Cost-Aware Generative Portfolios *

Lin William Cong Jingyu He Xin He Feng Ye

November 3, 2026

Abstract

We develop a *cost-aware Panel Tree* that generates tradable characteristic-managed basis-portfolios by incorporating transaction costs into split selection. Theoretically, we prove that the model recovers the mean–variance investment opportunity set of underlying individual assets asymptotically. Empirically, using 153 firm-level signals, we evaluate the cost-aware model in the U.S. equity market over 1984–2023 and find that it reduces turnover and improves net performance under mean–variance objectives. Most of the net-performance improvement arises from basis-portfolio generation rather than subsequent allocation. Cost-aware Panel Trees also selectively reallocate split capacity toward accounting-fundamental signals while retaining price-volume signals that survive transaction costs.

JEL Classification: G11, C61, C63.

Keywords: Asset Pricing, Generative Model, Panel Tree, Portfolio Selection, Transaction Costs.

*We are incredibly grateful to Zhou Lu (discussant), Dake Zhang (discussant) and Jun Zhang (discussant) for detailed comments and suggestions. We also thank seminar and conference participants at China Finance Annual Meeting 2026 Nankai University, Renmin University of China, Symposium on Artificial Intelligence and Financial Econometrics 2026 Xiamen University, and the 6th Great China Area Finance Conference 2026, for constructive discussions and feedback. Cong (Correspondence; E-mail: will.cong@ntu.edu.sg) is at Nanyang Business School, College of Computing and Data Science, Nanyang Technological University. He J. (E-mail: jingyuhe@cityu.edu.hk) is at City University of Hong Kong. He X. (E-mail: xin.he@ustc.edu.cn) is at University of Science and Technology of China. Ye (E-mail: fengye7-c@my.cityu.edu.hk) is at City University of Hong Kong. The authors acknowledge financial support from the Fundamental Research Funds for the Central Universities of USTC [WK2040250133], the National Natural Science Foundation of China [72595865, 7260030112] and Inquire Europe. All authors contributed equally. The author order is alphabetical.

1 Introduction

Since Markowitz (1952), mean–variance portfolio choice has been the central research question of finance and investment. Asset allocation among basis-portfolios is standard in the finance literature. Examples include the market, ME-sorted, and BE/ME-sorted portfolios (Fama and French, 1993), the correlation-clustered basis assets (Ahn et al., 2009), the country index and industry portfolios (DeMiguel et al., 2009), the characteristic efficient portfolios (Daniel et al., 2020), and the Panel Tree leaf basis-portfolios (Cong et al., 2025). Recent evidence shows that the gross investment utilities of factor- and characteristic-based strategies can be substantially reduced by liquidity, limits to arbitrage, price impact, and transaction costs; see Novy-Marx and Velikov (2016); DeMiguel et al. (2020); Li et al. (2024); Jensen et al. (2026).

This paper proposes the *cost-aware Panel Tree* as a generative modeling tool for portfolio choice. Unlike the cardinality constrained portfolios (Gao and Li, 2013; Xu et al., 2024) imposing hard constraints on the number of assets, our model partitions the cross-section of thousands of individual assets into a small number of non-overlapping leaf basis-portfolios, and then allocates capital among them. We aim to generate the optimal basis-portfolios for asset allocation using asset characteristics while accounting for realistic transaction costs. Further, we ask an empirical question: do these optimal characteristic-managed basis-portfolios remain attractive after accounting for the individual asset-level transaction costs?

An investor faces a cost-aware basis-portfolio generation problem and a cost-aware asset allocation problem. This paper argues that these two steps constitute a joint problem. Basis-portfolios sorted on fast-moving characteristics could enhance the gross utility; however, the benefit may be offset by substantial individual asset-level rebalancing.

In contrast, slow-moving characteristics may contribute little to the gross utility but incur much smaller transaction costs. Therefore, transaction costs affect the joint problem of basis-portfolio generation and asset allocation. The optimal basis-portfolios balance gross investment utility against transaction costs. We also ask which contributes more to investment utility: cost-aware basis-portfolio generation or cost-aware asset allocation.

