

基于树模型的有效前沿扩展

——学海拾珠系列之二百二十八

报告日期: 2025-03-20

分析师: 骆昱杉

执业证书号: S0010522110001

邮箱: luoyushan@hazq.com

分析师: 严佳炜

执业证书号: S0010520070001

邮箱: yanjw@hazq.com

相关报告

- 《使用深度强化学习解决高维多期环境下的组合配置 —— “学海拾珠”系列之二百二十七》
- 《风险规避型强化学习模型在投资组合优化中的应用 —— “学海拾珠”系列之二百二十六》
- 《贝塔异象的波动性之谜 —— “学海拾珠”系列之二百二十五》
- 《ETF 的资产配置与再平衡: 样本协方差对比 EWMA 与 GARCH 模型 —— 学海拾珠系列之二百二十四》
- 《市场对投资者情绪的反应 —— 学海拾珠系列之二百二十三》
- 《基于语境的财务信息解读 —— 学海拾珠系列之二百二十二》
- 《跟踪误差的构成成分、中期交易与基金业绩 —— 学海拾珠系列之二百二十一》

主要观点:

本篇是学海拾珠系列第二百二十八篇, 文章介绍了一种新的基于树的模型——Panel Tree (P-Tree), 用于分析个体资产收益的面板数据, 并在均值-方差有效框架下构建测试资产和潜在因子, 以扩展有效前沿。

● P-Tree 模型的构建与美国股票市场实证效果

P-Tree 模型通过高维特征对个体资产收益进行聚类, 生成测试资产和潜在因子, 显著推进了有效前沿。模型具有解释性强、计算效率高和稳健性好的特点, 为资产定价和投资管理提供了新的工具和方法。美国市场实证结果显示 P-Tree 模型能显著提升资产定价和投资管理的效果。

● P-Trees 模型拓展

Boosting P-Trees 通过逐步添加新的因子使投资组合的效率不断提高, 显著推进了有效前沿。同时 Boosting P-Trees 在投资策略中表现出色, 具有高夏普比率和显著的阿尔法值, 即使在样本外测试中也表现出色。随机 P-Trees 进一步扩展了 P-Tree 的模型复杂性, 提供了更稳健的 SDF 估计和特征重要性评估。

● P-Trees 模型评估

实证结果发现某些特征 (如 SUE、DOLVOL 和 BM_IA) 在基准 P-Tree 上进一步分裂时, 能够显著提高投资组合的表现在不同宏观经济状态下都能有效调整, 选择不同的特征进行均值-方差优化。此外, P-Tree 生成的测试资产在极端宏观经济状态下对现有因子模型的挑战更大。

● 文献来源

核心内容摘自 Lin William Cong、Guanhao Feng、Jingyu He 和 Xin He 于 2025 年 02 月 04 日在 Journal of Financial Economics 上的文章《Growing the Efficient Frontier on Panel Trees》。

● 风险提示

文献结论基于历史数据与海外文献进行总结; 不构成任何投资建议。

正文目录

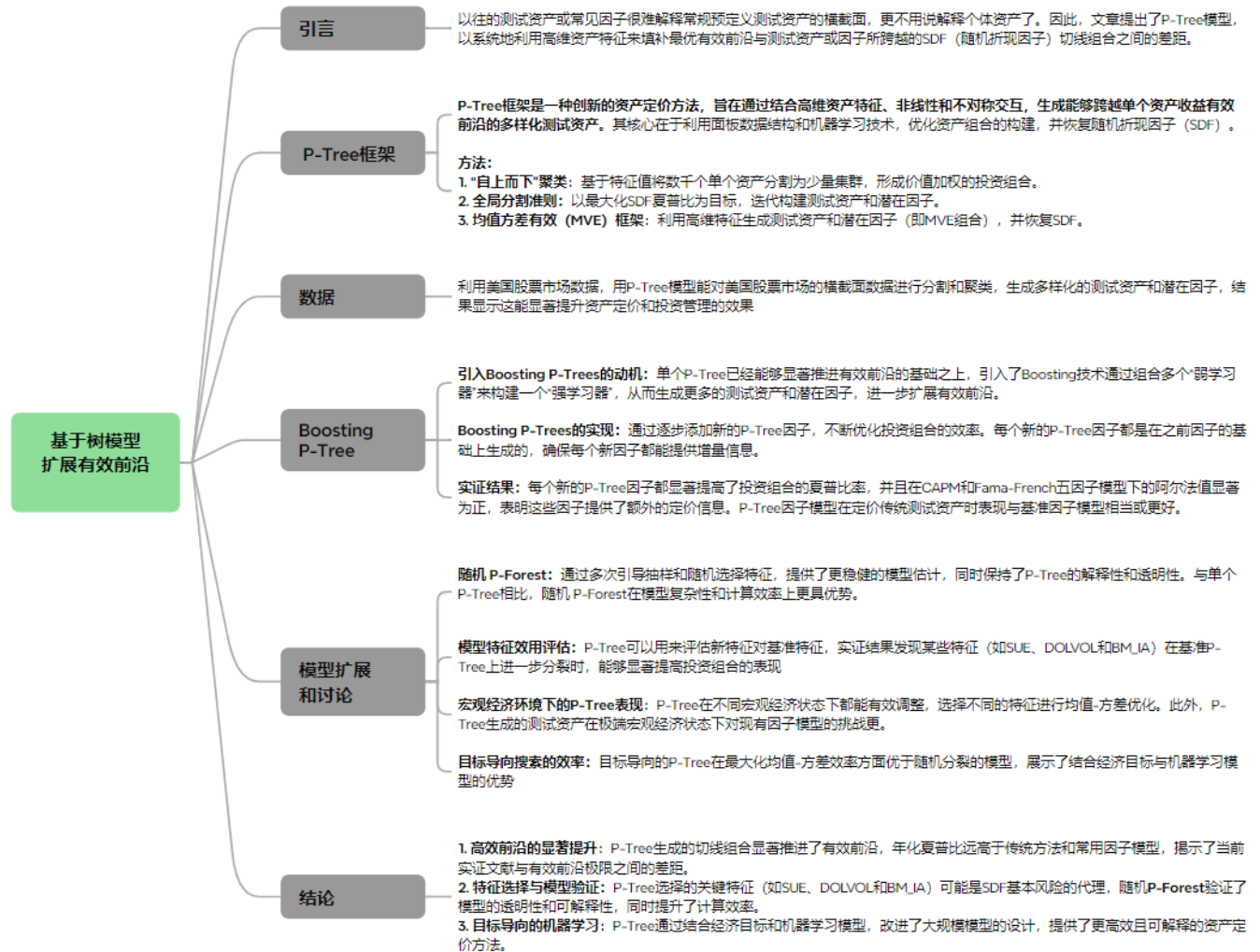
1 引言.....	4
2 用于资产定价的面板树.....	7
2.1 CART 和 P-TREE 创新.....	7
2.2 P-TREE 的生成.....	8
2.3 BOOSTING P-TREE.....	10
2.4 模拟试验.....	11
3 美国股票数据的横截面分割.....	12
3.1 美国股票数据.....	12
3.2 可视化单个 P-TREE.....	13
3.3 P-TREE 叶节点投资组合.....	15
4 用于多因子和测试资产的 BOOSTING P-TREE.....	18
4.1 拓展有效前沿.....	18
4.2 利用 BOOSTING P-TREE 生成测试资产.....	21
4.3 P-TREE 因子和资产定价表现.....	23
4.4 P-TREE 因子投资.....	24
5 模型扩展和讨论.....	26
5.1 随机 P-FOREST 与 P-TREE 的可解释性.....	26
5.2 随机 P-FOREST 和复杂性.....	27
5.3 用 P-TREE 评估股票特征.....	29
5.4 宏观经济环境下的 P-TREE.....	31
6 结论.....	32
风险提示:	32

图表目录

图表 1 文章框架	4
图表 2 决策树示例	7
图表 3 第一次拆分示例	9
图表 4 第二次拆分示例	10
图表 5 股票特征	12
图表 6 宏观经济变量	13
图表 7 1981 年到 2020 年的面板树	14
图表 8 分区图可视化非线性交互	15
图表 9 叶节点投资组合的评估	16
图表 10 多样化的 P-TREE 测试资产	17
图表 11 用 P-TREE 测试资产	18
图表 12 测试 BOOSTING P-TREE 的生成	20
图表 13 比较测试资产	22
图表 14 资产定价表现: 横截面 R^2	24
图表 15 通过 BOOSTED P-TREES. 进行因子投资	25
图表 16 通过选择概率确定特征重要性	27
图表 17 随机 P-FOREST SDF 的 OOS 性能	28
图表 18 ME 基线和 ME-BM 基线 P-TREE 的演示	29
图表 19 使用 P-TREE 评估股票特征	30
图表 20 政策变动下的 P-TREE 性能	31

1 引言

图表 1 文章框架



资料来源：华安证券研究所整理

估算均值-方差有效（MVE）边界对资产定价和投资管理至关重要。然而，利用数千个单个资产收益的非平衡面板估计切线组合（马科维茨，1952）是不切实际的。实证研究通常会考虑一组“多样化”的测试资产（如MEBM 25投资组合）来估计和评估因子模型，希望这些测试资产或少数几个共同因子能跨越与单个资产相同的有效边界。然而，目前流行的因子模型很难解释传统的预先指定测试资产的横截面（例如，Kozak, 2018；Lopez-Lira 和 Roussanov, 2020），更不用说这些测试资产的临时性阻碍了模型估计和评估的有效性（Lewellen 等，2010；Ang 等，2020）。例如，由于高维排序面临挑战（Cochrane, 2011），忽略了非线性和非对称交互作用（并非统一适用于所有资产），导致基于特征的测试资产往往局限于单变量和双变量排序的投资组合，即使是依赖排序也是如此（Daniel 等，1997）。通过分别评估不同的测试资产进行稳健性检查（如Fama 和 French, 1996）、选择测试资产（Daniel 等，2020；Giglio, 2023；Bryzgalova 等，2023）或合并测试资产（如Feng 等，2020）无法完全解决这些问题。

因子模型和测试资产的评估涉及均值-方差分散化，因此，使用这些测试资产或因子构建的切线组合的表现至关重要。Cochrane（2011）指出，预期收益、方差和协方差是资产特征（如规模、价值）的稳定函数，Kelly 等（2019）进一步指出，收益-因子协方差与资产特征相关。因此，填补最优有效前沿与测试资产或因子所跨越的 SDF（随机折现因子）切线组合之间差距的关键，在于系统地利用高维资产特征，这些特征包含了资产收益动态联合分布的丰富信息（例如，Kozak 和 Nagel, 2022）。

Panel Tree（P-Tree）框架：

为了实现这一目标，我们提出了一种新方法（P-Tree），将非线性和不对称交互与高维特征相结合，以创建能够跨越单个资产收益有效前沿的测试资产。Panel Tree（P-Tree）框架受到现代人工智能的启发，构成了一类多功能模型，通过对面板观测进行聚类来实现特定的经济目标。创建测试资产需要将单个资产聚类为组合以形成投资组合。在面向目标的全局分割准则下，P-Tree 对单个资产收益进行聚类并创建基础投资组合。它在均值方差有效（MVE）框架下利用高维特征，联合生成测试资产和一个潜在因子（即测试资产的 MVE 组合），并恢复随机折现因子（SDF）。因此，P-Tree 将回归树模型的范围扩展到纯预测任务之外，特别是生成了能够达到最终有效前沿的多样化测试资产。

具体而言，P-Tree 采用“自上而下”的方法（通常树的根在顶部），基于特征值将数千个单个资产的横截面分割并分组为少量集群，以形成（价值加权的）投资组合。在资产定价的指导下，我们通过迭代构建测试资产和 SDF 的潜在因子来生长 P-Tree，遵循最大化 SDF 夏普比的全局分割准则。构建的 SDF 的高夏普比反映了由生成的叶基础投资组合跨越的高 MVE 前沿。生成的叶基础投资组合和潜在因子为研究人员提供了一组多样化的测试资产和 SDF 模型。此外，P-Tree 直观且透明，允许经济解释，并识别出一组稀疏的有用特征，这些特征相互作用以共同生成测试资产和潜在因子。

从根本上说，P-Tree 是一种基于资产定价目标的贪婪搜索算法，用于在大规模建模空间内基于相似特征值进行最优聚类。在实证资产定价中，基础投资组合通常通过基于国家或行业分类的资产分组，或通过基于收益相关性或特征的无监督聚类（例如证券排序）来形成。相比之下，P-Tree 生成基于特征管理的叶基础投资组合，以最大化其 MVE 组合的夏普比，既不需要监督也不需要无监督。目标可以灵活指定，使得 P-Tree 框架适用于广泛的面板数据分析。

方法论创新：

首先，现成的机器学习方法，包括著名的分类与回归树（CART，Breiman 等，1984），通常假设观测值是独立同分布（i.i.d.）的，并非为分析面板数据而设计。尽管在特征-收益数据上直接拟合 CART 或树集成模型（如提升树或随机森林）显示出积极的预测性能（例如，Gu 等，2020），但它们忽略了面板数据结构。相比之下，P-Tree 完全拥抱面板数据结构，并为多期观测引入时间不变的树结构。时间不变的 P-Tree 在面板数据上构建机器学习模型时，允许经济可解释性。例如，类似于证券排序，时间不变的 P-Tree 允许所有时期使用同一组基于特征管理的叶基础投资组合。

其次，包括 CART 在内的标准基于树的机器学习模型仅关注预测。此外，基于 CART 的模型通过递归生长，优化预测的二次损失。这些模型在每个节点局部优化，

而不考虑兄弟节点，主要是为了提高计算效率。然而，这种“短视”策略通常会致过拟合，因为随着树的生长，每个节点上的观测值减少。相比之下，P-Tree 将应用范围扩展到预测之外，涵盖测试资产和潜在因子的生成。P-Tree 的迭代生长旨在利用整个横截面的数据，以防止使用局部分割准则生长的传统树所面临的过拟合问题。P-Tree 结合了经济原理和非线性机器学习算法，同时确保图形树图的可解释性，为分割横截面和生长有效前沿提供了统一的方法。

第三，P-Tree 框架可以集成机器学习的 boosting 或 bagging 策略，使多个 P-Tree 形成多因子模型。一方面，提升 P-Tree (Boosted P-Tree) 基于先前的 P-Tree 生长额外的树，为多因子模型提供了一个统一框架，使得额外的因子和测试资产必须提供增量贡献。集成方法使 P-Tree 成为一种多功能工具，可以通过在 MVE 框架下利用未跨越的有效前沿来增强任何预定因子模型的性能。当提升以生成多个因子时，P-Tree 提供了主成分分析 (PCA) 和深度神经网络的替代方案，具有更高的可解释性和稀疏性，同时捕捉 (不对称) 交互作用。另一方面，随机 P 森林 (Random P-Forest) 在随机引导样本上生成额外的 P-Tree，提供了一个统一的框架来创建多组不相关的因子和测试资产，可用于评估特征重要性。大规模随机 P 森林也与最近关于大规模和过参数化模型的文献相关 (例如，Didisheim 等，2024)。

实证研究结果:

首先，我们研究了 1981 年至 2020 年的美国月度股票收益，使用了 61 个公司特征。P-Tree 切线组合被构建为交易因子，显著推进了有效前沿，其年化夏普从单个 P-Tree (10 个投资组合) 的 6.37 到 20 个增强 P-Tree (200 个投资组合) 的 15.63 不等。这些数值显著高于传统基础投资组合 (单变量或双变量排序投资组合) 或常用因子模型构建的数值。这些发现强有力地证明了当前实证文献与有效前沿潜在极限之间的显著差距。此外，在统一的 MVE 框架下生成的 P-Tree 生成了多因子模型，能够有效定价横截面收益。同时，P-Tree 因子在样本外分析中表现出年化夏普比超过 3，并在过去预测未来和未来预测过去的测试中显示出显著的正阿尔法。

