

主动基金投资·配套学习资料

《主动基金投资》学习导引

从中国市场典型事实到可检验的研究与投资框架

刘洋溢、魏建榕、连祥斌、石川著 | 2026 年 8 月

先建立全书地图，再进入细节

适合具备基础金融知识、希望系统理解主动基金的读者

先读导引

- 把握四个典型事实与四类研究问题
- 理解每类指标在测量什么，而非背诵名称
- 识别结论的市场、数据与方法边界

再回到原书

- 用页脚的章、节编号定位详细论证
- 阅读公式、实证设定与代表性论文
- 将自测题作为是否继续深入的检查点

声明 本材料是阅读导航，不替代原书，也不构成基金或股票的投资建议。

被动投资兴起后，主动基金仍值得研究

问题不在于主动与被动二选一，而在于能否识别真正的基金能力

被动投资的优势

- 费率较低、透明度较高
- 稳定获取市场或风格因子收益
- 减少为非能力部分支付高额费用



主动研究仍有空间

- 中国主动股票型基金历史上平均来看可获取正 Alpha
- 基金之间的能力差异很大
- 优选的关键是把能力与风格、运气和规模分开

四个典型事实构成全书起点

事实一 | 平均正 Alpha

中国主动股票型基金在历史样本中，平均获得正的风险调整后超额收益。

事实二 | 基金赚钱、基民不赚钱

申购赎回时点和规模扩张，使投资者实际收益可能显著低于基金收益。

事实三 | 投资者追逐业绩

历史业绩是基金资金流最重要的解释变量之一。

事实四 | 简单抄作业不可靠

机械持有基金重仓股，并不自动带来优于市场的风险收益表现。

这些是样本中的经验事实，不是对未来收益的保证。

能力识别串联资金流、持仓与新方法

四类研究问题，加上一个中国市场实践落脚点



实践并非第五个等权研究主题

基金画像、分类、优选、配置与持仓测算，是把前述研究结论组合成投资流程的落脚点。

贯穿全书的判断标准：经济含义、样本外证据和中国市场适配性。

业绩预测的核心是识别基金能力

预测对象不是“过去谁涨得多”，而是“未来谁能创造经风险调整的收益”