In the literature, investors often convert a high-dimensional individual asset-level portfolio choice problem to a low-dimensional asset allocation problem over basis-portfolios. Characteristic-managed portfolios are natural building blocks for asset allocation, as Cochrane (2011) states that expected returns, variances, and covariances are stable functions of asset characteristics. Theorem 1 demonstrates that the optimal investment utility on the basis-portfolios equals that of the underlying individual assets with mild assumptions. The cost-aware Panel Tree is a myopic algorithm that recursively splits the cross-section of securities on lagged characteristics and generates value-weighted leaf basis-portfolios. Theorem 2 proves that the proposed model recovers the optimal basis-portfolios. When growing the Panel Tree, we evaluate each candidate split using the cost-aware objective, charging individual asset-level transaction costs. It includes the *cost-ignorant Panel Tree* as a special case in which transaction costs are set to zero. Theorem 3 establishes the value of integrating transaction costs directly into basis-portfolio generation—rather than merely reoptimizing downstream allocations—and identifies the conditions for a strictly positive improvement in population utility.

The cost-aware Panel Tree has three implications for portfolio choice. First, it endogenously constructs basis-portfolios in high-dimensional and unbalanced panel data, as securities are assigned to leaf basis-portfolios using lagged characteristics. Second, it is defined by a series of split rules, containing characteristic names and values, with clear

visualization in Figures 1 and 2. Third, it internalizes transaction costs at the basis-design stage, so transaction costs affect how basis-portfolios are generated, rather than affecting only asset allocation weights or ex-post net-of-cost utility reporting.

Empirical summary. We apply the cost-aware Panel Tree to the U.S. common stocks from 1974 to 2023 using 153 rank-standardized firm characteristics shared by Jensen et al. (2023), and compare it with the cost-ignorant counterpart. For each out-of-sample year from 1984 to 2023, we estimate the model with data from the preceding 120 months and apply it over the following year. Leaf membership and implied stock holdings are updated monthly as lagged characteristics evolve. The cost-aware and cost-ignorant Panel Trees use the same data, rolling-window protocol, and transaction-cost schedule; they differ only in whether transaction costs enter the objective. We consider an investor with mean–variance utility and another one with a maximum-Sharpe ratio objective. There are three empirical findings.

First, incorporating transaction costs in the objective function improves the net-of-cost investment utility out of sample. Under the maximum-Sharpe ratio objective, the average monthly transaction costs fall from 0.56% to 0.34%, turnover falls from 85.87% to 53.16%, and the net Sharpe ratio rises from 0.17 to 0.98 relative to an otherwise identical cost-ignorant Panel Tree. Under mean–variance utility with risk aversion $\delta = 2$, transaction costs fall from 2.31% to 1.38%, the monthly turnover falls from 351.92% to 220.86%, and the net objective rises from 0.62 to 1.40. Under both objectives, the higher net performance comes from higher net mean returns rather than lower volatility. The generation–allocation decomposition shows that the improvement is primarily the outcome of basis-portfolio construction. Pairing the basis-portfolios generated by the cost-aware Panel Tree

with the cost-ignorant allocation rule recovers 91.77% of the total maximum-SR improvement and essentially all of the mean–variance utility improvement, whereas applying the cost-aware allocation rule to the cost-ignorant basis-portfolios recovers only 32.26% and 26.13%, respectively.

Second, the cost-aware Panel Tree focuses more on the slow-moving accounting-fundamental signals and less on the fast-moving price-volume signals. Across 40 rolling estimation windows, the split score (variable importance) of accounting-fundamental signals rises from 50.89% to 61.72%, while that of the price-volume signals declines from 49.11% to 38.38%. For instance, Quality and Low Risk categories have the largest increases in split scores in the cost-aware Panel Tree, while Short-Term Reversal and Momentum categories have the largest decreases in split scores. As a result, the basis-portfolios of the cost-aware Panel Tree have more stable membership, incur lower turnover, and narrow the wedge between gross and net performance. This pattern is consistent with the regularization-portfolio perspective of Olivares-Nadal and DeMiguel (2018) and the economic feature-importance evidence in Jensen et al. (2026): the cost-aware objective drives the model parameters toward signals robust to transaction costs.