其次，这些多样化的 P-Tree 测试资产跨越了有效前沿，并对替代因子模型提出了重大挑战，突显了测试资产的重要性。我们识别了许多在经济和统计上未被解释的测试资产，表现为大多数月度阿尔法超过 1%，并且第一个 P-Tree 相对于 Fama-French 五因子模型 (Fama 和 French, 2015) 的 GRS 检验统计量高达 141.27 (加权定价误差)。这些不对称、交互式和非线性的特征管理叶节点投资组合难以用其他著名的可观测因子模型 (如 Hou 等，2021) 或基于机器学习的潜在因子模型 (如 Kelly 等，2019; Lettau 和 Pelger, 2020) 解释。鉴于单变量或双变量排序投资组合在模型测试中的 MVE 跨度不足和低门槛，我们建议在未来的模型评估中纳入这些多特征、系统排序的 P-Tree 测试资产。

最后，我们设计了一个带有 bagging 的随机 P 森林 (Random P-Forest) 来评估特征重要性，考虑变量选择风险，并验证透明且可解释的 P-Tree 模型。我们确认，P-Tree 选择的同一小组特征 (如 SUE、DOLVOL 和 BM_IA) 可能是 SDF 中真实基本风险的代理，这些风险在具有临时选择因子的线性因子模型中被忽视。此外，随机 P-Forest SDF 作为一个大型正则化模型，在样本外表现出优异的夏普比，与随机分割 SDF 显示出相似的模式，但在计算上显著更高效。从某种意义上说，目标导向搜索改进了带有统计和经济正则化的穷举法大模型，有效地将经济目标与 ML 模型相结合。最后，P-Trees 显示了所有分割规则，有助于研究人员理解特征之间的相互作用。

文章的其余部分如下：第 2 节介绍 PTree 模型。第 3 节说明了单个 P-Tree 在分割横截面和生成测试资产方面的经验应用。第 4 节展示了 Boosting 和多因子 P-Trees 的实证结果。第 5 节讨论了随机 P-Forest 和 P-Tree 对宏观制度的稳健性。第 6 节为结论，互联网附录包括模拟和其他实证讨论。

2 用于资产定价的面板树

第 2.1 节介绍了 P-Tree 与标准树状模型相比的创新之处。第 2.2 节深入探讨了 P-Tree 的生长过程，第 2.3 节介绍了用于多因素模型的 Boosting P-Tree。第 2.4 节介绍了模拟结果。

2.1 CART 和 P-Tree 创新

CART 模型 (Breiman 等, 1984 年) 及其变体将预测空间划分为不同的、非重叠的叶节点区域，并为每个区域分配一个恒定的叶参数：假设数据构成 K 个预测变量，并且第 i 个观测值用 $\mathbf{z}_i = (z_{i,1}, \dots, z_{i,K})$ 表示。决策树的第 j 条拆分规则用 $\tilde{c}^{(j)} = (z_{i,k}, c_j)$ 表示，它通过检查第 k 个预测因子 $z_{i,k}$ 的值是否大于或小于值 c_j 来拆分数据。CART 只考虑二进制拆分，因为任何多向拆分树都可以表示为具有多个二进制拆分的树。选择最佳拆分规则是为了最小化预测误差。

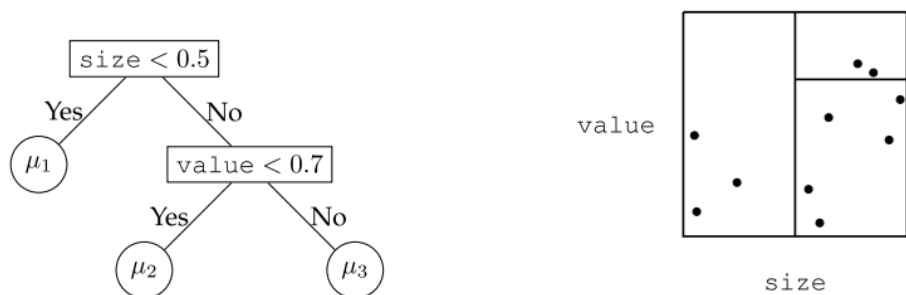
经过 j 次分割后，CART 将预测空间划分为 $j+1$ 个叶节点，用 $\{\mathcal{R}_n\}_{n=1}^{j+1}$ 表示，并为每个节点分配一个常数叶参数 μ_n 回归树 \mathcal{T} ，参数 $\Theta_j = \{\{\tilde{c}^{(j)}\}_{j=1}^j, \{\mu_n\}_{n=1}^{j+1}\}$ ，构成了一个高维阶跃函数：

$$\mathcal{T}(\mathbf{z}_i | \Theta_j) = \sum_{n=1}^{j+1} \mu_n \mathbb{I}\{\mathbf{z}_i \in \mathcal{R}_j\} \quad (1)$$

指标 $\mathbb{I}\{\mathbf{z}_i \in \mathcal{R}_j\}$ 对一个叶节点取 1，对其他叶节点取 0。阶跃函数中的叶参数是通过平均每个叶节点内的训练数据估算出来的。这种非参数方法使 CART 能够适应低信噪比环境和小样本量。树状模型通过定位相应的树叶并使用其参数作为训练后的预测值来预测新的观测结果。

图表 2 展示了传统的树状模型 (如 CART) 如何通过递归将特征空间划分为群组 (小盘股、大盘成长股和大盘价值股) 来预测具有公司特征的股票收益。底部节点通常称为叶节点，与恒定的叶参数相关联，这意味着同一节点内的资产共享相同的回报预测。

图表 2 决策树示例



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：左图决策树具有两个拆分、三个叶节点和三个叶参数。右图为值和大小预测器空间样本生成相应的分区图。

然而，由于存在多种限制，这些方法不适合经济或金融面板数据，也不适合生成测试资产或定价核。首先，它们将面板资产收益视为 i.i.d 独立同分布观察值。其次，它们侧重于使用局部拆分标准进行预测，如最小化特定拆分子节点的平方误差目标之和。因此，它们不包含面板结构或经济约束。随着树的生长，每个节点中的观测数据数量会减少，特异性噪声会导致过度拟合。虽然树集合有助于缓解过度拟合，但其可解释性不如单一决策树。

P-Tree 通过以下方式解决了这两个问题：(i) 采用时间不变的树形结构；(ii) 采用经济目标，考虑所有观察结果，而不仅仅是父节点中的观察结果，从而指导树的生长。这样，(i) 就能彻底提取面板数据以构建叶节点投资组合，(ii) 利用全局分割标准迭代构建 **P-Trees** 以防止过度拟合，从而将树的范围从纯粹的预测扩展到以目标为导向的聚类，以生成测试资产和因子。

2.2 P-Tree 的生成

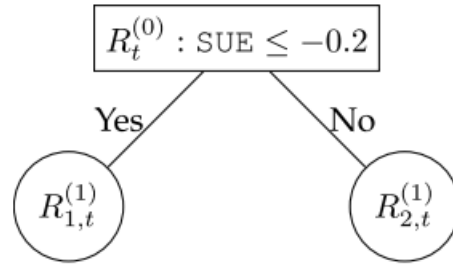
P-Tree 根据排序特征的值将单个资产的分组划分为不重叠的叶节点，并且在 T 个周期内保持时间不变的结构。叶节点投资组合是基于时间不变树结构中每个叶节点内股票构建的（价值加权）投资组合。**P-Tree** 从根节点开始依次分割，并在每次额外分割后生成一个新的叶节点基础投资组合。具体来说，在第 j 次分割后，它会生成 $j + 1$ 个叶节点投资组合，从而将数千个单个资产的维度减少到 $j + 1$ 个投资组合。这些基于特征管理的叶节点投资组合扩展了特征排序投资组合，以适应多个特征之间的不对称交互作用。

$\mathbf{R}_t^{(j)}$ 表示树的第 j 次分割后叶节点投资组合的所有超额收益向量，让 $f_t^{(j)}$ 表示 $\mathbf{R}_t^{(j)}$ 所跨因子。树的所有叶节点都有 $j+1$ 个基础组合，用 $\mathbf{R}_t^{(j)} = [R_{1,t}^{(j)}, \dots, R_{n,t}^{(j)}, \dots, R_{j+1,t}^{(j)}]$ 表示其 $R_{n,t}^{(j)}$ 表示第 n 叶节点在 T 期内的收益的长度为 T 向量。

为了找到第一次拆分，我们从所有资产的投资组合开始，记为 $\mathbf{R}_t^{(0)} = [\mathbf{R}_t^{(0)}]$ ，作为 **P-Tree** 的根节点。通过对 $\{\mathbf{R}_t^{(j)}, f_t^{(j)}\}$ 的迭代更新来扩展该树。首先，叶基组合 $\mathbf{R}_t^{(j)}$ 会随着树的生长而扩展。其次，利用扩展后的叶基组合，我们重新估计 **P-Tree** 因子 $f_t^{(j)}$ 。作为一种贪婪算法，与全面枚举可能的排序相比，**P-Tree** 的顺序增长确保了计算的可行性。

第一次分割：根节点包含市值加权市场因子对应的所有资产。 公司特征在每个时期内横截面统一归一化为 $[-1, 1]$ 。我们评估了特征 $z_{i,K}$ 的不同拆分阈值 c_m 方案，如五分位拆分 $c_m \in [-0.6, -0.2, 0.2, 0.6]$ 。考虑到每一个可选择的分叉 $\check{c}_{k,m} = (z_{\cdot,k}, c_m)$ 如图表 3。根据 $z_{\cdot,k}$ 是否小于或等于 c_m ，将根节点划分为左侧和右侧子节点。

图表 3 第一次拆分示例



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：为了确定最佳特征和切点值，本研究评估了一组候选拆分（即标准化意外季度收益（SUE） ≤ -0.2 ；40% 百分位数）以计算拆分标准值。

每个潜在子叶内的股票收益观测值构成一个叶节点投资组合，其收益向量分别为 $\mathbf{R}_{1,t}^{(1)}$ 和 $\mathbf{R}_{2,t}^{(1)}$ 。请注意，它们是收益向量，因此保持了资产收益的面板结构。P-Tree 因子是作为所有叶节点投资组合的 MVE 投资组合估算的。

$$\mathbf{f}_t^{(1)} = \mathbf{w}^{(1)'} \mathbf{R}_t^{(1)}, \quad \mathbf{w}^{(1)} \propto \hat{E}[\mathbf{R}_t^{(1)} \mathbf{R}_t^{(1)'}]^{-1} \hat{E}[\mathbf{R}_t^{(1)}] \quad (2)$$

其中， $\mathbf{R}_t^{(1)} = [\mathbf{R}_{1,t}^{(1)}, \mathbf{R}_{2,t}^{(1)}]$ 是第一次拆分后两个叶节点投资组合的收益矩阵、

$\hat{E}[\mathbf{R}_t^{(1)}]$ 和 $\hat{E}[\mathbf{R}_t^{(1)} \mathbf{R}_t^{(1)'}]$ 是叶节点投资组合的样本平均值和第二矩阵。生成潜在因子与 P-Tree 拆分标准一致，因为 P-Tree 因子代表所有叶基组合的 MVE 组合。

因此，在基线规范中，我们的拆分标准旨在最大化 MVE 投资组合的夏普比率：

$$\mathcal{L}(\tilde{\mathbf{c}}_{k,m}) = \sqrt{\hat{\mu}_F' \hat{\Sigma}_F^{-1} \hat{\mu}_F} \quad (4)$$

其中， $\mathbf{F} = \mathbf{f}_t^{(1)}$ 是生成的潜在因子，样本均值为 $\hat{\mu}_F$ ，协方差矩阵为 $\hat{\Sigma}_F$ 。该准则允许 P-Tree 同时构建和估计基础投资组合、潜在因子和有效投资组合权重，以实现构建有效前沿的总体目标。

每个候选拆分 $\tilde{\mathbf{c}}_{k,m}$ 都会产生不同的数据分区，从而产生独特的叶基组合、相应的 P-Tree 因子，并最终产生式子 (4) 中不同的拆分标准估值。因此，我们对所有候选拆分进行循环，并选择拆分标准（夏普比率）最大化的拆分作为初始拆分规则。

第二次拆分：第二次拆分可能发生在根的左侧或右侧子节点上。我们会评估两个叶子节点所有候选拆分的拆分标准，并选择最大化标准的拆分，如式子 (4) 所示。图表 4 描述了第二次拆分的候选树。无论是哪种情况，都会有一个叶节点分裂，成为内部节点，并产生两个新的叶节点。然后根据所有三个叶基组合构建 P-Tree 因子（因此是全局的）：

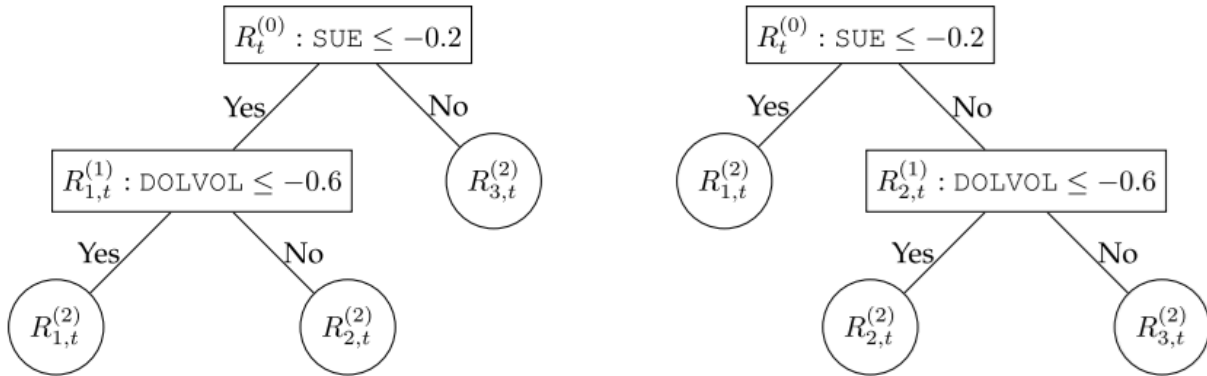
$$\text{Cross-sectional } R^2 = 1 - \frac{\sum_{i=1}^N (\bar{R}_i - \widehat{R}_i)^2}{\sum_{i=1}^N \bar{R}_i^2} \quad (5)$$

其中 $\hat{E}[\mathbf{R}_t^{(2)}]$ 和 $\hat{E}[\mathbf{R}_t^{(2)} \mathbf{R}_t^{(2)'}]$ 是第二次分裂后叶子基础组合 $\mathbf{R}_t^{(2)} = [\mathbf{R}_{1,t}^{(2)}, \mathbf{R}_{2,t}^{(2)},$

$\mathbf{R}_{3,t}^{(2)}]$ 的样本平均数和第二矩阵。如图 3 所示，三个基础组合的构建取决于候选者

分裂的节点。将更新后的 P-Tree 因子插入 (4) 中以最大化夏普比率，其中 $F = f_t^{(2)}$ 是更新后的 P-Tree 因子。

图表 4 第二次拆分示例



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：第二次拆分有两个可能的候选项，每个候选项都拆分一个原始叶节点。第二次拆分后，将生成三个叶节点投资组合以更新潜在因子。

值得注意的是，我们提出的目标是一个全局分割标准，因为它在构建 F （全局 MVE 投资组合）时考虑了所有叶节点投资组合。**CART** 专注于特定的叶节点，而我们的模型则不同，它会探索所有叶节点中的所有候选拆分，以找到投资改进最大的那个。

生长终止：所有后续的拆分过程与此类似。确定树叶数量是树生长的一个自然转折点，也是 **P-Tree** 唯一需要调整的参数。在基准规范中，我们认为 **P-Tree** 的叶子节点数为 $J+1=10$ 。此外，由于树叶需要作为基础组合，我们规定树叶生长的最小尺寸为 20，任何不符合这一标准的树叶都不会被进一步拆分。树生长过程结束后，会输出 **P-Tree** 分割序列、叶子基础组合 $\mathbf{R}_t^{(J)} = [\mathbf{R}_{1,t}^{(J)}, \dots, \mathbf{R}_{J+1,t}^{(J)}]$ 以及 **P-Tree** 因子 $f_t^{(J)}$ 。注意，每个叶组合都可以通过顺序拆分揭示的经济基本面来解释。