**剥离风格、费用、运气和规模影响后，
历史业绩才可能反映能力。**

历史业绩 + 基金特征 + 持仓与交易 → 能力代理变量 → 样本外检验

基金超额收益来自能力、因子与噪声

先分解收益，再讨论基金经理是否创造了价值

$$\text{基金超额收益} = \text{Alpha} + \text{因子相关收益} + \text{随机扰动}$$

Alpha

选定因子模型或基准无法解释的收益。

因子相关收益

系统性因子暴露乘以对应因子收益。

随机扰动

模型未解释的当期噪声，不代表稳定能力。

使用边界 Alpha 取决于模型；原始收益高不等于能力强。

基准选择会改变基金 Alpha 的判断

基准决定哪些收益被视为风险补偿，哪些被归入能力

常见基准选择

- 基金合同披露的业绩比较基准
- 市场或多因子资产定价模型
- 按持仓、风格或文本分类得到的同类基金

判断前先问

- 基准是否覆盖基金真实投资范围？
- 基准是事前可得，还是事后选择？
- 结论对替代基准是否稳健？

结论：基金能力评价不是一个脱离基准的绝对排名问题。

主动偏离可能反映私有信息

理性的基金经理只应在拥有信息优势时偏离基准

获得私有信息 → 形成不同判断 → 持仓偏离基准 → 潜在 Alpha

理论含义

主动性可以成为私有信息和基金能力的代理变量。

不能反推

偏离基准也可能来自风格漂移、赌博动机或错误判断。

使用边界 主动性是有理论支撑的预测信号，不是能力的直接观测。

业绩预测指标分为简单特征与主动性指标

简单基金特征

- 基础特征：规模、费率、年龄、资金流
- 经理特征：团队结构等
- 业绩特征：动量、调整后业绩

数据较易获得，经济机制可能不唯一。

基金主动性指标

- 收益率相对基准的偏离
 - 持仓相对基准的偏离
 - 交易、风险和管理能力代理变量
- 更贴近基金特殊性，但数据与构造要求更高。

使用边界 分类帮助理解信息来源；最终仍需用未来业绩检验预测能力。

简单特征易用，却难稳定刻画能力

每个特征都可能同时承载能力、约束和选择效应

规模

期末基金净资产，单位通常为元；既反映历史资金流，也受策略容量约束。

费率

管理、托管及销售等费用以资产比例计；直接减少投资者净收益。

年龄与团队

由成立日和经理任职信息构造；可能反映经验，也可能混入幸存者选择。

中国市场结论

本书实证显示，绝大多数简单特征的预测能力并不显著；不能仅凭特征标签机械打分。

动量和资金流兼具信息与内生性

基金动量

含义：过去一段窗口内的基金业绩或 Alpha。

频率：常以月度收益构造，预测下月或下一季度。

输入：基金净值收益、无风险收益和因子收益。

基金资金流

含义：剔除净值涨跌后，基金净资产的净申购赎回变化。

频率：常用月度或季度数据（中国市场目前只有季度数据，美国市场提供月度数据）。

输入：期初、期末净资产及同期基金收益率。

使用边界 资金流既可能反映投资者识别能力，也会扩大基金规模并影响其未来业绩。

收益率主动性衡量基金偏离基准的程度

R^2

基金超额收益回归于因子模型的拟合优度，范围为 0 至 1；本书以过去 24 个月月度收益滚动估计。数值越低，基准解释力越弱。

跟踪误差

基金收益减基准收益的波动率；单位为每期或年化收益波动。需要同期基金与基准收益。

使用边界 R^2 较低或跟踪误差较高，只说明基金偏离基准更大；这种偏离可能来自能力，也可能来自额外风险。

持仓主动性直接比较基金与基准

行业集中度 (ICI)

比较基金与基准在各行业上的权重差异，再把所有行业的偏离程度汇总为一个指标。同一报告期计算，数值越高表示基金越集中地偏离基准行业配置。

主动份额

基金与基准各证券权重绝对差的一半，范围通常为 0 至 100%。需要同一时点的完整持仓及明确基准；数值越高表示持仓差异越大。

中国市场的数据约束

季报通常仅披露前十大重仓股，完整持仓主要来自半年报和年报；高换手率会进一步放大测量误差。

管理能力需结合净值、持仓与资金流

现金管理能力

考查资金流冲击下的交易与现金安排；需要收益、规模和持仓数据。

收益缺口

实际基金收益减以上期披露持仓计算的同期持仓收益；单位为同期收益率。

隐形交易

用更高频收益与持仓信息修正收益缺口，以捕捉两个披露日之间未被观察到的交易。

使用边界 收益缺口还会受到费用、交易成本和披露时滞影响，不能全部解释为交易能力。

规模报酬递减会侵蚀未来 Alpha

策略容量把基金搜寻机会的能力与可管理规模连接起来



交易冲击、流动性约束与机会稀释，压低未来业绩

使用边界 主流证据支持个体基金存在规模报酬递减，但其大小、行业层面效应和识别方法仍有争议。

中国市场不能照搬海外指标定义

积极证据

- 基金动量等少数简单特征表现较突出
- 行业集中度、收益缺口和时变能力有正向预测证据
- 中国情境下修正的隐形交易与现金管理指标值得关注

不能忽视的限制

- 持仓披露频率低且不完整，基金换手率较高
- 不同因子模型会改变 Alpha 和指标表现
- 发表后套利、竞争和资金拥挤会削弱样本外收益

最终判断应基于滚动、样本外和多基准检验，而非单次历史回测。

自测：业绩等于能力吗？

先独立作答

- ① 一只基金原始收益很高，为何仍不能断定基金经理能力强？
- ② 主动份额很高，为什么既可能是好信号，也可能是坏信号？
- ③ 为什么历史业绩优异且规模迅速扩张的基金需要格外谨慎？

核对思路

- ① 收益可能来自因子暴露、运气或不恰当的基准。
- ② 偏离既可能源于私有信息，也可能源于赌博或错误判断。
- ③ 资金流会扩大规模；超过策略容量后，未来 Alpha 可能下降。