Third, the cost-aware basis-portfolios carry more asset pricing information than the cost-ignorant counterparts. Before costs, both cost-aware and cost-ignorant Panel Tree strategies earn positive and statistically significant alphas across all six benchmark factor models. After deducting transaction costs, the cost-aware net alpha ranges from 0.26% to 0.43% per month under the maximum-SR objective and from 1.17% to 2.15% under mean–variance utility, remaining positive and statistically significant across all six factor models. By contrast, the cost-ignorant net alpha is close to zero or even significantly negative against a battery of factor models. Cross-strategy spanning tests reveal the same asym-

metry. The cost-aware net return earns an alpha of 0.42% under maximum SR and 1.67% under mean–variance utility, when the cost-ignorant net return serves as the explanatory factor. In the reverse regressions, there are no significant alphas for the cost-ignorant net returns against the cost-aware net return as the explanatory factor model.

Literature. The paper contributes to three strands of research. The first studies portfolio selection over a pre-specified set of basis-portfolios. Since Markowitz (1952), this literature has developed regularized and robust portfolio methods to improve out-of-sample performance in the presence of estimation uncertainty in expected returns and covariances, including Sharpe (1971); Britten-Jones (1999); Jacobs et al. (2005); DeMiguel and Nogales (2009); DeMiguel et al. (2009). Estimation uncertainty is especially problematic for covariance matrices, motivating shrinkage and high-frequency-data methods such as those of Ledoit and Wolf (2004, 2017); Ao et al. (2019); Ledoit and Wolf (2020); Ding et al. (2021). In large asset universes, sparse and cardinality-constrained approaches select a subset of informative assets while setting other weights to zero (Gao and Li, 2013; Bertsimas and Cory-Wright, 2022; Moreira Costa et al., 2022; Xu et al., 2024). Recently, learning-based approaches incorporate investment objectives into model training to learn predictors or portfolio policies, including through deep reinforcement learning (Elmachtoub and Grigas, 2022; Babier et al., 2025; Chen et al., 2026; Simon et al., 2026; **?**). These approaches learn predictors or allocation policies for a given investment universe, whereas we use the investor's cost-aware objective to generate the menu of characteristic-managed basis-portfolios.

The second strand asks which test assets should be used to evaluate asset pricing models and how informative basis-portfolios should be constructed. Lewellen et al. (2010)

show that strong pricing results on the standard ME–B/M portfolios provide weak support for an asset pricing model. Giglio et al. (2025) show more broadly that empirical conclusions can depend on the choice of test assets especially when factors are weak. This literature shifts attention from pricing a fixed set of portfolios to designing portfolios that better span the relevant opportunity set. Existing approaches include correlation-based clustering, portfolio sorting, and machine-learning procedures. Ahn, Conrad, and Dittmar (2009) form basis assets by clustering securities on return comovement to represent the investment opportunity set and sharpen inference about the risk–return trade-off, and Bryzgalova, Pelger, and Zhu (2025) use decision trees to select low-dimensional characteristic-managed cross-sections that span the projected stochastic discount factor.

Our closest methodological benchmark is Cong, Feng, He, and He (2025), who use Panel Trees to generate characteristic-based test assets and recover a latent stochastic discount factor. Their economic problem is test-asset design: constructing portfolios that expand the mean–variance frontier and sharpen inference about the pricing kernel. In contrast, we study implementable basis-portfolio generation. The goal is to construct characteristic-managed basis-portfolios that an investor can hold and rebalance when trading is costly, so that the resulting opportunity set improves an *ex-ante* cost-aware objective. We retain the Panel Tree architecture but replace the global split criterion with a cost-aware objective that maps basis-portfolio allocations into individual asset-level holdings and evaluates the net-objective. While Cong et al. (2025) study efficient-frontier expansion for stochastic discount factor estimation, we recast Panel Trees as a cost-aware basis-design method for portfolio selection.

The third strand studies how trading frictions reshape portfolio choice. Strategies built from characteristics, factors, and machine-learning methods could be exposed to

high turnover and low net performance once costs are included (Novy-Marx and Velikov, 2016; Avramov et al., 2023; Li et al., 2024). Olivares-Nadal and DeMiguel (2018) show that incorporating transaction costs into portfolio selection is equivalent to robust portfolio optimization, regularized regression, and Bayesian portfolio choice, giving transaction-cost penalties an interpretation as stabilization against estimation error. DeMiguel, Martín-Utrera, Nogales, and Uppal (2020) show that transaction costs can change which characteristics matter jointly for investors because combining characteristics can reduce rebalancing costs through offsetting trades and increase the effective dimension of the cross section. Fan, Medeiros, Yang, and Yang (2024) develop a high-dimensional finite-horizon mean–variance estimator with proportional and quadratic transaction costs. Their framework treats transaction costs as part of the investment decision; our focus is earlier in the basis-design step on how transaction costs shape the characteristic-managed basis-portfolios before the asset allocation problem.