可能的扩展：**P-Tree** 框架非常灵活，可适应式子 (4) 中使用单个资产回报估算 MVE 投资组合的全局拆分标准之外的其他目标。例如，我们可以创建最小方差投资组合，或开发解释测试资产的潜在因素，如个人或基础投资组合回报。关键在于定义一个明确的经济目标，并利用树内的贪婪增长算法来探索广泛的观察聚类空间并优化目标。**P-Trees** 还允许灵活的贪婪性，可以在每次迭代中考虑多个分割，而不是传统的单一分割，或者通过沿特征的某些线性组合进行分割，沿多个特征同时进行分割。这些都留待今后探索。

2.3 Boosting P-Tree

Boosting 是一种将弱学习器组合成强学习器的 ML 技术（例如，Freund 和 Schapire, 1997; Rossi 和 Timmermann, 2015; Capponi 和 Yu, 2024）。在 MVE 框架下，**Boosting P-Trees** 会依次生长出一系列加法树，这些加法树会增强之前的树，这有助于利用额外的基础投资组合进一步跨越有效前沿，并生成多因子模型。考虑到现有因子，**Boosting** 程序通过创建额外的 **P-Tree** 因子无缝集成到 **P-Tree** 框架中，以最大化切线组合的集体夏普比率。**Boosting** 程序概述如下。

1. 初始因子用表示 $f_{1,t}$ ，可以是第一个单一 **P-Tree** 生成的因子，也可以是研

究人员选择的基准因子（如市场因子）。

2. 第二个因子 $f_{2,t}$ 由第二个 P-Tree 生成，以补充第一个因子。第二棵树的生长遵循 (4) 中的分割标准，但 $F = [f_{1,t}, f_{2,t}]$ 。MVE 投资组合由所有因子直接生成。

3. 重复步骤 2，依次生成 P 个因子 $F_t = [f_{1,t}, \dots, f_{p,t}]$ ，直到触发停止标准。

Boosting P-Tree 中的块结构：Boosting P-Tree 采用两步流程，通过整合多组叶基组合来生成 MVE 组合或 SDF。首先，计算每个 P-Tree 中所有叶基组合的树特定 MVE 组合，以确定当前的 P-Tree 因子。随后，将该因子与之前的所有 P-Tree 因子或基准合并，为多因子 P-Tree 模型创建一个全树 MVE 组合。通过顺序树块结构，Boosting P-Tree 间接对全树 MVE 投资组合权重进行了正则化，即使由此产生的夏普比率可能不会超过直接从所有 P-Tree 中的所有基础投资组合中产生的夏普比率。这避免了高维估计问题，进而实现了稳健的估计和样本外性能。

2.4 模拟试验

本节中的模拟练习强调了在应用 P-Tree 时使用样本外检验和包括大多数可观测特征的重要性，并展示了 P-Tree 捕捉产生数据的真实基本特征及其非线性交互作用的能力。

我们引入了三个真实的潜在特征、它们的交互作用以及非线性项。我们根据历史数据校准了收益生成过程（下一节将介绍），并包含了一组与生成数据的真实特征相关但不驱动收益的冗余特征。随后，我们进行了以下四组实验。

首先，我们比较了 P-Tree 测试资产与传统特征排序投资组合的夏普比率和阿尔法，后者未纳入特征交互作用。在各种信噪比水平下，无论是在样本内还是样本外，单个 P-Tree 和 Boosted P-Tree 始终生成具有更高夏普比率和更大未解释定价误差的叶投资组合。即使直接对真实特征进行排序，单变量或双变量排序投资组合的表现仍然不如基于高维特征生成的 P-Tree 测试资产。这一结果在后文的实证分析中得到了验证，即 P-Tree 测试资产比排序投资组合实现了更好的均值-方差效率（第 4.2 节），P-Tree 确实捕捉到了数据生成过程中的重要交互作用。

其次，一棵或几棵 P-Tree 的可解释性取决于其选择重要变量的能力。P-Tree 有助于从大量冗余或无用的变量中区分出真正数据生成的稀疏特征。通过反复模拟，P-Tree 在最初的几次拆分中就能高概率地选择出真正的特征。在信噪比较高的情况下，选择率的提高尤为明显。这一发现只能通过模拟来辨别，因为真实世界数据的稀疏特征仍然是未知的。一致的选择行为与第 5.1 节中评估特征重要性有关。

其次，单个 P-Tree 或少数 P-Tree 的可解释性依赖于其选择重要变量的能力。P-Tree 有助于从大量冗余或无用的特征中区分出真实数据生成过程中的稀疏特征。通过重复模拟，P-Tree 在前几次分割中以高概率选择了真实特征。这种改进的选择率在高信噪比场景中尤为明显。这一发现只能通过模拟来识别，因为真实数据的稀疏特征仍然是未知的。这种一致的选择行为与第 5.1 节中评估特征重要性的内容相关。

第三，我们将 P-Tree 模型在样本内和样本外表现之间的差异分解为两个部分：过拟合和学习限制，正如 Didisheim 等（2024）所讨论的。在真实数据中，由于市场数据生成过程未知，估计真实可预测性具有挑战性。然而，在我们的模拟中，真实可预测性被定义为使用“预言”特征（oracle characteristics）的 P-Tree 在样本外实现的夏普比率。我们发现过拟合和学习限制的正值，并且随着提升 P-Tree 数量的增加而增加，这支持了 Didisheim 等（2024）的结论，即随着模型参数的增加和观测数量的有限，过拟合和学习限制变得更加严重。在后文的实证分析中，我们

再次观察到 P-Tree 模型在样本内和样本外表现之间的巨大差距，这证明了我们关注样本外指标的合理性。

第四，我们评估了当 P-Tree 在输入特征中遗漏某些真实特征（但存在大量冗余或无用的特征）时的效率损失。这一实验评估了模型在不完整或无用信息下的表现，这种情况可能由多种原因引起。在我们校准的收益生成过程中，夏普比率和 CAPM 阿尔法显著下降，P-Tree 更频繁地选择冗余或无用的特征。因此，我们应包含所有可用的预测变量。在后文的历史数据分析中，我们确实使用了数据集中所有可用的公司特征。当 P-Tree 遗漏真实特征时，其性能显著下降，表明在资产定价中应尽可能包含所有可用的预测变量。

3 美国股票数据的横截面分割

P-Tree 框架可以轻松应用于公开股票或公司债券。我们以美国股票为例进行说明性应用。值得注意的是，在大型模型兴起的背景下，P-Tree 提供了一种可解释的替代方案，且不需要过多的计算资源。

3.1 美国股票数据

股票数据与特征：对美国股票的标准筛选条件（例如，与 Fama-French 因子构建相同）进行了应用。该股票池仅包括在 NYSE、AMEX 或 NASDAQ 上市超过一年的股票，并使用 CRSP 股票代码为 10 或 11 的公司的观测值。我们排除了账面价值为负或滞后市值为负的股票。我们使用 61 个公司特征，每月对每只股票进行观测，涵盖六个主要类别：**动量、价值、投资、盈利能力、摩擦（或规模）和无形资产。特征在横截面上标准化为范围[-1, 1]。**图表 5 列出了这些输入变量。

月度数据的时间范围为 1981 年至 2020 年。前 20 年的平均和月度股票观测值分别为 5265 和 4925，后 20 年分别为 4110 和 3837。我们对个体股票收益进行 1% 和 99% 的横截面缩尾处理，以减轻异常值的影响。整个 40 年样本用于基准分析。样本分为两部分——1981 年至 2000 年的前 20 年和 2001 年至 2020 年的最近 20 年——用于子样本稳健性检查以及训练和测试样本。

图表 5 股票特征

No.	Char.	Description	Category				
1	ABR	Abnormal returns around earnings announ.	Momentum	32	MOM12M	Cumulative returns in the past (2-12) months	Momentum
2	ACC	Operating accruals	Investment	33	MOM1M	Previous month return	Momentum
3	ADM	Advertising expense-to-market	Intangibles	34	MOM36M	Cumulative returns in the past (13-35) months	Momentum
4	AGR	Asset growth	Investment	35	MOM60M	Cumulative returns in the past (13-60) months	Momentum
5	ALM	Quarterly asset liquidity	Intangibles	36	MOM6M	Cumulative returns in the past (2-6) months	Momentum
6	ATO	Asset turnover	Profitability	37	NI	Net equity issue	Investment
7	BASPREAD	Bid-ask spread (3 months)	Frictions	38	NINCR	Number of earnings increases	Momentum
8	BETA	Beta (3 months)	Frictions	39	NOA	Net operating assets	Investment
9	BM	Book-to-market equity	Value-versus-growth	40	OP	Operating profitability	Profitability
10	BM_IA	Industry-adjusted book to market	Value-versus-growth	41	PCTACC	Percent operating accruals	Investment
11	CASH	Cash holdings	Value-versus-growth	42	PM	Profit margin	Profitability
12	CASHDEBT	Cash to debt	Value-versus-growth	43	PS	Performance Score	Profitability
13	CFP	Cashflow-to-price	Value-versus-growth	44	RD_SALE	R&D-to-sales	Intangibles
14	CHCSHO	Change in shares outstanding	Investment	45	RDM	R&D-to-market	Intangibles
15	CHPM	Change in Profit margin	Profitability	46	RE	Revisions in analysts' earnings forecasts	Intangibles
16	CHTX	Change in tax expense	Momentum	47	RNA	Return on net operating assets	Profitability
17	CINVEST	Corporate investment	Investment	48	ROA	Return on assets	Profitability
18	DEPR	Depreciation/PP&E	Momentum	49	ROE	Return on equity	Profitability
19	DOLVOL	Dollar trading volume	Frictions	50	RSUP	Revenue surprise	Momentum
20	DY	Dividend yield	Value-versus-growth	51	RVAR_CAPM	Idiosyncratic volatility - CAPM (3 months)	Frictions
21	EP	Earnings-to-price	Value-versus-growth	52	RVAR_FF3	Res. var. - Fama-French 3 factors (3 months)	Frictions
22	GMA	Gross profitability	Investment	53	SVAR	Return variance (3 months)	Frictions
23	GRLTNOA	Growth in long-term net operating assets	Investment	54	SEAS1A	1-Year Seasonality	Intangibles
24	HERF	Industry sales concentration	Intangibles	55	SGR	Sales growth	Value-versus-growth
25	HIRE	Employee growth rate	Intangibles	56	SP	Sales-to-price	Value-versus-growth
26	ILL	Illiquidity rolling (3 months)	Frictions	57	STD_DOLVOL	Std of dollar trading volume (3 months)	Frictions
27	LEV	Leverage	Value-versus-growth	58	STD_TURN	Std. of Share turnover (3 months)	Frictions
28	LGR	Growth in long-term debt	Investment	59	SUE	Standardized unexpected quarterly earnings	Momentum
29	MAXRET	Maximum daily returns (3 months)	Frictions	60	TURN	Shares turnover	Frictions
30	ME	Market equity	Frictions	61	ZEROTRADE	Number of zero-trading days (3 months)	Frictions
31	ME_IA	Industry-adjusted size	Frictions				

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

宏观经济变量：此外，我们使用 10 个宏观变量来捕捉潜在的经济体制变化。图表 6 总结了这些宏观变量，包括市场时机的宏观预测变量、债券市场预测变量以及 S&P 500 的总体特征。我们通过过去 10 年的历史百分位数对这些宏观预测变量数据进行标准化。这种滚动窗口数据标准化在比较预测变量水平以检测不同宏观经济体制时非常有用。

图表 6 宏观经济变量

No.	Variable name	Description
1	DFY	Default yield
2	DY	Dividend yield of S&P 500
3	EP	Earnings-to-price of S&P 500
4	ILL	Pastor-Stambaugh illiquidity
5	INFL	Inflation
6	LEV	Leverage of S&P 500
7	NI	Net equity issuance of S&P 500
8	SMVAR	Stock Market (S&P 500) Volatility
9	TBL	Three-month treasury bill rate
10	TMS	Term spread

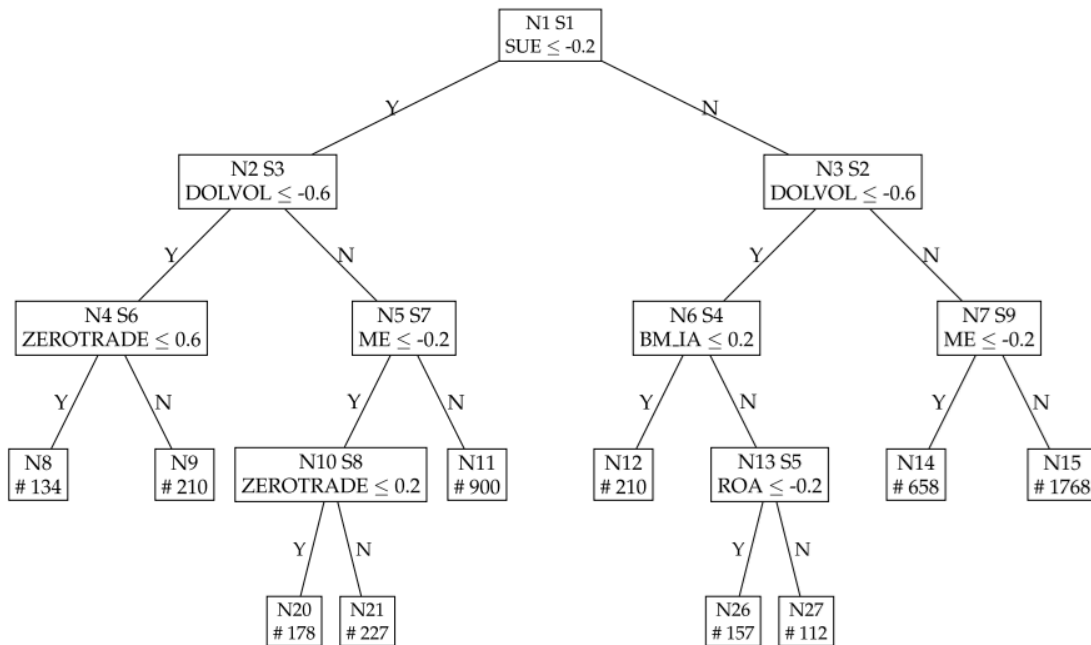
资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：上表列出了实证研究中使用的宏变量的描述。

3.2 可视化单个 P-Tree

图表 7 是 P-Tree 图。在每个叶节点中，S#代表连续拆分的顺序，N#是节点索引。我们在末端叶片中提供了投资组合规模的数字：股票观察的月度中位数。拆分规则是选定的拆分特征和横截面五分位切点[-0.6, -0.2, 0.2, 0.6]。在第一次拆分前，树从根节点（N1）开始生长，其叶子基础组合代表价值加权市场组合。

图表 7 1981 年到 2020 年的面板树



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：上图提供拆分特征、每个父节点的切割点值，以及它们各自的节点和拆分索引。例如，节点 N1 在 -0.2（40% 百分位）处被 SUE 作为第一个拆分 S1，第二个拆分 S2 在 -0.6（20% 百分位数）的节点上被 DOLVOL 拆分。终端叶节点投资组合中的每月资产数量中位数也包括在内。例如，节点 N8 的每月中位数有 134 只股票。图表 5 描述了股票特征。

数据驱动的 P-Tree 首先沿着标准化意外季度收益（SUE, Rendleman 等，1982 年）在 -0.2（40% 百分位数）处分割。分割后，40% 的股票进入左侧叶片（N2），60% 的股票进入右侧叶片（N3）。然后，第二次拆分是根据美元交易量（DOLVOL, Chordia 等，2001 年）进行的，位于右侧叶片（N3）的 -0.6 处，第三次拆分也是根据 DOLVOL 进行的，位于左侧叶片（N2）的 -0.6 处。此外，后续拆分还包括行业调整后的账面市值比（BM_IA）、股本回报率（ROA）、零交易（ZEROTRADE）和市场股本（ME）。经过九次拆分后，我们停止 P-Tree 增长，得到 10 个叶基组合。