资金流连接投资者决策与基金业绩

投资者行为并非外围话题，而是业绩预测链条的一部分



短期交易压力与长期规模报酬递减共同影响业绩

使用边界 基金净资金流是全部投资者决策的汇总代理变量，无法揭示个体投资者之间的异质性。

追逐历史赢家未必带来持续业绩

业绩追逐

投资者倾向申购过去表现较好的基金，并赎回过去表现较差的基金。实证上常用资金流对滞后业绩的敏感性刻画。



业绩持续性没有那么强

历史业绩混合了能力、因子暴露、运气与规模变化；过去赢家不一定继续领先。

这组张力是理解“基金赚钱、基民不赚钱”的关键起点。

业绩追逐既可能是理性学习，也可能是追逐简单业绩信号

理性贝叶斯学习

- 历史业绩是基金能力的含噪声信号
- 投资者结合先验和新信息更新能力判断
- 资金流使基金扩张，直至净 Alpha 被规模效应压低

简单业绩信号追逐

- 投资者依赖评级、排名等醒目信号
- 对风险调整、基准选择和能力来源关注不足
- 基金经理可能迎合信号并扭曲风险承担

不同专业度的投资者共存于市场

“投资者是否理性” 不应被压缩成一个全市场平均答案

高专业度

行为更加专业，如采用更复杂、更专业的模型来评估基金能力。

有限专业度

关注 CAPM-Alpha 等简单风险调整指标，但较少使用多因子信息。

低专业度

主要依据评级、短期排名和显著信息选基。

使用边界 基金层面的净资金流把不同投资者混合在一起；账户级别与流入、流出分项数据有助于识别异质性。

行为偏差也会影响基金投资者

彩票偏好

偏好小概率极高回报的基金或持有彩票型股票较多的基金，可能接受更低的期望收益。

前景理论

投资者以参考点衡量盈亏，并对损失和收益表现出不对称反应。

有限关注

更容易处理显眼、简单和近期的信息，忽视复杂但可能更有价值的能力指标。

因子收益追逐

把风格或因子带来的历史收益误认为基金经理能力，推动资金在风格间迁移。

互联网平台可能放大热门基金追逐

可能改善决策

- 降低基金信息获取与交易成本
- 扩大产品可得性和比较范围
- 为个性化投资者教育提供数据基础

也可能放大偏差

- 排名、推荐位和关注度强化显著性
- 热门基金获得更多流量与资金流
- 基金经理可能迎合平台偏好与短期业绩

平台效应取决于信息如何展示，而不只是信息数量是否增加。

自测：追涨必然意味着不理性吗？

先独立作答

- ① 为什么业绩追逐既可能符合贝叶斯学习，也可能反映专业度不足？
- ② 资金流为何会改变被追逐基金的未来业绩？
- ③ 互联网平台提供更多信息，为何仍可能降低决策质量？

核对思路

- ① 关键在于投资者如何处理业绩信号，而不是是否使用历史业绩。
- ② 流入扩大规模并触发交易，可能加剧规模报酬递减。
- ③ 排名和推荐机制可能放大显著性与有限关注。

机械复制重仓股忽略基金能力差异

持仓披露有信息，但“所有基金、所有重仓股等同处理”会稀释信息质量

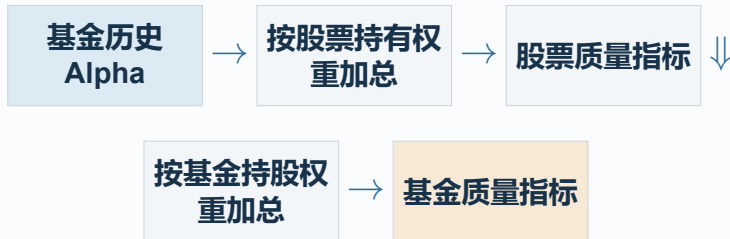
简单抄作业的问题

- 大量基金并不具有稳定选股能力
- 持仓披露存在时滞且季报并不完整
- 重仓可能来自风格暴露、流动性或监管约束

信息何时更有用

- 先识别有能力的基金
- 将基金持仓与业绩联合使用
- 把基金信号作为自身策略的补充验证

基金质量与股票质量双向提炼信息



经济含义

被能力更强的基金以更高权重持有的股票，可能包含更高质量的机构信息；再把股票质量聚合回基金，可补充单一基金 Alpha。

基金隐含收益只是增量信号

较稳妥的使用顺序

- ① 形成自身的股票或基金研究信号
- ② 估计持有者的基金能力与信息质量
- ③ 检查基金持仓是否提供一致或冲突证据
- ④ 对冲突信号进一步追溯基本面和时点