Our paper is also closely related to Jensen, Kelly, Malamud, and Pedersen (2026), who combine machine-learning return forecasts with cost-aware portfolio selection and evaluate strategies on the implementable efficient frontier. Their framework takes the tradable asset universe as given and learns implementable individual asset-level portfolio weights from an economic objective. We instead study the tradable menu itself. The cost-aware Panel Tree forms non-overlapping, value-weighted leaf portfolios and retains a split only if it improves the investor's *ex-ante* cost-aware objective after accounting for the individual asset-level turnover. In Jensen et al. (2026), transaction costs discipline asset allocation conditional on a pre-specified tradable universe; in our setting, we discipline basis-portfolio generation. This distinction matters because characteristic-driven asset migration across leaves is a source of turnover before portfolio weights are chosen.

A further practical distinction is the investment universe: Jensen et al. (2026) restrict attention to a balanced panel of NYSE stocks to align with their fixed-universe assumption, whereas our cost-aware Panel Tree can handle the whole unbalanced panel of US equities.

The remainder of the article is as follows. Section 2 formulates the learning problem, defines basis-portfolios and individual asset-level transaction costs, and develops the transaction-cost-aware Panel Tree. It also establishes when characteristic-managed basis-portfolios preserve the relevant individual asset-level opportunity set and when the Panel Tree's investment utility converges to that of the optimal characteristic-managed basis-portfolios, and characterizes when transaction costs change the optimal basis-portfolio design. Section 3 reports the empirical results: Subsection 3.1 describes the data and empirical design, Subsection 3.2 examines tree structure and characteristic selection, Subsection 3.3 evaluates net-of-cost performance and decomposes the improvement into basis generation and downstream allocation, and Subsection 3.4 studies asset pricing implications through factor-model and cross-strategy spanning tests. Section 4 concludes.

2 A Cost-Aware Panel Tree Framework

2.1 Economic Setup

Consider a trading universe $\mathcal{I}_t$ containing $N_t := |\mathcal{I}_t|$ individual assets, indexed by $i = 1, \ldots, N_t$, for each month t. Let $r_{i,t}$ be the excess return of asset $i \in \mathcal{I}_t$ at month t, and $r_{f,t}$ be the risk-free return. The lagged characteristic vector is $\boldsymbol{z}_{i,t-1} := (z_{i,t-1,1}, \ldots, z_{i,t-1,K})^\top$, where K is the number of characteristics and $z_{i,t-1,p} \in [0, 1]$ is the cross-sectionally rank-standardized value of characteristic p, for $p = 1, \ldots, K$.

At the beginning of month t, an investor observes asset returns $r_{i,t-1}$ and characteris-

tics $\boldsymbol{z}_{i,t-1}$ over the historical sample $\mathcal{W}_T$, where T denotes updated date. Let $W_T := |\mathcal{W}_T|$ denote the number of months in the window. In this context, an investor trains a Panel Tree model and allocates a budget across the leaf basis-portfolios generated by the Panel Tree. The tree structure and basis-portfolio allocation weights are fixed for the test sample. During the test period, assets may move across leaves as their lagged characteristics evolve, while asset weights within each leaf are proportional to their lagged market equities.

2.2 Generating Leaf Basis-Portfolios

Let Θ denote a Panel Tree model, $\mathcal{L} := \mathcal{L}(\Theta)$ denote the set of leaf basis-portfolios generated by it, and $m = |\mathcal{L}|$ denote the number of basis-portfolios, indexed by $j = 1, \ldots, m$. Let $\Theta_s = (\xi_1, \ldots, \xi_s)$ denote the tree model after s splits. Each candidate split is $\xi = (p, c, \mathcal{L}_j)$, where $p \in \{1, \ldots, K\}$ denotes a characteristic, $c \in \mathcal{Q}$ is a cutpoint from the threshold grid $\mathcal{Q}$, and $\mathcal{L}_j \in \mathcal{L}(\Theta_s)$ is the leaf to be split. The notation $\Theta_s \oplus \xi_{s+1}$ appends ξ_{s+1} to the current tree and produces Θ_{s+1}.

