analysts Journal*, 58(4), 36–52.
- [5] Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), 703–708.
- [6] Jobson, J. D., & Korkie, B. M. (1981). Performance Hypothesis Testing with the Sharpe and Treynor Measures. *Journal of Finance*, 36(4), 889–908.
- [7] Memmel, C. (2003). Performance Hypothesis Testing with the Sharpe Ratio. *Finance Letters*, 1, 21–23.
- [8] Bailey, D. H., & López de Prado, M. (2014). The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting and Non-Normality. *Journal of Portfolio Management*, 40(5), 94–107.
- [9] Bailey, D. H., Borwein, J. M., López de Prado, M., & Zhu, Q. J. (2017). The Probability of Backtest Overfitting. *Journal of Computational Finance*, 20(4), 39–69.
- [10] Booth, D. G., & Fama, E. F. (1992). Diversification Returns and Asset Contributions. *Financial Analysts Journal*, 48(3), 26–32.
- [11] Willenbrock, S. (2011). Diversification Return, Portfolio Rebalancing, and the Commodity Return Puzzle. *Financial Analysts Journal*, 67(4), 42–49.
- [12] Harvey, C. R., & Liu, Y. (2015). Backtesting. *Journal of Portfolio Management*, 42(1), 13–28.
- [13] McLean, R. D., & Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability? *Journal of Finance*, 71(1), 5–32.
- [14] 深圳证券信息有限公司、国证指数官网. 国证价值 100、国证成长 100 指数编制方案与历史行情数据. <https://hq.cnindex.com.cn>.