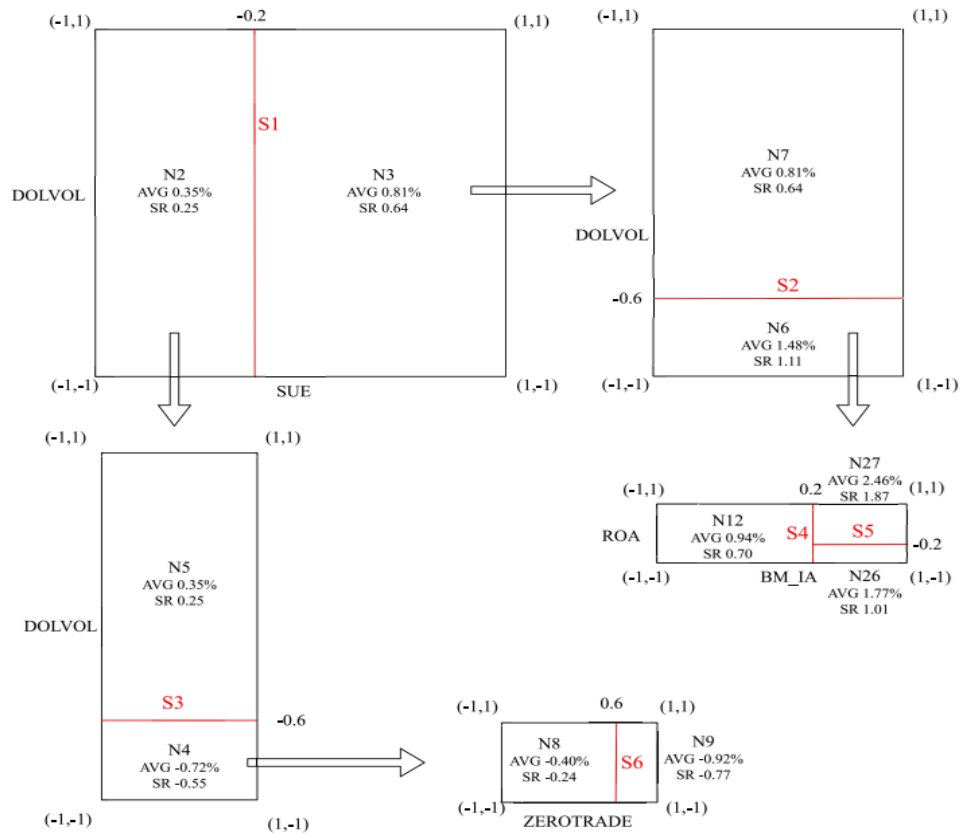
聚类模式和非对称互动：P-Tree 根据基本特征对类似资产进行聚类，揭示了均值-方差多样化的来源。 图表 7 显示了用于分割横截面的特征（非对称）相互作用。通过共同定义与叶节点相对应的分区，P-Tree 可以了解同一路径中出现的特征之间的相互作用。例如，流动性的 ZEROTRADE（Liu, 2006）是进一步分割低 SUE 低 DOLVOL 股票的重要指标。然而，对于低 SUE 非低 DOLVOL 股票，规模 ME（Banz, 1981）可能是 MVE 框架下进一步拆分的更好指标。

目前的文献已经认识到两个特征之间的相互作用，但还需要更系统的研究。例如，Lee 和 Swaminathan（2000）通过与交易量的交互作用来增强动量策略，Da 等（2014）发现连续信息公司的动量效应比离散信息公司更强。我们的框架允许探索多种特征之间的相互作用，超越了典型的二元关系。图表 7 演示了这一点，图中可以识别出至少涉及四个特征（如 SUE、DOLVOL、ME 和 ZEROTRADE）的交互路径。

图表 8 中的分区图是可视化聚类和非对称模式的另一种方法。我们报告了每个

叶片的月平均超额收益和年化夏普比率。首先，P-Tree 在 SUE 为-0.2 时进行分割，从而得到一个低 SUE 投资组合，包括横截面中 40% 的股票，以及一个由 60% 的股票组成的高 SUE 投资组合。N2 和 N3 之间的月预期收益率差为 0.46%，夏普比率介于 0.25 和 0.64 之间。

图表 8 分区图可视化非线性交互



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：此图说明了图表 7 中树结构的前几次拆分的分区。第一次拆分（S1）发生在整个 stock universe 上 SUE 的 40% 处，第二次拆分发生在高 SUE 分区上 DOLVOL 的 80% 处。提供了每个分区的投资组合结果（月平均超额回报率（AVG）和年化夏普比率（SR））。箭头表示在与上一个分割的分区区域上实施下一个分割。

其次，我们对 N3 的 DOLVOL 以-0.6 进行拆分，以获得 N6 和 N7，其中低 DOLVOL 投资组合 N6 的预期收益和夏普比率均高于 N7。第三，我们将 DOLVOL 拆分为 N2 上的-0.6，收获 N4 和 N5，其中低 DOLVOL 投资组合 N4 的预期收益和夏普比率低于 N5。我们发现 DOLVOL 对 N2 和 N3 有不同的影响。在高 SUE 方面，DOLVOL 与资产回报呈正相关。然而，在低 SUE 方面，DOLVOL 与资产回报负相关。这是 SUE 与 DOLVOL 之间非对称互动的一个例子。一个简单的交易策略是做空 N4（低 SUE - 低 DOLVOL 投资组合），做多 N6（高 SUE - 低 DOLVOL 投资组合），每月预期收益率超过 2%。叶片之间的收益差距表明，通过特征的非对称交互作用分割横截面是非常有用的。

3.3 P-Tree 叶节点投资组合

资产聚类：P-Tree 生成叶片，又称叶节点投资组合。与 CART 中纯粹预测

的标量输出不同，P-Tree 的叶子代表了来自时间不变的 PTree 结构的投资组合回报的时间序列。这些叶节点投资组合是非线性、交互式和高维特征管理投资组合。图表9面板 A 总结了图表5中第一个 P-Tree 生成的叶节点投资组合：节点指数、叶节点投资组合中股票数量中位数、平均回报、CAPM α (%)、 β 、时间序列回归 R^2 ，以及与 FF5、Q5、RP-PCA 和 IPCA 五因子模型 (Fama and French, 2015; Hou et al, 2021; Kelly 等, 2019; Lettau 和 Pelger, 2020)。

图表9 叶节点投资组合的评估

ID	# Median	AVG	STD	α_{CAPM}	β_{CAPM}	R^2_{CAPM}	α_{FF5}	α_{Q5}	α_{RPS}	α_{IPS}
Panel A: 40 Years (1981–2020)										
N8	134	-0.40	5.76	-0.97***	0.84	0.42	-0.91***	-0.66***	-1.51***	-0.58*
N9	210	-0.92***	4.14	-1.35***	0.63	0.47	-1.46***	-1.25***	-1.84***	-0.36*
N20	178	-0.34	8.89	-1.34***	1.45	0.53	-1.10	-0.49*	-1.55***	-1.78***
N21	227	-1.14***	5.90	-1.83***	1.00	0.58	-1.83***	-1.56***	-2.34***	-1.19***
N11	900	0.36*	4.81	-0.35***	1.04	0.93	-0.30***	-0.15**	-0.29***	0.16
N12	210	0.94***	4.63	0.42**	0.76	0.54	0.29**	0.48***	-0.22	1.17***
N26	157	1.77***	6.04	1.16***	0.88	0.43	1.08***	1.37***	0.40**	1.40***
N27	112	2.46***	4.55	1.97***	0.71	0.49	1.80***	2.00***	1.18***	2.18***
N14	658	1.03***	6.88	0.21	1.18	0.59	0.32*	0.64***	-0.33**	0.06
N15	1768	0.81***	4.41	0.14***	0.98	0.98	0.10***	0.03	-0.26***	0.29
Panel B: 20 Years (1981–2000)										
N16	97	0.25	5.99	-0.34	0.81	0.36	-0.27	-0.08	-0.27	-0.02
N17	256	-0.74***	4.40	-1.16***	0.57	0.33	-1.36***	-1.17***	-1.56***	-1.47***
N18	95	1.72***	5.72	1.17***	0.75	0.34	1.07***	1.27***	0.84**	0.03
N19	48	2.91***	4.97	2.44***	0.64	0.33	2.13***	2.50***	2.36***	1.87***
N5	3443	0.76***	4.39	0.03*	0.99	1.00	-0.02*	-0.04***	-0.49***	1.34***
N24	160	-3.53***	7.75	-4.36***	1.13	0.42	-3.90***	-3.29***	-3.58***	0.79
N25	250	-2.82***	6.85	-3.58***	1.04	0.45	-3.20***	-2.54***	-1.89***	1.55*
N13	286	-1.75***	9.94	-2.86***	1.52	0.46	-1.58***	-1.19***	-1.29***	2.22**
N14	217	-1.21*	9.60	-2.21***	1.37	0.40	-1.16***	-0.63	-0.09	4.52**
N15	116	1.23*	10.07	0.16	1.46	0.41	1.12***	1.52***	1.39**	3.58
Panel C: 20 Years (2001–2020)										
N8	153	-0.20	4.44	-0.63**	0.67	0.45	-0.63**	-0.46**	-0.93***	-0.10
N18	58	-0.31	6.82	-0.96***	1.01	0.44	-0.93***	-0.53*	-1.36***	-0.88**
N19	112	-0.90***	4.72	-1.39***	0.76	0.53	-1.39***	-1.31***	-1.89***	-0.93***
N5	1206	0.35	5.23	-0.38***	1.13	0.95	-0.27***	-0.25***	-0.21***	0.37*
N24	171	0.94**	6.47	0.26	1.05	0.53	0.31	0.57**	-0.13	0.59**
N25	104	2.44***	6.80	1.79***	1.02	0.45	1.88***	2.17***	1.16***	1.43***
N52	47	1.68***	5.28	1.17***	0.79	0.46	1.17***	1.45***	0.52**	1.56***
N53	76	3.53***	5.27	2.97***	0.86	0.54	2.88***	3.10***	2.24***	3.19***
N27	76	1.79***	4.05	1.36***	0.66	0.53	1.31***	1.37***	0.92***	1.6***
N7	1896	0.76***	4.27	0.16***	0.94	0.98	0.10**	0.09**	-0.05	0.63***

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：*、** 和 *** 分别表示 10%、5% 和 1% 的显著性水平。面板 A 对应于图表 7，面板 B 和 C 分别对应于子集的结果。

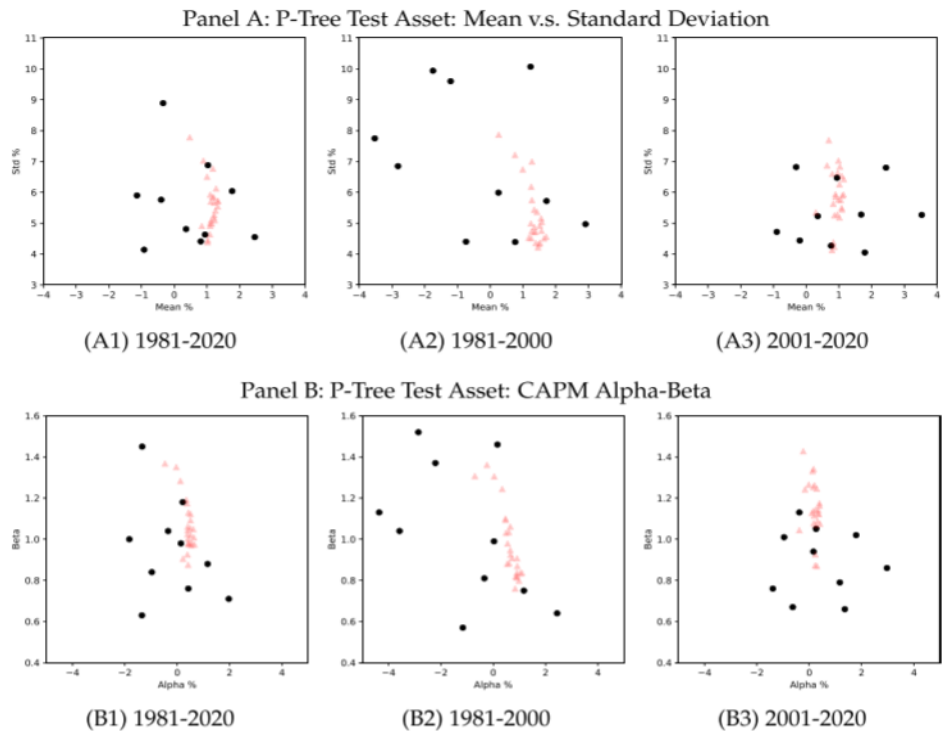
我们观察到两片大叶子：N11 包含 900 只股票，N15 包含每月中位数超过 1700 只股票。它们的 CAPM R^2 超过 90%，而 β 接近 1，这意味着它们的回报时间序列与市场回报高度相关。这两片叶子的夏普比率较高，但在 MVE 框架下，其他叶子的投资改善程度较高。例如，N21 代表低 SUE、高 DOLVOL、低 ME 和高 ZEROTRADE 叶子，其平均回报率为 -1.14%，CAPM α 为 -1.83%；N27 代表高 SUE、低 DOLVOL、高 BM_IA 和高 ROA 叶子，其平均回报率为 2.46%，CAPM α 为 1.97%。通过采用买入 N27 和卖空 N21 的简单月度再平衡策略，可以预期平均回报率为 3.6%。

这些叶节点投资组合是通过 61 个特征的交互式非线性拆分生成的。我们预计，使用流行的线性因子模型很难对它们进行定价。在 10 个叶节点投资组合中，有 9 个叶节点投资组合具有显著的 CAPM 指数。在 10% 的置信水平下，FF5、Q5、RP-PCA 5 和 IPCA 5 因子模型在统计上显著的非零值分别为 9、9、9 和 7。P-Tree 是一种以目标为导向的聚类算法，可最大化。

Q-Tree 是一种以目标为导向的聚类算法，可最大限度地提高基线规范中的夏普比率。因此，我们预计这些测试资产与那些“不足”的基准因子模型相比，在经济上具有较大的统计显著性。有了这样独特的基础投资组合，我们预计它们将实现更高的均值-方差效率，跨越基准因子模型未能跨越的有效前沿。图表 9 的 B 和 C 面板提供了前 20 年和后 20 年的子样本分析。虽然由于重新拟合，树形结构和叶片特征发生了变化，但我们仍然发现了类似的模式。例如，使用基准因子模型仍然很难根据叶子因子对这些投资组合进行定价。

多样化测试资产。 图表 10 面板 A 描述了 P-Tree 测试资产的绩效统计。子图 A1 用黑色圆圈表示 1981 年至 2020 年 40 年样本中第一个 P-Tree 的叶节点投资组合的均值-标准差散点图。预期收益率在 -2% 到 3% 之间，标准差在 4% 到 9% 之间。我们观察到这些投资组合的预期收益和风险存在很大差异，这意味着它们在 MVE 框架下是分散的。在 A2 和 A3 的子样本分析中，投资组合之间存在显著差异，较早的样本显示出更大的分散性，而较晚的样本则更加集中。相比之下，浅红色三角形中的 5 × 5 ME-BM 投资组合的平均回报率集中在同一水平，其分散程度远低于 P-Tree 检验资产。这一结果表明，在 MVE 框架下，P-Tree 测试资产比 ME-BM 投资组合更加分散。

图表 10 多样化的 P-Tree 测试资产



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：此图报告了 P-Tree 测试资产(如图表 9 所示)和 5 × 5 ME-BM 投资组合的性能。10 个黑色圆圈表示 P-Tree 测试资产，25 个浅红色三角形表示 ME-BM 投资组合。面板 A 显示了投资组合收益的平均值和标准差的散点图，单位为百分比。面板 B 显示了 CAPM α 和 β 的散点图，其中 α 以百分比表示。