Let $\mathcal{I}_t^j$ denote the set of $N_t^j := |\mathcal{I}_t^j|$ underlying individual assets in the leaf-j basis portfolio during month t. For each asset $i \in \mathcal{I}_t^j$, the within-leaf value weight based on lagged market value, $mv_{i,t-1}$, is

$$\pi_{i,t} := \frac{mv_{i,t-1}}{\sum_{u=1}^{N_t^j} mv_{u,t-1}}. \tag{1}$$

The excess return on basis-portfolio j in month t is

$$R_{j,t} := \sum_{i=1}^{N_t^j} \pi_{i,t} r_{i,t}. \tag{2}$$

These portfolio returns are aggregated into the vector $\mathbf{R}_t := \left[R_{1,t}, ..., R_{m,t}\right]^\top \in \mathbb{R}^m$.

2.3 Individual Asset-Level Holdings and Proportional Transaction Costs

Given a basis-portfolio allocation vector $\boldsymbol{w} := [w_1, \ldots, w_m]^\top$, the implied holding in individual asset $i \in \mathcal{I}_t^j$ is

$$h_{i,t} = w_j \pi_{i,t}. \tag{3}$$

The strategy's gross excess return in month t is therefore $R_t^{gross} = \boldsymbol{w}^\top \mathbf{R}_t = \sum_{i=1}^{N_t} h_{i,t} r_{i,t}$, yielding a pre-cost gross wealth growth of

$$G_t := 1 + r_{f,t} + R_t^{gross}. \tag{4}$$

Prior to rebalancing, the drifted asset holding entering month $t+1$ is

$$h_{i,t}^+ := \frac{h_{i,t}(1 + r_{f,t} + r_{i,t})}{G_t}. \tag{5}$$

The trade required to achieve the target holding in month $t+1$ is $\Delta h_{i,t+1} = h_{i,t+1} - h_{i,t}^+$.

Following Brandt et al. (2009), we define the realized transaction cost in month t as

$$C_t := \sum_{i=1}^{\widetilde{N}_t} \kappa_{i,t} \left|\Delta h_{i,t}\right|, \tag{6}$$

where $\kappa_{i,t} \in \mathbb{R}^+$ denotes the proportional cost rate of asset i in month t. [1] $\kappa_{i,t}$ describes variation in transaction-cost rates over time across securities. Let $\psi_t \in \mathbb{R}^+$ denote the time-varying multiplier and $b_{i,t} \in \mathbb{R}^+$ denote the individual asset-specific component.

[1] The count $\widetilde{N}_t := |\mathcal{I}_{t-1} \cup \mathcal{I}_t|$ denotes the number of assets in the combined universe for months $t-1$ and t. For any asset i leaving the trading universe in month t, we assume $\kappa_{i,t} = \kappa_{i,t-1}$.

Their product gives the proportional cost rate: [2]

$$\kappa_{i,t} = \psi_t b_{i,t}. \tag{7}$$

Consequently, the corresponding net strategy return is

$$R_t^{\text{net}} := R_t^{\text{gross}} - C_t. \tag{8}$$

2.4 Cost-Aware Basis-Portfolio Allocation

We assume the investor considers two classic objectives $\tau \in \{\text{MVU}, \text{SR}\}$, where MVU denotes mean–variance utility and SR denotes the maximum Sharpe ratio. Let $\mathcal{A}_\tau$ denote the feasible set for the basis-portfolio allocation under objective τ. [3] For MVU, we use the ℓ_2 constraint to limit concentration, following the portfolio-norm approach in DeMiguel et al. (2009). To control leverage, we choose $16/m$ so that total gross basis-portfolio exposure is smaller than four. Hence, the feasible set is

$$\mathcal{A}_{\text{MVU}} = \left\{ \boldsymbol{w} \in \mathbb{R}^m : \mathbf{1}^\top \boldsymbol{w} = 1,\ \|\boldsymbol{w}\|_2^2 \le 16/m \right\}.$$

For SR, we normalize gross exposure to one because the gross objective value is invariant to gross exposure. Thus, we have

$$\mathcal{A}_{\text{SR}} = \{ \boldsymbol{w} \in \mathbb{R}^m : \|\boldsymbol{w}\|_1 = 1 \}.$$

[2] Empirically, following Brandt et al. (2009), we set $b_{i,t} := 0.006 - 0.0025\, mv_{i,t-1}^{rank}$, where $mv_{i,t-1}^{rank}$ denotes rank-standardized lagged market value. The scalar ψ_t decreases linearly from 4 in January 1974 to 1 in January 2002, remaining fixed at 1 thereafter.