不能忽视的摩擦

- 披露日与实际建仓时点不同
- 交易成本、拥挤与价格冲击可能吞噬收益
- 基金可能在披露前进行窗饰或调整

无信息资金流也会影响价格

资产定价效应来自交易冲击，而不一定来自资金流的信息含量



短期含义

价格压力可能制造业绩持续、聪明钱效应或动量的表象。

长期检验

需区分价格冲击、信息扩散与基本面变化，不能只看同期相关。

共同资金流冲击形成系统性定价因子

共同资金流冲击

先从基金资金流中剔除历史业绩、滞后资金流等可预测成分，再将基金按规模分为五组，从规模加权平均冲击中提取第一主成分。

资金流 Beta

股票收益对共同资金流因子的敏感度。本书介绍的美国证据表明，这种暴露与预期股票收益相关，因而成为被定价的股票特征。

使用边界 中国市场的资金流—业绩凸性、投资者结构和交易约束不同，定价方向与强度仍需本土检验。

基金网络提供持仓之外的信息

重仓股网络

共同重仓连接基金，网络内可能发生信息扩散。

竞争网络

风格越相近，竞争越强；过度竞争可能阻碍交流。

社交网络

地理、校友或社交联系可能传递私有信息，但识别困难。

使用边界 网络位置与业绩相关，并不自动证明信息传递；共同风格、地理选择和内生连接都可能构成替代解释。

自测：基金信息为何不能机械使用？

先独立作答

- ① 简单复制全部基金重仓股为什么会稀释信息？
- ② 无信息的资金流为何仍可产生资产定价效应？
- ③ 基金网络与收益相关时，主要识别风险是什么？

核对思路

- ① 基金能力异质，持仓又有时滞和约束。
- ② 申赎迫使基金交易既有持仓，形成价格压力。
- ③ 网络可能同时代理共同风格、选择效应或真实信息传递。

基金特征与未来业绩可能非线性

机器学习的价值，不只是“放入更多变量”，而是识别交互与非线性

线性模型通常假设

每个指标以相对稳定、可加总的边际效应影响基金未来业绩。

例子

规模的影响可能取决于持仓流动性，动量的作用也可能随资金流和市场状态变化。

非线性模型可以刻画

阈值、变量交互、状态依赖，以及不同基金群体中的异质关系。

代价

模型更复杂、解释更困难，也更容易在有限样本中拟合噪声。

复杂模型中，简单特征仍很重要

基金动量

历史 Alpha 或价值创造能力常有较高预测重要性。

基金资金流

有助于短期预测，也包含规模扩张和交易压力。

主动性信息

R^2 等简单主动性指标可补充历史业绩。

正确解读

“简单特征重要”不代表建立一条固定打分规则即可持续获利；模型设定、预测对象、再平衡频率和样本期均会改变结果。

持仓异象与深度学习取决于预测目标

持仓异象特征

将股票层面的价值、盈利、投资、动量等特征按基金持股权重聚合，形成基金层面的异象投资指标；需要同一时点持仓与股票特征。

深度学习

同时使用基金、持股和基金家族特征，通过神经网络捕捉非线性交互；需要严格的训练、验证和样本外测试。

使用边界 总超额收益、Alpha 与 FFC 四因子调整后的当期异常收益口径不同；比较模型前必须统一预测目标。

文本分析将策略表述转为研究变量

核心文本来源是募集说明书和基金定期报告

投资策略文本

- 测度策略文本的独特性与创新程度
- 按文本相似性重新进行基金分类
- 区分量化基金与主观基金

定期报告文本

- 提取基金经理的经济与市场判断
- 分析风险披露及其投资行为后果
- 检验文本信息能否预测基金或资产表现

前沿方法不能把预测等同于机制

样本外衰减

特征发表、策略采用和竞争会压缩后续收益。

数据泄漏

文本、持仓和特征在预测时点必须已经可得。

解释不唯一

变量重要不等于因果；
文本也可能只是披露风格。

最低检验标准

时间顺序明确 | 滚动或扩展窗口训练 | 独立验证集 | 交易成本与换手率 | 替代模型和特征组 | 中国市场再检验