图表 10 的 B 面板显示了 CAPM 模型推导的阿尔法-贝塔图。在子集样本中，

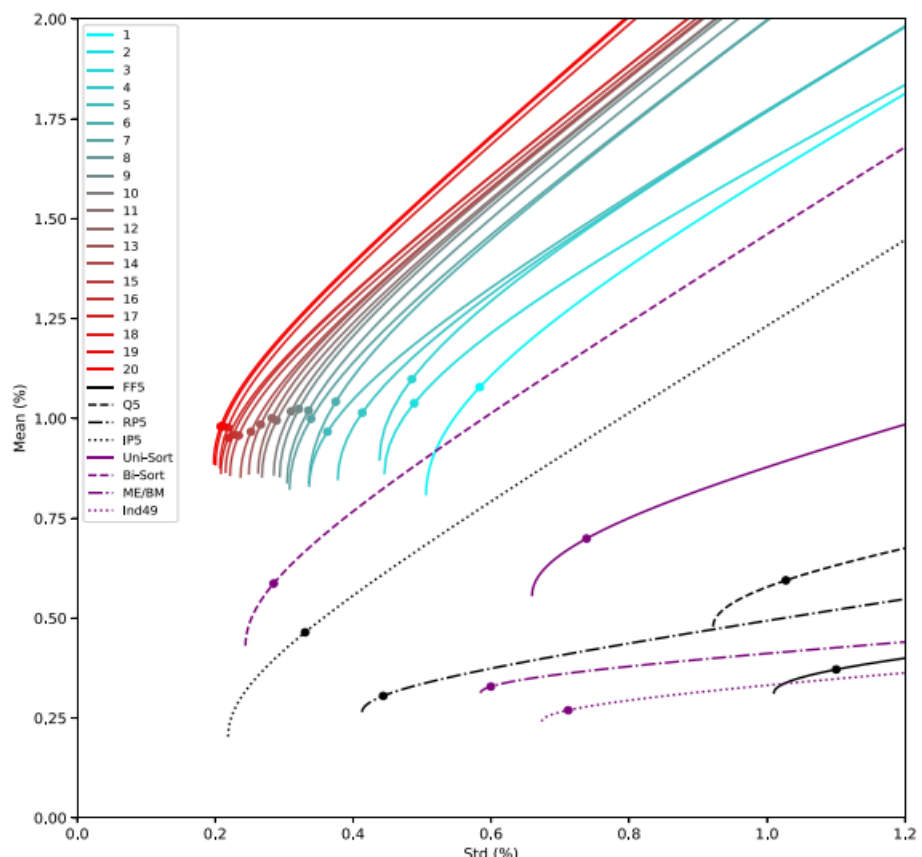
我们发现 P-Tree 测试资产的 CAPM β S 值分散在 0.6 到 1.6 的范围内。CAPM 均值约为零，范围在 -2% 到 2% 之间。这些基础投资组合的 CAPM α 和 β 变化很大。在 B1 中，前 20 年的阿尔法和贝塔值范围较大，而在 B2 中，后 20 年的变化较小。与 P-Tree 测试资产相比，5×5 ME-BM 投资组合更容易用 CAPM 模型解释，因为在所有数据中，这些投资组合都与阿尔法为零的垂直线非常接近，多样化程度更容易忽略不计。总之，这是一个非常积极的证据，表明 P-Tree 检验资产的分散化模式优于传统资产，即使是用 CAPM 定价时也是如此。

4 用于多因子和测试资产的 Boosting P-Tree

4.1 拓展有效前沿

我们的研究重点是使用 (Boosting) P-Tree 生成的测试资产所跨越的有效前沿。许多 P-Tree 测试资产比多个众所周知的基准因子模型具有显著的优势，这表明这些因子模型不能有效地跨越有效前沿。**Boosting 技术增强了我们生成更有效测试资产的能力，从而更好地扩展有效前沿，即“在 P-Tree 上扩展有效前沿”。**

图表 11 用 P-Tree 测试资产



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：此图显示了第一个 P 树的 MVE 边界、提升的 P 树因子模型、基准因子模型和基准测试资产的顺序。边界上的点表示相切组合。平均值和标准差以月为单位。样本期为 1981 年至 2020 年。

未跨越的有效前沿：图表 11 展示了 1981 年至 2020 年样本期内 P-Tree 和提升 P-Tree 的 MVE 前沿，颜色从蓝色渐变到红色。基准因子模型以黑色显示，基准测试资产以紫色显示。值得注意的是，P-Tree1 前沿（由第一个 P-Tree 测试资产生成）已经比任何基准因子模型和测试资产跨越的前沿更有效。在基准测试资产中，285 个双变量排序投资组合跨越的前沿与 P-Tree1 前沿非常接近。然而，P-Tree1 仅包含 10 个投资组合，远少于 285 个。在因子模型中，IPCA 五因子模型是一个强有力的候选者，但仍然不如 P-Tree1 前沿有效。总体而言，这 10 个叶基础投资组合是跨越有效前沿的有希望的测试资产，在 MVE 框架下比基准更有效。**换句话说，列出的基准因子模型和测试资产不足以跨越真正的前沿。**

Boosting 经验前沿：在第一个 P-Tree 之后，Boosting 的 P-Tree 将投资组合效率从一个因子提高到 20 个因子。图表 11 中，跨越的前沿向均值-方差图的左上角移动。在 20 个 P-Trees 中，我们的方法一直在提高有效前沿。由于维度诅咒，在一个 MVE 投资组合中组合 200 种资产具有挑战性。与 Ait-Sahalia 和 Xiu（2017 年）使用外生行业分类构建块对角模式不同，我们的 Boosting P-Trees 在收益协方差中采用了内生块对角结构。我们将每个 P-Tree 的 10 个叶基组合合并为一个 P-Tree 因子，然后估计 Boosting 后的 P-Tree 因子的切线组合。**因此，在 MVE 框架下，我们可以生成多个因子，并在一个前沿中组合多组测试资产。**

从一棵 P-Tree 到 20 棵 P-Tree 会产生更有效的前沿，使投资者在承担相同风险的情况下获得更高的预期收益。虽然 P-Tree1 和 P-Tree10 的前沿之间存在显著差异，但 P-Tree10 之后的前沿线增长缓慢，这表明随着每一次 Boosting，均值-方差效率的改善幅度较小，可以忽略不计。

我们采用资产定价测试来证明 Boosting P-Tree 的增量效率。首先，我们展示了每个 Boosting 的 P-Tree 因子和多因子模型的夏普比率。其次，我们用 CAPM 和 FF5 对每个 Boosting 因子进行评估，检查是否存在显著的阿尔法。图表 9 所示，P-Tree 检验资产很难定价，因此我们希望 P-Tree 因子能通过 CAPM 检验和 FF5 检验。第三，我们将每个 P-Tree 因子与之前为扩展因子测试生成的所有 Boosting P-Tree 因子进行回归。如果之前的 P-Tree 因子没有跨越新增的 P-Tree 因子，前沿的效率就会提高。

测试结果见图表 12，每个 Boosting 因子的夏普比率都高于 1，即使是第 20 个因子也是如此。同时，累计多因子夏普比率从 6 到 15 单调递增，这意味着 **Boosting 后的 P-Tree 的测试资产在现有测试资产的基础上增加了增量定价信息。**此外，所有 CAPM 和 FF5 均为正且高度显著，表明所有 P-Tree 增强因子都无法用 CAPM 或 FF5 解释。在扩展因子检验中，我们发现了显著的系数、较大的 t 统计量和 R^2 较高的时间序列。在应用嵌套资产定价检验（Barillas 和 Shanken，2017 年）后，我们没有发现任何证据支持有足够的因子模型来解释 Boosting P-Tree 检验中的资产。**相比之下，用于检验资产定价模型的传统单变量或双变量检验资产并不能完全覆盖有效前沿。**

图表 12 测试 Boosting P-Tree 的生成

	Sharpe ratio		CAPM test		FF5 test		Expanding factors test			BS test
	Single	Cumu.	α (%)	t-stat	α (%)	t-stat	α (%)	t-stat	R ²	p-value
Panel A: 40 Years (1981–2020)										
1	6.37	6.37	1.39	35.36	1.37	35.81	-	-	-	
2	3.20	7.35	0.52	17.65	0.48	15.27	0.62	9.55	0.01	0.00
3	1.18	7.80	0.34	5.25	0.18	3.28	-0.86	-5.43	0.26	0.00
4	2.06	8.46	0.44	11.22	0.38	9.56	0.69	7.63	0.16	0.00
5	1.99	9.18	0.48	10.25	0.41	9.25	0.84	5.61	0.21	0.00
6	1.01	9.57	0.18	4.24	0.08	2.94	-0.50	-5.47	0.45	0.00
7	1.42	10.11	0.28	7.62	0.22	6.67	0.63	7.03	0.36	0.00
8	1.32	10.40	0.28	7.14	0.20	4.95	-0.50	-5.33	0.41	0.00
9	1.83	10.88	0.53	10.11	0.43	9.38	0.85	6.88	0.34	0.00
10	1.48	11.20	0.44	7.35	0.30	7.30	-0.68	-5.49	0.46	0.00
11	1.78	11.72	0.38	10.33	0.31	8.62	0.72	6.70	0.30	0.00
12	1.02	12.06	0.20	4.68	0.10	2.99	-0.55	-5.76	0.55	0.00
13	1.37	12.57	0.29	8.38	0.22	6.48	0.76	6.01	0.33	0.00
14	1.37	13.01	0.48	5.93	0.32	5.86	-0.91	-5.68	0.60	0.00
15	1.37	13.81	0.31	6.58	0.21	5.66	0.97	7.11	0.48	0.00
16	1.24	14.28	0.28	6.23	0.17	4.11	-0.74	-6.24	0.54	0.00
17	1.54	14.60	0.46	8.16	0.40	7.52	-0.89	-4.80	0.40	0.00
18	1.64	14.92	0.32	8.43	0.27	7.78	-0.65	-5.43	0.34	0.00
19	1.48	15.43	0.43	8.63	0.36	7.47	1.18	5.86	0.34	0.00
20	1.35	15.63	0.33	7.38	0.23	6.19	-0.59	-4.19	0.44	0.00
Panel B: 20 Years (1981–2000)										
1	7.12	7.12	1.86	29.62	1.78	30.22	-	-	-	
2	1.09	7.75	0.25	3.66	0.11	1.91	-0.83	-4.80	0.28	0.00
3	2.26	9.45	0.39	9.72	0.34	8.48	0.92	8.19	0.17	0.00
4	6.48	11.35	1.80	26.96	1.76	27.06	1.75	10.16	0.03	0.00
5	1.59	12.71	0.33	6.50	0.28	5.77	1.11	6.85	0.36	0.00
6	1.65	13.67	0.37	5.87	0.26	4.57	-1.03	-5.34	0.40	0.00
7	1.72	14.94	0.33	6.99	0.32	6.17	1.25	5.67	0.20	0.00
8	1.52	15.78	0.39	6.10	0.31	4.93	-1.32	-5.12	0.35	0.00
9	3.02	17.90	1.03	14.00	0.93	13.08	2.69	8.42	0.26	0.00
10	1.93	19.18	0.83	8.12	0.67	7.14	-2.41	-6.27	0.52	0.00
11	1.34	20.72	0.38	4.93	0.19	2.76	-2.18	-7.33	0.45	0.00
12	2.22	22.89	0.36	9.38	0.31	8.03	1.46	7.36	0.39	0.00
13	1.51	24.14	0.35	4.88	0.23	3.69	1.82	6.06	0.43	0.00
14	2.11	25.71	0.58	7.80	0.48	7.57	-2.43	-5.93	0.40	0.00
15	2.39	28.37	0.87	10.14	0.76	8.70	3.69	7.64	0.52	0.00
16	2.53	30.37	0.53	9.58	0.44	7.91	2.40	7.12	0.31	0.00
17	1.50	32.42	0.30	5.66	0.31	5.08	2.65	7.18	0.35	0.00
18	1.63	33.89	0.37	5.95	0.28	4.99	-2.79	-5.64	0.37	0.00
19	1.84	35.78	0.47	7.04	0.37	5.93	3.54	7.35	0.30	0.00
20	2.70	37.94	0.89	9.85	0.79	9.06	-4.67	-6.91	0.44	0.00
Panel C: 20 Years (2001–2020)										
1	5.82	5.82	1.51	24.22	1.47	24.26	-	-	-	
2	2.14	6.64	0.58	8.74	0.52	8.26	0.91	6.44	0.03	0.00
3	0.91	7.53	0.18	3.28	0.09	2.01	-0.86	-5.82	0.36	0.00
4	1.35	8.60	0.29	6.18	0.31	6.04	1.00	8.65	0.17	0.00
5	1.08	9.30	0.31	4.14	0.26	3.88	-1.02	-6.38	0.42	0.00
6	2.63	10.36	0.60	12.15	0.56	11.11	1.04	7.61	0.12	0.00
7	1.39	11.30	0.29	5.53	0.26	5.16	1.00	7.23	0.28	0.00
8	1.60	11.93	0.40	5.76	0.30	5.47	-0.85	-5.62	0.44	0.00
9	1.04	12.73	0.21	4.53	0.18	3.83	-1.01	-7.44	0.36	0.00
10	1.64	14.32	0.30	5.96	0.25	5.31	1.11	7.80	0.44	0.00
11	1.69	15.73	0.36	6.73	0.30	6.20	1.43	6.79	0.33	0.00
12	1.58	16.82	0.60	6.08	0.51	5.36	2.48	5.72	0.24	0.00
13	1.48	18.06	0.41	5.29	0.29	4.47	-1.77	-5.52	0.47	0.00
14	1.39	19.46	0.32	5.22	0.26	4.62	1.68	6.05	0.40	0.00
15	1.48	20.60	0.35	5.51	0.21	4.34	-1.38	-6.25	0.58	0.00
16	1.40	21.91	0.34	5.84	0.27	4.82	1.84	6.29	0.46	0.00
17	1.27	23.41	0.39	4.82	0.25	3.46	-2.55	-6.99	0.51	0.00
18	1.44	24.36	0.34	5.16	0.28	4.72	-1.61	-4.97	0.48	0.00
19	1.58	25.45	0.33	6.59	0.28	6.08	-1.74	-5.20	0.42	0.00
20	1.88	26.52	0.58	7.12	0.50	6.53	-2.76	-4.96	0.39	0.00

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：测试 Boosting P-Tree 生长。上表显示了每个顺序生成的提升 P 树因子的性能，包括夏普比率、CAPM 检验、FF5 检验、膨胀因子检验和 (Barillas 和 Shanken, 2017) 检验。面板 A、B 和 C 报告不同的样本周期。“Single”列显示单个因子和“Cumu”的夏普比率。(累积)列显示从第一个 P 树因子到当前 P 树因子的多个因子跨越的 MVE 的夏普比率。这些列报告 CAPM 和 FF5 跨度回归测试的 α (%) 和 t 统计量。对于扩展因子检验，我们将每个 P 树因子回归到所有先前提升的 P 树因子上，并报告 α (%)、t 统计量和 R²。

4.2 利用 Boosting P-Tree 生成测试资产

首先，我们展示了叶节点投资组合的定价性能，这意味着它们可以作为多样化的测试资产，并对 FF5 模型构成挑战。然后，我们利用 Boosting P-Trees 生成多组基础组合。根据经验，我们最多可以生成 20 棵 P-Trees，产生 200 个叶节点投资组合。

P-Tree 测试资产。图 13 面板 A 列出了投资组合的数量、GRS 检验统计量及其 p 值、PY 检验的 p 值、平均绝对 α 、均方根 α 、平均时间序列回归 R^2 以及 10%、5% 和 1% 置信水平的显著 alpha 百分比。在各行中，P-Tree1 代表一组测试资产，包括第一个 P-Tree 中的 10 个叶节点投资组合。P-Tree1-5 包含前五个 P-Tree 生成的所有叶节点投资组合，P-Tree6-10 包含第 6 个至第 10 个 P-Tree。类似地，我们定义了 P-Tree11-15、P-Tree16-20 和 P-Tree1-20。

我们发现 GRS 检验和 PY 检验总是拒绝所有 PTree 检验资产规格下检验资产的阿尔法共同为零的零假设，这意味着 **P-Tree 检验资产的预期收益率难以用 FF5 模型解释**。此外，最后三列显示，许多资产的阿尔法值在统计意义上是显著的。为便于比较，我们纳入了文献中常用的四组测试资产：150 个 10×1 单变量排序投资组合、285 个双变量排序投资组合、 5×5 ME-BM 投资组合和 49 个行业投资组合，所有这些都可以从 Ken French 的网站上下载。与这些基准测试资产相比，P-Tree 测试资产的 GRS 测试统计量和阿尔法值更大。