[3] This set $\mathcal{A}_\tau$ is conditional on the current tree and its number of leaves, with that dependence suppressed locally; the later comparisons between tree models writes $\mathcal{A}(\Theta)$ explicitly.

The investor estimates $\widehat{\boldsymbol{\mu}}_T = \frac{1}{W_T}\sum_{t=1}^{W_T}\mathbf{R}_t$ and $\widehat{\boldsymbol{\Sigma}}_T = \widehat{\boldsymbol{S}}_T + \gamma\mathbf{I}_m$, where $\widehat{\boldsymbol{S}}_T$ is the sample covariance matrix, $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ is the identity matrix, and $\gamma > 0$ is the shrinkage parameter. This ridge-shrinkage specification ensures that $\widehat{\boldsymbol{\Sigma}}_T$ is positive definite, following Ledoit and Wolf (2004). The MVU training criterion penalizes gross-return variance, while the SR training criterion uses gross-return covariance in its denominator and deducts the training cost term in its numerator. Based on these, the investor optimizes the two cost-aware objective functions [4] :

$$\boldsymbol{w}^*_{\mathrm{MVU},T} = \arg\max_{\boldsymbol{w}\in\mathcal{A}_{\mathrm{MVU}}} \quad \widehat{\boldsymbol{\mu}}_T^\top \boldsymbol{w} - \frac{\delta}{2}\boldsymbol{w}^\top\widehat{\boldsymbol{\Sigma}}_T\boldsymbol{w} - \widehat{C}_T(\boldsymbol{w}), \tag{9}$$

$$\boldsymbol{w}^*_{\mathrm{SR},T} = \arg\max_{\boldsymbol{w}\in\mathcal{A}_{\mathrm{SR}}} \quad \frac{\widehat{\boldsymbol{\mu}}_T^\top \boldsymbol{w} - \widehat{C}_T(\boldsymbol{w})}{\sqrt{\boldsymbol{w}^\top\widehat{\boldsymbol{\Sigma}}_T\boldsymbol{w}}}, \tag{10}$$

$$\widehat{C}_T(\boldsymbol{w}) = \frac{1}{W_T}\sum_{t=1}^{W_T}\widetilde{C}_t(\boldsymbol{w}), \tag{11}$$

where $\widehat{C}_T$ denotes the average training cost [5] and $\delta \in \mathbb{R}^+$ denotes the risk aversion parameter.

[4] The MVU problem is a convex QCQP and the SR problem is equivalent to a SOCP. Internet Appendix C.2 gives the specific formulations, the SR recovery rule, and the convexity discussion.

[5] Computing sample-average costs from exact return-induced drift in $h^+_{i,t}$ would introduce fractional dependence on $\boldsymbol{w}$, preventing direct use of the convex programs. We therefore evaluate the wealth-growth denominator at the cost-ignorant allocation $\boldsymbol{w}^{\mathrm{CI},*}$ (in Internet Appendix C.1), while retaining $\boldsymbol{w}$ in the numerator: $h^+_{i,t-1}(\boldsymbol{w}) \approx \widetilde{h}^+_{i,t-1}(\boldsymbol{w}) := \frac{h_{i,t-1}(\boldsymbol{w})(1+r_{f,t-1}+r_{i,t-1})}{1+r_{f,t-1}+(\boldsymbol{w}^{\mathrm{CI},*}_{r,T})^\top\mathbf{R}_{t-1}}$. Define the approximate trades and per-period training costs by

$$\widetilde{\Delta h}_{i,t}(\boldsymbol{w}) := h_{i,t}(\boldsymbol{w}) - \widetilde{h}^+_{i,t-1}(\boldsymbol{w}),$$

$$\widetilde{C}_t(\boldsymbol{w}) := \sum_{i=1}^{\widetilde{N}_t}\kappa_{i,t}\left|\widetilde{\Delta h}_{i,t}(\boldsymbol{w})\right|.$$

The term $\widehat{C}_T$ averages $\widetilde{C}_t$ across sample period, while exact C_t retains the economic definition in Equation (6).