自测：机器学习何时真正有价值？

先独立作答

- ① 非线性模型相较单指标排序，主要增加了什么能力？
- ② 为什么持仓特征能预测总收益，却未必能预测 Alpha？
- ③ 文本指标显著时，为什么仍不能直接推断基金经理具有信息优势？

核对思路

- ① 识别阈值、交互和状态依赖关系。
- ② 持仓特征可能主要刻画系统性风险暴露。
- ③ 文本还可能反映披露习惯、营销或内生选择。

实践需要可持续的组合流程

研究结论只有嵌入比较、配置、持仓估计和持续监测，才可能转化为决策质量



流程原则

五步对应第 12 章五节；风格、规模、风险和数据质量监测贯穿各环节，并形成反馈闭环。

基金画像与分类决定评价是否公平

基金画像

综合收益率、波动、回撤、因子暴露、行业与个股持仓、换手率和基金经理信息，描述基金“如何赚钱、承担什么风险”。

基金分类

将画像转为特征向量，以距离和相似性网络构建可比组，可用 Leiden 社区发现形成多级标签；分类应允许风格迁移。

使用边界

若分类过粗，风格差异会被误认为能力；若分类过细，样本过少又会使排名不稳定。

基金优选与因子配置应协同

横截面优选

在同类基金中组合历史业绩、主动性、规模约束和风险控制指标；用信息系数、分组收益和样本外表现检验稳定性。

因子配置

先判断不同风格或因子的相对环境，再在相应类别中优选基金；配置与选基应协同，而非分别追逐各自历史最优。

使用边界 本书中的回测结果来自特定样本、基准、费率和调仓规则，不能直接视为未来可实现收益。

持仓测算弥补披露空档，但不创造信息

目标是在披露日之间估计基金的仓位和因子暴露

主要输入

基金高频净值收益、已披露持仓、行业与资产比例、股票或因子收益，以及半年报和年报的完整持仓锚点。

本书中的方法

二次规划通过约束优化匹配基金收益；卡尔曼滤波允许持仓或因子暴露随时间逐步变化。

使用边界 估计误差会在高换手、风格突变和收益共线时上升；测算结果应与后续披露交叉验证。

四类研究共同支持可检验的基金判断

能力识别



行为与资金流



持仓与定价



新数据与方法

研究问题

指标测量什么？为何有效？有无中国样本外证据？

投资问题

改善哪一步决策？是否考虑规模、成本、风险与时点？

纪律要求

区分事实、机制、预测与因果；保留不确定性。

不同读者重点不同，但误区相同

优先阅读章节

- **个人投资者：**优先第 1、5、7、9、12 章
- **从业者：**优先第 1、3、7、9、12 章
- **研究者：**全书全部内容

五个常见误区

- 把短期收益当成稳定能力
- 把风格暴露当成 Alpha
- 忽视规模报酬递减和费用
- 机械复制重仓股或模型排名
- 把回测相关性写成经济机制

带着术语与边界阅读原书

术语速查

- **Alpha**: 选定模型或基准无法解释的风险调整后收益
- **Beta**: 资产收益对系统性因子收益的敏感度
- **基金资金流**: 剔除净值涨跌后的规模净变化
- **主动性**: 基金收益或持仓相对基准的偏离程度

代表性文献

- Berk and Green (2004); Chen et al. (2004)
- Kacperczyk et al. (2005, 2008, 2014)
- Cohen et al. (2005); Lou (2012)
- DeMiguel et al. (2023); Kaniel et al. (2023)
- Kostovetsky and Warner (2020)

本导引为《主动基金投资》配套学习资料，仅供学习交流。历史研究和回测不代表未来表现；投资决策应结合自身目标、风险承受能力和最新公开信息独立作出。