此外，我们还观察到从第一个 P-Tree 到随后的 Boosting P-Trees，阿尔法值都在下降，这表明向有效前沿极限的边际收敛是递减的。与后续 P-Tree 生成的测试资产相比，第一个 P-Tree 的测试资产具有更大的 GRS 统计量、更高的平均阿尔法值和更高的显著阿尔法值比例。此外，我们还看到

总体而言，由 20 个高维特征 P-Tree 创建的这 200 个测试资产对 FF5 模型提出了挑战。它们比基准测试资产更难定价，为测试因子模型设定了高标准，回应了实证研究的关切 (Lewellen 等, 2010; Ang 等, 2020)。

为稳健起见，图表 13 面板 B 和面板 C 报告了前 20 年和后 20 年的子样本。在最近的 20 年中，FF5 在基准测试资产的定价方面表现良好；也就是说，GRS 和 PY 测试不能在 10% 的水平上被拒绝，49 个行业组合和单变量排序组合的显著性较低。然而，P-Tree 检验资产总是拒绝 GRS 和 PY 检验的零假设，并且在两个子样本中的定价始终具有挑战性。

图表 13 比较测试资产

	N	GRS	p-GRS	p-PY	$ \bar{\alpha} $	$\sqrt{\alpha^2}$	\bar{R}^2	% $\alpha_{10\%}$	% $\alpha_{5\%}$	% $\alpha_{1\%}$
Panel A: 40 Years (1981-2020)										
P-Tree1	10	141.27	0.00	0.00	0.92	1.11	75	100	90	80
P-Tree1-5	50	60.32	0.00	0.00	0.44	0.61	80	70	62	44
P-Tree6-10	50	4.60	0.00	0.00	0.29	0.37	79	56	50	34
P-Tree11-15	50	4.74	0.00	0.00	0.20	0.26	80	38	36	24
P-Tree16-20	50	4.21	0.00	0.00	0.31	0.42	77	52	44	30
P-Tree1-20	200	41.31	0.00	0.00	0.31	0.43	79	54	48	33
Uni-Sort	150	1.62	0.00	0.00	0.10	0.14	88	25	18	7
Bi-Sort	285	2.50	0.00	0.00	0.12	0.17	89	30	23	15
ME-BM	25	5.01	0.00	0.00	0.12	0.16	92	36	28	20
Ind49	49	1.99	0.00	0.00	0.28	0.35	60	39	31	18
Panel B: 20 Years (1981-2000)										
P-Tree1	10	84.36	0.00	0.00	1.58	1.95	70	90	80	80
P-Tree1-5	50	50.84	0.00	0.00	0.79	1.26	76	62	60	52
P-Tree6-10	50	7.27	0.00	0.00	0.58	0.87	75	56	44	38
P-Tree11-15	50	6.39	0.00	0.00	0.55	0.82	76	66	60	42
P-Tree16-20	50	8.42	0.00	0.00	0.52	0.74	76	62	54	50
P-Tree1-20	200	112.90	0.00	0.00	0.61	0.95	76	62	55	46
Uni-Sort	150	1.94	0.00	0.00	0.17	0.23	88	35	27	19
Bi-Sort	285	-	-	0.00	0.21	0.30	89	42	33	21
ME-BM	25	4.75	0.00	0.00	0.21	0.25	91	56	48	32
Ind49	49	2.44	0.00	0.00	0.52	0.61	61	59	49	27
Panel C: 20 Years (2001-2020)										
P-Tree1	10	56.76	0.00	0.00	1.09	1.35	68	90	90	90
P-Tree1-5	50	30.35	0.00	0.00	0.43	0.68	76	52	38	24
P-Tree6-10	50	5.17	0.00	0.00	0.29	0.37	75	34	28	14
P-Tree11-15	50	2.20	0.00	0.00	0.27	0.35	75	30	22	10
P-Tree16-20	50	2.52	0.00	0.00	0.31	0.40	76	42	28	10
P-Tree1-20	200	83.91	0.00	0.00	0.33	0.47	76	40	29	14
Uni-Sort	150	1.46	0.03	0.94	0.09	0.12	89	12	7	0
Bi-Sort	285	-	-	0.01	0.11	0.15	91	21	15	6
ME-BM	25	2.58	0.00	0.10	0.11	0.14	93	24	8	8
Ind49	49	1.29	0.11	0.36	0.25	0.32	62	18	8	2

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：该表显示了 P-Tree 测试资产和其他资产的绩效统计数据，以供比较，包括 P-Tree1，它由第一个 P-Tree 中的 10 个投资组合组成。P-Tree1-5、P-Tree6-10、P-Tree11-15、P-Tree16-20 和 P-Tree1-20 是一组测试资产，它们分别组合了前 5、6-10、11-15、16-20 和 1-20 个 P-Tree 生成的叶基础投资组合。"Uni-Sort" 有 150 个单变量排序的投资组合，"Bi-Sort" 有 285 个二元排序的投资组合，ME/BE 有 25 个投资组合，"Ind49" 包括 49 个行业投资组合。报告的统计数据包括测试资产的数量、GRS 测试统计数据 and p 值 (Gibbons 等, 1989 年)、PY 测试的 p 值 (Pesaran 和 Yamagata, 2023 年)、平均绝对 α (%)、均方根 α (%)、根据 Fama-French 五因素回归投资组合的平均 R^2 (%)，以及无法解释的显著 alpha 低于 10% 的测试资产的比例，5% 和 1% 显著性水平。这三个面板报告了不同样本周期的重新训练模型。

P-Tree 集群的优势。P-Tree 有两个优势，使这些测试资产变得高效。首先，**P-Tree 是一种以目标为导向的聚类算法，专门用于迭代跨越有效边界。**每片叶子都根据相似的特征值对单个资产收益进行聚类，汇总为投资组合收益的时间序列。其目标是最大化叶节点投资组合所跨越的 MVE 投资组合的集体夏普比率。P-Tree 的经济目标与 CART 和其他现成的 ML 模型大相径庭，因为 CART 和其他现成的 ML 模型只关注统计收益预测，不关注经济效用，根本无法完成增长有效前沿的任务。

其次，**P-Tree 利用了高维特征上复杂的非对称交互和非线性特性。**常用的测试资产是按特征排序的投资组合，如单变量和双变量排序的投资组合，它们只考虑十等分或五等分排序，以及根据研究人员的领域专长临时选择的最多三个特征。相比之下，P-Tree 更为灵活，允许在一个统一的框架内进行具有更多特征和非对称

交互的一般顺序和同步排序。图表 7 所示，P-Trees 可以包含 N15 这样的大叶子，进一步拆分不会导致更高的夏普比率，也可以包含 N8 这样的小叶子。在 MVE 框架下，非对称拆分允许 P-Trees 专注于更细的信息资产集群，而较少关注非信息资产。

4.3 P-Tree 因子和资产定价表现

利用前面的分析结果，我们确定了 P-Tree 测试资产有别于传统测试资产。接下来，我们将在不同的测试资产（包括 PTree 叶基组合）上交叉检验 P-Tree 因子和各种基准因子模型。我们采用横截面 R^2 来评估资产定价性能：

$$\text{Cross-sectional } R^2 = 1 - \frac{\sum_{i=1}^N (\bar{R}_i - \hat{\bar{R}}_i)^2}{\sum_{i=1}^N (\bar{R}_i)^2} \quad (6)$$

其中， \bar{R} 表示平均回报 $\hat{\bar{R}}_i = \beta'_i \lambda$ 是因子模型隐含的资产 i 的平均回报。风险溢价估计采用因子的横截面回归估计，因子载荷是从时间序列回归估计的。横截面 R^2 代表资产平均回报的派系，由模型隐含的预期回报解释。使用各种测试资产（包括 P 树测试资产、常规排序投资组合和行业投资组合）来检验因子模型。在图表 14 中，我们在列中显示因子模型，以解释行中测试资产的预期回报。面板 A 提供了 40 年期样本的资产定价表现，面板 B 和面板 C 是 20 年期子样本的报告。

首先，P-Tree1 因子（来自第一个 P-Tree）解释了 P-Tree1 叶基组合（10 种资产）横截面收益率变化的 98.3%，而基准因子模型的表现要低得多。这一发现并不奇怪，因为第一个 P-Tree 因子是与这些测试资产共同构建的。然而，FF5、Q5 和 RP-PCA 五因子模型的 R^2 远远小于 P-Tree1 因子，表明在解释高度分散的 P-Tree1 测试资产的预期收益时存在较大的定价误差。一个例外是，IPCA 五因子模型在 P-Tree1 测试资产的定价方面表现良好， R^2 为 73.5%。

其次，P-Tree 五因子模型（P-Tree1-5，理解为包含第一到第五个因子的模型）包含了前五个 P-Tree 中的五个 PTree 因子，在资产定价方面表现出色。在对前 10 个 P-Tree（P-Tree1-10）测试资产和 5×5 ME-BM 投资组合进行定价时，它的表现优于基准五因子模型。对于测试资产的其他规格，P-Tree1-5 在所有五因子模型中显示出相当水平的横截面 R^2 ，可进行公平比较。在整个样本和子样本分析中，面板 A 至面板 C 的卓越定价性能保持一致。

第三，这些内生生成的 P-Tree 测试资产更好地扩展了 MVE 边界，提高了测试资产定价模型的标准。常用的测试资产为测试因子模型设定了较低的标准，因此基准因子模型似乎足以为其定价。从图表 14 的面板 A 中我们可以看到，列出的基准五因子模型在基准测试资产上的横截面 R^2 超过 90%，在面板 B 和面板 C 的子样本分析中，这些数字也一直很大。然而，在对 P-Tree 测试资产进行测试时，其 R^2 显著下降。P-Tree 测试资产使用了大量特征之间的非线性和非对称交互作用，而基准测试资产只涉及最多两个特征 1。总之，P-Tree 因子模型能准确地为 P-Tree 测试资产定价，而基准因子则不能，并且在为基准测试资产定价时，P-Tree 因子模型的表现与基准因子模型相似或更好。

总结，P-Tree 测试资产为测试资产定价模型设立了一个高标准，超越了普通分类和行业投资组合。

图表 14 资产定价表现：横截面 R^2

	P-Tree1F	P-Tree5F	P-Tree10F	P-Tree20F	FF5	Q5	RP5	IP5
Panel A: 40 Years (1981–2020)								
P-Tree1	98.3	99.2	–	–	36.8	33.9	51.1	73.5
P-Tree1–5	65.0	82.6	90.2	95.9	54.3	55.9	56.8	49.6
P-Tree1–10	68.5	74.0	87.9	92.1	66.9	70.0	69.1	62.4
P-Tree1–20	72.1	77.9	84.3	88.9	73.8	76.8	77.8	68.6
Uni-Sort	92.7	97.3	98.5	98.7	97.0	97.9	98.0	95.3
Bi-Sort	88.7	97.2	98.0	98.5	96.1	97.5	97.3	92.9
ME-BM	88.4	98.6	98.8	99.4	96.8	97.4	97.0	96.6
Ind49	82.0	92.9	95.3	97.8	96.1	95.9	95.7	91.1
Panel B: 20 Years (1981–2000)								
P-Tree1	99.3	99.3	–	–	54.9	59.7	37.6	79.5
P-Tree1–5	49.2	88.5	92.6	96.0	25.8	31.8	41.1	48.6
P-Tree1–10	47.7	59.7	77.8	82.3	37.7	40.5	47.8	46.8
P-Tree1–20	42.8	54.0	70.2	74.5	42.2	45.6	52.5	50.2
Uni-Sort	80.4	96.5	96.9	97.6	94.6	95.9	97.0	94.5
Bi-Sort	66.4	92.1	93.9	94.9	89.1	91.5	92.7	91.1
ME-BM	66.6	94.0	96.2	99.0	93.4	95.3	96.4	94.5
Ind49	60.9	86.5	87.9	95.0	91.0	89.7	92.1	84.1
Panel C: 20 Year (2001–2020)								
P-Tree1	99.2	99.7	–	–	52.2	43.6	69.7	53.9
P-Tree1–5	71.4	87.6	92.3	94.3	60.0	60.7	64.9	61.3
P-Tree1–10	70.4	78.1	86.4	88.4	66.4	65.7	69.8	65.4
P-Tree1–20	70.5	76.4	79.8	82.8	71.7	69.9	75.4	65.2
Uni-Sort	87.5	96.6	97.6	97.9	97.3	97.2	97.4	95.7
Bi-Sort	93.0	97.2	97.9	98.4	97.9	97.5	98.0	96.9
ME-BM	92.9	95.7	98.6	99.3	97.5	97.5	97.4	96.1
Ind49	75.3	92.6	94.8	97.7	90.2	90.8	92.8	86.0

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：此表显示了按 P-Tree 因子模型和基准因子模型为不同测试资产集定价的横截面 R^2 （%）。因子模型列在列中，测试资产的不同规范列在行中。具体来说，该表在列中包括 1、5、10 和 20 个 P-Tree 因子模型以及 FF5、Q5、RP-PCA 和 IPCA 五因子模型。在行中，测试资产来自前 1、5、10 和 20 个 P-Tree、单变量排序的投资组合、双变量排序的投资组合、ME-BM 5×5 个投资组合和 49 个行业投资组合。

4.4 P-Tree 因子投资

P-Tree 框架将数千种资产聚类为若干叶节点投资组合，降低了投资组合优化的复杂性。这些 P-Tree 测试资产可直接用作因子，或用作 MVE 投资组合的构建模块。此外，重新平衡 P-Tree 投资策略的频率可降至每季度或每年一次，以降低交易成本。

全样本评估。 我们通过将 P-Tree 因子组合成一个 P-Tree 因子切线组合来评估 P-Tree 因子的投资绩效。图表 14 面板 A 报告了 P-Tree 投资策略在 40 年样本中的年化夏普比率和月度阿尔法值。单因子策略的夏普比率分别从 6.37 增加到 9.19、11.21、13.83 和 15.64，因子数量分别增加到 5、10、15 和 20 个。为了评估其额外的投资信息，我们进一步评估了这些 P-Tree 投资策略与基准因子模型 的比较。这些经模型调整的系数均大于 0.80%，且在统计上非常显著。

图表 15 通过 Boosted P-Trees. 进行因子投资

	SR	α_{CAPM}	α_{FF5}	α_{Q5}	α_{RP5}	α_{IP5}
Panel A: 40 Years (1981-2020)						
P-Tree1	6.37	1.39	1.37	1.36	1.28	1.12
P-Tree1-5	9.19	0.97	0.95	0.93	0.87	0.82
P-Tree1-10	11.21	1.01	1.00	0.98	0.93	0.89
P-Tree1-15	13.83	0.95	0.94	0.93	0.90	0.87
P-Tree1-20	15.64	0.97	0.96	0.95	0.93	0.90
Panel B1: 20 Years In-Sample (1981-2000)						
P-Tree1	7.13	1.86	1.78	1.72	1.62	1.59
P-Tree1-5	12.74	1.54	1.51	1.48	1.37	1.41
P-Tree1-10	19.22	1.51	1.49	1.49	1.43	1.43
P-Tree1-15	28.43	1.42	1.41	1.40	1.37	1.39
P-Tree1-20	38.01	1.36	1.35	1.34	1.32	1.34
Panel B2: 20 Years Out-of-Sample (2001-2020)						
P-Tree1	3.23	1.35	1.31	1.23	1.04	0.93
P-Tree1-5	3.41	1.02	1.00	0.95	0.77	0.62
P-Tree1-10	3.21	0.95	0.94	0.89	0.74	0.56
P-Tree1-15	3.12	0.89	0.89	0.83	0.69	0.48
P-Tree1-20	3.13	0.85	0.84	0.78	0.66	0.49
Panel C1: 20 Years In-Sample (2001-2020)						
P-Tree1	5.83	1.51	1.47	1.50	1.52	1.69
P-Tree1-5	9.32	1.30	1.29	1.28	1.30	1.31
P-Tree1-10	14.35	1.12	1.11	1.11	1.11	1.09
P-Tree1-15	20.64	1.08	1.07	1.08	1.10	1.05
P-Tree1-20	26.57	1.09	1.08	1.08	1.10	1.11
Panel C2: 20 Years Out-of-Sample (1981-2000)						
P-Tree1	4.35	1.50	1.42	1.35	1.60	1.58
P-Tree1-5	3.87	1.18	1.05	0.96	1.23	1.24
P-Tree1-10	4.29	1.02	0.93	0.85	1.14	1.10
P-Tree1-15	4.03	0.96	0.86	0.80	1.07	1.02
P-Tree1-20	3.88	0.96	0.87	0.81	1.08	1.03

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：该表显示了 P-Tree 产生的初始因子和 Boosting P-Trees 产生的多因子 MVE 投资组合的投资表现。报告的统计数据包括关于 CAPM、FF5、Q5、RP-PCA 和 IPCA 五因素模型的年化夏普比率和 alphas。图 A 显示了从 1981 年到 2020 年的 40 年样本。图 B1 和 B2 分别显示了 1981 年至 2000 年和 2001 年至 2020 年的样本内和样本外结果。图 C1 和 C2 分别显示了 2001 年至 2020 年和 1981 年至 2000 年的样品内和样品外结果。我们发现所有 α 在 1% 的置信水平下都显著。

现有研究发现，P-Tree 投资策略具有极具竞争力的夏普比率。Ahn 等（2009）基于相关性的聚类产生的基础投资组合在 ME 和 BM 上的表现低于双变量排序投资组合。此外，Daniel 等（2020）通过构建相应的特征效率投资组合，将特征排序投资组合的夏普比率从 1.17 提高到 2.13，几乎翻了一番。不过，从 1981 年到 2020 年的 40 年间，单一 P-Tree 的夏普比率为 6.37，与 Ahn 等（2009）和 Daniel 等（2020）相比有显著提高，尽管实证样本不同。

样本外评估。 图表 15 面板 A 中的结果基于全样本分析，这可能会引起对样本内过拟合的担忧。因此，我们提供了两个关于 样本外 投资策略的练习。我们效仿 Kozak 等（2020）的做法，将训练样本和测试样本对半分。当过去样本和未来样本相互预测时，只需证明 样本外 性能即可。当未来样本训练的模型预测过去样本时，这种方法进一步减轻了对前瞻性偏差的担忧。一方面，面板 B1 和 B2 报告了过去预测未来的结果。前二十年样本内的夏普比率甚至高于面板 A 中的四十年样本内的夏普比率，样本外 夏普比率超过 3。尽管业绩有所下降，但经基准模型调整后，许多 样本外 系数仍接近 1%，且非常显著。

另一方面，对于未来预测过去的结果，样本外 夏普比率大于 3.87，模型调整后的阿尔法值超过 0.80%。这些研究结果表明，P-Tree 测试资产和因子在跨越 MVE 边界时具有很强的 样本外 证据。

本文报告的样本外投资业绩结果与近期研究结果类似。Cong 等（2021）的深度强化学习策略从 1990 年到 2021 年的夏普比率一直高于 2，即使在剔除小型股票之后也是如此。Bryzgalova 等（2023）基于规模、营业利润率和投资排序的投资组合，在 2004 年至 2016 年期间，正则化投资组合的年化夏普比率达到 2.39。Feng 等（2024）的深度神经网络方法报告了 2002 年至 2021 年期间 2.95 至 3.00 的夏普比率。作为测试资产生成、资产定价和投资管理的全方位框架，P-Tree 策略因其透明度高、易于计算和灵活性而脱颖而出。

5 模型扩展和讨论

我们沿着几个方向扩展了 P-Tree 框架和应用，包括可解释性、模型复杂性、特征效用和宏观经济制度。

5.1 随机 P-Forest 与 P-Tree 的可解释性

随机森林（Breiman, 2001 年）在引导训练样本上生长多棵“装饰相关”的树，这些样本是从完整训练集中随机抽取的观测样本和变量子集。这就是所谓的“bagging”，它有助于减轻模型的过度拟合，量化模型估计的不确定性。我们可以采用这种集合方案，在引导训练样本上生长多个“装饰相关”的 P-Tree，形成随机 P-Tree。因此，随机 P-Forest 有两个直接应用：评估特征重要性和使用大量“装饰相关”的 P-Trees 恢复稳健的 SDF。

特征重要性，**尽管存在冗余或无用的特征，非线性和交互式 P-Tree 仍能恢复真实的特征集**，从而改进有效前沿估计的风险代理。这与线性因子模型形成了鲜明对比，后者可能会忽略这些特征。传统的顺序或独立资产排序方案通常最多关注三个特征。枚举具有多个特征的所有可能排序情况是 NP-hard 的，但 P-Tree 的迭代增长算法有效地克服了这一问题，使计算变得可行。

构建随机 P-Forest 的集合步骤如下：(i) 对时间序列范围内的数据进行引导替换，并保留所选时间段内面板数据的完整截面，以利用资产收益之间的低序列相关性。(ii) 随机选择 20 个特征，并在每个自举样本上独立生长 P-Trees。这两个步骤重复 1000 次，以创建由 1000 个 P-Trees 组成的森林。

我们研究在随机 P-Tree 中选择特征进行拆分的频率。选择频率越高的特征对均值-方差效率最大化越重要。在每个 bootstrap 样本中，随机抽取 20 个特征来生长 P-Tree，其中只有部分特征被选为分割变量。我们计算第 l 个特定 z_l 特征在前 J 次分裂中的使用次数和出现的总次数。我们对特征重要性的度量定义如下：

$$\text{Select. Prob.}(z_l) = \frac{\#(z_l \text{ is selected at first } J \text{ splits})}{\#(z_l \text{ appears in all bootstrap subsamples)}}, \quad (7)$$

图表 16 总结了 $J=1, 2, 3$ 时特征的选择概率。选择概率随着 J 从 1 到 3 的增加而增加，因此我们必须在行内进行比较。在每一行中，我们列出了被选中的前五个特征。

盈利惊喜（SUE）是最重要的基本要素之一，在 bootstrap 样本中，它有 50% 的概率成为第一个分割特征，有 66% 的概率成为前两个分割特征之一。虽然 FF5 和 Q5 等线性因子模型不包含 SUE，但 SUE 在非线性互动模型中对均值-方差效率最大化最有价值。其他经常被选中的特征包括收益波动率（SVAR）、美元交易量（DOLVOL）、特异波动率（RVAR_CAPM）、利润率变化（CHPM）和行业

调整后市净率 (BM_IA)。图表 16 中的特征涵盖了四个主要类别--动量、摩擦、盈利能力、价值--的顶部拆分。请注意，未被选中的特征并不一定意味着它们毫无用处，因为本表只显示了前三名的拆分。

图表 16 通过选择概率确定特征重要性

	1	2	3	4	5
Top1	SUE	SVAR	CHPM	RVAR_CAPM	BASPREAD
	0.51	0.34	0.25	0.23	0.21
Top2	SUE	SVAR	DOLVOL	CHPM	BM_IA
	0.66	0.38	0.34	0.31	0.31
Top3	SUE	DOLVOL	BM_IA	SVAR	ME_IA
	0.71	0.53	0.48	0.41	0.41

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：此表报告了从 1000 棵树的随机 P-Tree 中选择的最频繁特征。"Top 1" 行仅计算 1000 棵树的第一个拆分。"Top 2" 或 "Top 3" 行只计算前两个或三个拆分。报告的数字是从 1000 个融合中选出的这些主要特征的选择频率。特性描述如图表 6 所示。

可解释性 VS 过度拟合。 随机 P-Forest 可能参数过大，在可视化和解释方面具有挑战性，但它有几个优点。(i) 随机 P-Forest 解决了模型的不确定性问题，可以通过多个自举样本来评估特征的重要性。换句话说，单一 P-Tree 很容易解释，因为它直观地描述了非线性和非对称交互作用的特征，但如果我们不能证明模型没有过拟合，那么模型的功效可能会受到限制。相比之下，随机 P-Forest 利用多个随机引导样本，有助于降低过拟合的风险。(ii) 随机 P-Forest 为创建多个测试资产提供了一种替代 boosting 的方法。与按顺序生成资产的 boosting 不同，随机 P-Forest 是跨 P-Tree 独立创建资产。

除了图表 16 中用于评估模型过拟合的样本外评估外，我们还比较了随机 P-Forest 和单一 P-Tree 的关键特征。随机 P-Forest 也显示了单一 P-Tree 所选的重要特征，这表明单一 P-Tree 并没有过度拟合。这种一致的结果可能归功于 P-Tree 的全局分割标准，它与 ML 组合方法类似，可以防止过度拟合，同时提供更高的可解释性和对非线性和非对称交互作用的透明洞察力。图表 7 所示，从随机 P-Forest 中获得的重要特征，如盈利惊喜 (SUE)、美元交易量 (DOLVOL) 和行业调整后市净率 (BM_IA)，也被选入 P-Tree 40 年的顶级分割中。此外，图表 16 中列出的重要特征也被选入子样本分析的 P-Tree 图中显示的顶部分割中。

5.2 随机 P-Forest 和复杂性

我们的随机 P-Forest 还与金融领域日益流行的大型模型有关。在最近的这些文献中，研究人员通常会质疑模型的性能是否会随着模型规模的增加而提高，这与最近关于“良性过拟合”和“高维插值”（如 Belkin 等，2019；Hastie 等，2022）以及金融应用中的“复杂性的美德”（如 Kelly 等，2022，2024；Didisheim 等，2024）的统计文献不谋而合。这些文献表明，大型正则化模型（相对于观测值而言参数数量较多）也能在样本外表现良好，其表现甚至会随着参数数量的增加而改善。这与统计学和金融学中使用简约模型的传统智慧背道而驰。

随机 P-Forest 在随机引导样本上独立构建了大量不相关的 P-Trees，使其可以扩展为一个具有统计和经济正则化的大型模型。这些数百个 P-Trees 和数千个

叶节点投资组合与 Didisheim 等 (2024) 中随机生成的投资组合类似。我们可以使用森林中的所有叶基组合估算 SDF (切线组合), 并报告其样本外性能, 以展示大型模型的性能。

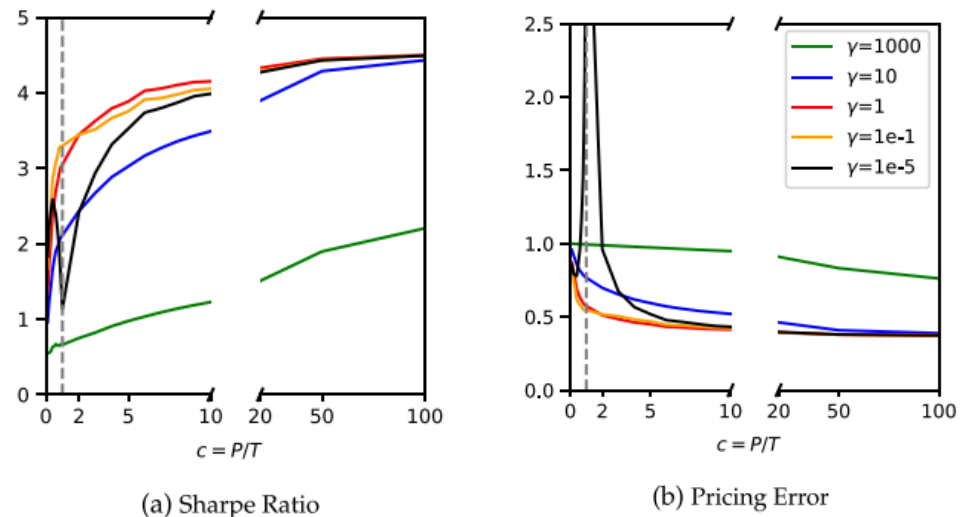
我们通过以下方法创建随机 P-森林 SDF: (i) 为一棵 P-Tree 随机选择 L 特性, 用分割标准 (4) 生长 P-Tree。 (ii) 重复(i)程序 B 次, 以并行计算的方式形成 B P-Tree 组成的森林。 (iii) 根据样本数据, 对随机 P-森林中所有叶基组合的 SDF 权重进行估计 (8)。我们采用 Didisheim 等 (2024 年) 的脊 (无脊) SDF 估计器来估计叶基组合的 SDF 权重 \hat{w} :

$$\hat{w}(\gamma) = \operatorname{argmin}_w E \left[(1 - w' \mathbf{R}_t)^2 + \gamma \|w\|^2 \right] \quad (8)$$

其中, γ 是收缩参数。我们还研究了大型模型的性能, 并将 $c=P/T$ 定义为参数化程度或复杂性, 其中 P 是叶基组合的数量, $T = 240$ 是固定估计滚动窗口。我们检验了广泛参数化和收缩估计对不同 c 和 γ 值的随机 PForest SDF 的影响。除了样本外夏普比率, 我们还报告了样本外定价误差 $E_{Oos}[(1 - \hat{w}' \mathbf{R}_t)^2]$ 。定价误差是样本外 HJ 距离 (Hansen 和 Jagannathan, 1997 年), 当 $P > T_{Oos}$, 且两者都足够大时, 其表达式如上。

图表 17 显示了 2001 年至 2020 年随机 P-Forest SDF 的 OOS 性能, 其中 $L = 10$ 个随机特征。随着 c 的增加, 我们发现: (i) 低缩水情况下的夏普比率呈现双倍上升, 高缩水情况下的夏普比率呈现永久上升; (ii) 高缩水情况下的定价误差减小, 低缩水情况下的 $c = 1$ 左右出现峰值。这些模式与 Didisheim 等 (2024) 的研究类似, 表明随机性和大模型规模对随机 P-Forest SDF 的样本外性能有好处。当 $c = 10$ 时, 夏普比率约为 4.0, 定价误差低于 0.44。当 $c = 100$ 时, 夏普比率约为 4.5, 定价误差低于 0.38。

图表 17 随机 P-Forest SDF 的 OOS 性能



资料来源: 《Growing the efficient frontier on panel trees》, 华安证券研究所

注: 该图报告了已实现的样本外 SDF 投资组合的夏普比率和定价误差 (HJD)。随机 P 森林中的每个 P-Tree 都提供 10 个特征, 从 61 个特征中随机抽取而不替换, 并生成 10 个叶基础投资组合。这些 P-Tree 是独立训练的。P-Tree 使用面向目标的标准进行拆分, 以最大化叶节点投资组合的集体夏普比率。叶节点投资组合的总数表示为 P。横轴显示模型复杂度 $c = P/T$, c 范围为 0.1 到 100, $T = 240$ 个月。我们报告了 5 个收缩参数 γ [10⁻⁵, 10⁻¹, 1, 10,

1000] 的规格。

目标导向 P-Tree 与随机分割 P-Tree 的比较。随机 P-Tree SDF 涉及随机性，因为每个 P-Tree 都会随机抽取一个特征子集，但拆分决策并不是随机做出的。P-Tree 中以目标为导向的搜索有助于叶节点投资组合有效地最大化集体夏普比率。然而，Didisheim 等（2024）的研究表明，如果投资组合的数量足够多，随机投资组合也能生成具有出色样本外性能的 SDF。我们通过提出另一种 SDF 规格进行比较，说明了目标导向搜索在参数过多情况下实现均值-方差效率最大化的重要性。具体来说，随机拆分 SDF 使用随机拆分标准来创建叶基组合，没有特定的拆分标准，每个决策的特征和拆分值都是随机选择的，从而生成 10 个叶基组合。

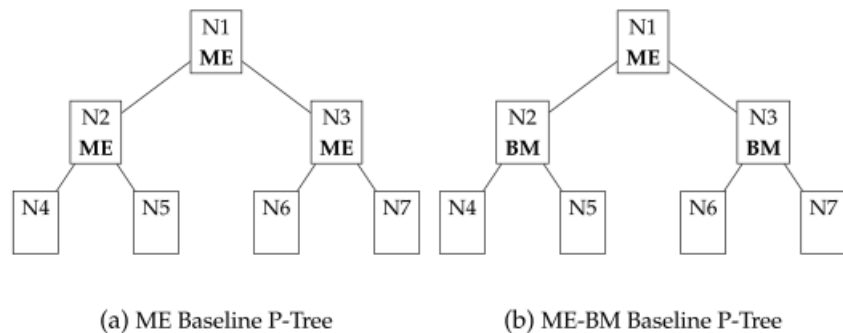
随机 P-Forest SDF 的模式与随机拆分 SDF 类似，但随机 P-Forest 在参数化和计算费用方面明显更有效。具体来说，要达到与随机拆分 SDF 相同的夏普比率和定价误差水平，所需的复杂度更低。因此，目标导向搜索提高了具有统计和经济正则化的大型模型的效率，符合金融研究中将经济目标与 ML 模型相结合的理念。

最后，从测试资产构建的角度来看，目标导向 P-Tree 为资产定价模型的估计和评估提供了实际有用的测试资产。测试资产的数量通常少于 100 个，因为当数量较多时，资产定价测试的统计能力较低（Fan 等，2015）。

5.3 用 P-Tree 评估股票特征

作为另一种定价内核，P-Tree 可以直接评估新特征相对于基准的增量（和非线性）贡献，从而对文献进行补充。研究人员在实证研究中根据特定特征进行分类，开发出新的因子，然后通过跨度回归控制基准模型，检验其独立信息。然而，实证研究很少考察相关特征是否提供了新信息，可能会造成信息损失。为了解决这个问题，P-Tree 提供了一个新的视角，即在给定一组基准特征（如 ME 和 BM）的情况下，评估特征在生成测试资产和潜因模型中的有用性。

图表 18 ME 基线和 ME-BM 基线 P-Tree 的演示



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：ME 基线和 ME-BM 基线 P-Tree 演示。

在图表 17 中，我们展示了两种基准 P 树：一种是三层 ME 基准 P 树，所有分层都基于 ME；另一种是三层 ME-BM 基准 PT 树，第一层基于 ME，第二层基于 BM。为了评估根据 ME 估算特征有效前沿的增量效用，我们扩展了 ME 基线 P-Tree，对感兴趣的特征或 ME 进行了额外的拆分。对于浅层 P-Tree 来说，我们可能会认为拆分会提高 MVE 投资组合的夏普比率。因此，我们只通过比较在基线 P-Tree 中添加某一特征与对 ME 进行拆分时的夏普比率改善情况，来评估某一特

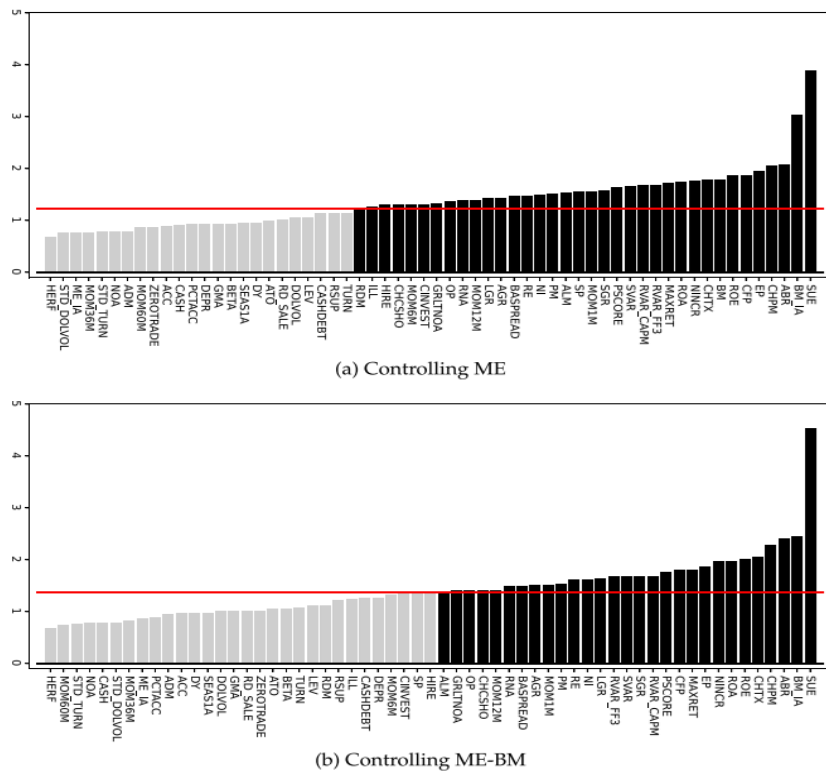
征是否提供了新的信息。这种方法有助于我们衡量相关特征的增量影响。

在图表 19 中，子图 (a) 显示了每个特征在 ME 基线 P-Tree 上的夏普比率增量。我们用黑色标出表现优于 ME 的特征，用灰色标出表现劣于 ME 的特征。我们还按照夏普比率对特征进行了分类。即使对于浅层 P-Tree 来说，24 个特征也不能提供与 ME 基线 P-Tree 相比的增量信息。此外，我们还确定了除规模以外的其他类别，包括动量 (SUE、ABR、NINCR)、价值 (BM_IA、BM、CFP、EP)、盈利能力 (CHPM、ROA、ROE) 和摩擦/波动性 (MAXRET、SVAR)，这些类别通过对 ME 基线 P-Tree 进行简单的进一步拆分，大大提高了投资绩效。

总之，我们通过控制 ME 和/或 BM 对投资目标的边际贡献来评估基准 P-Tree 的每个特征。如果将更深的树作为基准，我们可以预期这些边际贡献会更小。当基准 P-Tree 或控制特征发生变化时，特征的相对重要性也会发生变化，这突出了 P-Tree 评估的适应性。在评估特征时，这种评估可以补充传统的因子跨度回归。

同样，我们使用图表 19 子图 (b) 中所示的 ME-BM 基准 P-Tree 评估每个特征。该 P-Tree 有三层，基于 ME 和 BM。有了更强的基准，一些特征，如 RDM、ILL、HIRE、MOM6M、CINVEST、SP，在这棵升级的 ME-BM 基线 P-Tree 上并不能提供额外的信息。与 ME-BM 基线 P-Tree 相比，我们发现 30 个特征没有为估计有效前沿提供增量信息。与 ME 基线 P-Tree 的结果一致，动量、价值和盈利能力特征在非线性互动模型中是有用的。

图表 19 使用 P-Tree 评估股票特征



资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：该图显示了基于 ME 或 ME-BM 基线 P-Tree 的特征拆分后 P-Tree 基础投资组合的夏普比率。子图 (a) 和 (b) 分别显示了 ME 和 ME-BM 基线 P-Tree 的结果。我们按特征的夏普比率升序对特征进行排序。红线显示了 P-Tree 基础投资组合在控制特征上进一步拆分后的夏普比率。红线上方的条形表示针对基准特征的增量信息，颜色为黑色。否则，灰色条表示特征不提供增量信息。

5.4 宏观经济环境下的 P-Tree

越来越多的经验证据表明，在不同的宏观经济条件下，因子风险溢价和载荷会有很大的不同。鉴于横截面驱动力可能会随着时间的推移而发生变化，研究 P-Tree 在不同状态下的表现以及选择哪些特征进行均值-方差分散是非常有价值的。本小节介绍了在各种宏观制度下实施 P-Trees 的子样本分析，并展示了其稳健的表现。

制度下的特征评估。我们通过分析图表 5 中的十个宏观变量，对不同制度下的子样本进行稳健性检验。不同制度下 P-Tree 模型的实证结果见图表 20。我们使用 50% 的 10 年滚动窗口百分位数将样本划分为上层和下层制度，并为每种制度拟合一个单独的模型。这样，我们就可以评估 P-Tree 模型在宏观经济利好和利空环境下的有效性。

首先，在不同的宏观经济条件下，有用的特征可能会有所不同。P-Tree 表现良好，可根据宏观经济变化进行调整，并为前三个拆分选择不同的特征。例如，当期限利差较低或市场波动性较高时，SUE 和 DOVOL 对公司基本面和交易量的作用对横截面均值-方差分散化更为重要。然而，代表流动性和波动性摩擦的 BASPREAD 和 RVAR_CAPM 等特征通常被用于高股票净发行量、低违约收益率、低市场流动性或低通胀的情景。

图表 20 政策变动下的 P-Tree 性能

	Regime	AVG	SR	GRS	p-value	$ \alpha $	$\sqrt{\alpha^2}$	\bar{R}^2	Top three characteristics		
DFY	Top	1.78	6.14	71.84	0.00	1.27	1.57	52.69	SUE	DOLVOL	CFP
	Btm	1.21	4.56	40.01	0.00	1.08	1.44	60.05	ABR	RVAR_CAPM	STD_DOLVOL
	Com	1.49	5.16								
DY	Top	1.45	5.84	43.17	0.00	1.03	1.30	73.16	CHPM	BM_IA	ME
	Btm	1.30	6.21	101.56	0.00	1.05	1.21	51.45	ABR	ZEROTRADE	DOLVOL
	Com	1.35	6.03								
EP	Top	0.99	5.51	34.85	0.00	0.94	1.28	78.21	RVAR_CAPM	ME_IA	ME
	Btm	1.34	6.04	98.62	0.00	0.94	1.18	56.89	SUE	DOLVOL	BM_IA
	Com	1.24	5.76								
ILL	Top	1.10	5.90	89.15	0.00	0.81	1.04	62.31	RVAR_CAPM	SUE	STD_DOLVOL
	Btm	1.50	7.44	72.13	0.00	1.03	1.30	59.38	SUE	DOLVOL	SP
	Com	1.24	6.21								
INFL	Top	1.16	4.83	30.10	0.00	0.99	1.29	75.20	CFP	ATO	ME
	Btm	1.29	6.61	107.40	0.00	0.93	1.26	57.67	SUE	DOLVOL	CFP
	Com	1.25	5.89								
LEV	Top	1.33	5.10	38.77	0.00	0.89	1.14	65.03	CFP	ALM	STD_DOLVOL
	Btm	1.39	5.36	66.94	0.00	1.23	1.46	52.23	NI	MOM12M	BM_IA
	Com	1.37	5.26								
NI	Top	1.37	5.05	45.70	0.00	1.67	2.16	53.03	RVAR_CAPM	ME_IA	RD_SALE
	Btm	1.51	6.71	93.02	0.00	1.18	1.38	61.79	ROA	DOLVOL	SUE
	Com	1.45	5.81								
SVAR	Top	1.24	6.22	72.03	0.00	0.94	1.24	62.66	SUE	DOLVOL	CFP
	Btm	1.23	4.78	43.92	0.00	1.18	1.38	59.41	ROA	ZEROTRADE	ABR
	Com	1.23	5.36								
TBL	Top	1.09	7.15	59.53	0.00	0.75	0.91	61.98	CHCSHO	SUE	DOLVOL
	Btm	1.30	7.05	134.30	0.00	0.93	1.10	63.44	SUE	DOLVOL	CFP
	Com	1.24	6.97								
TMS	Top	1.17	6.01	67.58	0.00	0.99	1.28	63.45	EP	STD_DOLVOL	SUE
	Btm	1.23	5.82	70.00	0.00	0.86	1.17	52.52	SUE	DOLVOL	BM_IA
	Com	1.20	5.91								

资料来源：《Growing the efficient frontier on panel trees》，华安证券研究所

注：上表显示了 P-Tree 模型在各种宏观经济状态下的性能。每个宏变量的时间序列使用 50% 的 10 年滚动窗口百分位数分为 "Top" 和 "Btm" 状态。然后为每个制度构建一个简单的 P-Tree，图表 13 中报告了类似的性能统计数据，以及前三个分裂特征。标记为 "Com" 的行报告了 "Top" 和 "Btm" 方案的组合 P-Tree 因子的平均回报和夏普比率。

其次，P-Tree 测试资产在资产定价方面始终提供了有参考价值的结果：针对

法马-法式五因子的 GRS 测试在每个制度下都被拒绝，P-Tree 测试资产所跨的 MVE 投资组合的夏普比率均大于 7。有趣的是，GRS 统计（加权定价误差）在涉及低通胀率或低 t-bill 利率的情景下明显更大，这意味着捕捉横截面信号的常用因子的矩最弱。图表 13 面板 A 中的平均值大于“PTree1”的平均值，这意味着极端宏观制度下的测试资产更难通过通用因子进行定价。最后，我们将两种制度下的 P-Tree 因子结合起来，发现图表 14 面板 A 中的“‘Com’”（组合）因子比第一个 P-Tree 因子的夏普比率更大，因为后者执行的是无条件模型。

6 结论

利用单个资产回报估算均值-方差有效前沿以及为资产定价模型评估创建多样化测试资产是长期存在的实证挑战（例如，马科维茨，1952；Lewellen 等，2010 年；丹尼尔等，2020）。我们的论文引入了一类新的 P-Tree，它能有效地应对这些实证挑战，并通过概括多特征证券排序和分割横截面来分析单个资产收益的（非平衡）面板。在均值-方差效率的全局标准下，P-Tree 利用包含资产收益联合分布丰富信息的高维特征，生成特征管理投资组合并恢复随机贴现因子。更广泛地说，PTrees 将基于树的模型和其他 ML 模型扩展到纯粹预测之外，通过经济导向目标有效地分析面板数据，同时保持可解释性，并处理低信噪比环境下的非对称和非线性相互作用。

我们对美国股票的实证研究表明，与那些常用的测试资产（如双变量或单变量分类投资组合）相比，P-Tree 测试资产大大推进了有效前沿，并且与基准模型（如法玛-法式因子）相比，表现出显著的未解释优势，凸显了测试资产的重要性。这些结论在最近二十年的样本外评估中保持一致和稳健。其次，P-Tree 切线组合是作为交易因子模型构建的，在因子投资和横截面定价性能方面优于流行的可观察和 ML 潜在因子模型。第三，我们展示了基于稀疏树模型在经济解释中的多功能性与实用性。例如，我们发现 SUE、DOLVOL 和 BM_IA 是相互作用解释资产收益截面的关键特征。最后，P-Tree 利用稀疏性捕捉到了面板股票回报的复杂性，实现了接近超参数化大型模型的卓越样本外夏普比率。除了资产定价之外，我们的框架还为基于深度学习的人工智能在大型建模空间中进行目标导向搜索提供了一种可解释且计算高效的替代方法。

文献来源：

核心内容摘选自 Lin William Cong、Guanhao Feng、Jingyu He 和 Xin He 于 2025 年 02 月 04 日在 Journal of Financial Economics 上的文章《Growing the Efficient Frontier on Panel Trees》

风险提示：

文献结论基于历史数据与海外文献进行总结；不构成任何投资建议。

重要声明

分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的执业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，本报告所采用的数据和信息均来自市场公开信息，本人对这些信息的准确性或完整性不做任何保证，也不保证所包含的信息和建议不会发生任何变更。报告中的信息和意见仅供参考。本人过去不曾与、现在不与、未来也将不会因本报告中的具体推荐意见或观点而直接或间接接收任何形式的补偿，分析结论不受任何第三方的授意或影响，特此声明。

免责声明

华安证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。本报告中的信息均来源于合规渠道，华安证券研究所力求准确、可靠，但对这些信息的准确性及完整性均不做任何保证。在任何情况下，本报告中的信息或表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。华安证券及其所属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投资银行服务或其他服务。

本报告仅向特定客户传送，未经华安证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。如欲引用或转载本文内容，务必联络华安证券研究所并获得许可，并需注明出处为华安证券研究所，且不得对本文进行有悖原意的引用和删改。如未经本公司授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司并保留追究其法律责任的权利。

投资评级说明

以本报告发布之日起 6 个月内，证券（或行业指数）相对于同期沪深 300 指数的涨跌幅为标准，定义如下：

行业评级体系

- 增持—未来 6 个月的投资收益率领先沪深 300 指数 5%以上；
- 中性—未来 6 个月的投资收益率与沪深 300 指数的变动幅度相差-5%至 5%；
- 减持—未来 6 个月的投资收益率落后沪深 300 指数 5%以上；

公司评级体系

- 买入—未来 6-12 个月的投资收益率领先市场基准指数 15%以上；
- 增持—未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%；
- 中性—未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；
- 减持—未来 6-12 个月的投资收益率落后市场基准指数 5%至；
- 卖出—未来 6-12 个月的投资收益率落后市场基准指数 15%以上；
- 无评级—因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。市场基准指数为沪深 300 指数。