2.5 Growing the Cost-Aware Panel Tree

For each candidate split, the objective value at $\boldsymbol{w}^*_{\tau,T}$ is:

$$\mathcal{U}_{\tau,T} = \begin{cases} \widehat{\boldsymbol{\mu}}_T^\top \boldsymbol{w}^*_{\tau,T} - \dfrac{\delta}{2} \boldsymbol{w}^{*\top}_{\tau,T} \widehat{\boldsymbol{\Sigma}}_T \boldsymbol{w}^*_{\tau,T} - \widehat{C}_T(\boldsymbol{w}^*_{\tau,T}), & \tau = \text{MVU}, \\ \dfrac{\widehat{\boldsymbol{\mu}}_T^\top \boldsymbol{w}^*_{\tau,T} - \widehat{C}_T(\boldsymbol{w}^*_{\tau,T})}{\sqrt{\boldsymbol{w}^{*\top}_{\tau,T} \widehat{\boldsymbol{\Sigma}}_T \boldsymbol{w}^*_{\tau,T}}}, & \tau = \text{SR}. \end{cases} \tag{12}$$

Using the split notation defined in Section 2.2, the cost-aware gain from split ξ is:

$$\Delta\mathcal{U}_{\tau,T}(\xi; \Theta_s) = \mathcal{U}_{\tau,T}(\Theta_s \oplus \xi) - \mathcal{U}_{\tau,T}(\Theta_s). \tag{13}$$

At each step, the cost-aware Panel Tree selects the admissible split with the largest positive gain:

$$\xi^*_{s+1} \in \arg\max_{\xi} \Delta\mathcal{U}_{\tau,T}(\xi; \Theta_s), \qquad \Delta\mathcal{U}_{\tau,T}(\xi^*_{s+1}; \Theta_s) > 0. \tag{14}$$

Tree growth stops when the split budget is exhausted or no admissible split yields a positive cost-aware gain. Algorithm 1 summarizes the procedure.

2.6 Theoretical Foundations for Basis-Portfolio Generation

We derive three theorems to justify the theoretical soundness of cost-aware Panel Tree in basis-portfolio generation and asset allocation. Theorem 1 relates the optimal portfolio choice in individual assets to the optimal portfolio choice in characteristic-managed basis-portfolios. Theorem 2 bounds the utility gap between the leaf basis-portfolios generated by Panel Tree and the optimal characteristic-managed basis-portfolios. Theorem 3 shows that transaction costs reshape leaf basis-portfolio generation.

The theoretical analysis focuses on the population mean–variance utility (MVU) criterion. Let g denote a characteristic-based generation rule, with $g = \Theta$ corresponding to a Panel Tree model. Let $\mathcal{A} = \mathcal{A}_{\tau=\text{MVU}}$ denote the feasible set of asset allocation weights among leaf basis-portfolios, and $\boldsymbol{w} \in \mathcal{A}$ denote a vector of portfolio weights. Then, we have the strategy's gross excess return:

$$R_t(g, \boldsymbol{w}) := \boldsymbol{w}^\top \mathbf{R}_t(g). \tag{15}$$

The population objective before transaction costs is

$$\mathcal{U}_0(g, \boldsymbol{w}) := \mathbb{E}[R_t(g, \boldsymbol{w})] - \frac{\delta}{2} \operatorname{Var}[R_t(g, \boldsymbol{w})]. \tag{16}$$

2.6.1 Spanning by characteristic-managed basis-portfolios

The underlying asset weights in characteristic-managed basis-portfolios depend on lagged asset characteristics. According to Cochrane (2011), expected returns, variances, and covariances are stable functions of asset characteristics. Therefore, characteristics guide the investor's decision making. We write the optimal portfolio weight of asset i as $h^*_{i,t} = q_{i,t} f^*(\boldsymbol{z}_{i,t-1})$, where $q_{i,t} := mv_{i,t-1} / \sum_{u=1}^{N_t} mv_{u,t-1}$ is its market-portfolio benchmark weight and f^* is a general latent function that maps asset characteristics to a multiplicative adjustment to this weight.

Every period, assets are assigned to m non-overlapping characteristic-managed portfolios, indexed by $j = 1, \ldots, m$. Portfolio j contains N_t^j assets, and $\sum_{j=1}^{m} N_t^j = N_t$. Individual assets are clustered by similarity in their characteristics and the implied expected returns and covariances.

Let $\phi_{j,t}$ denote a general latent function shared by the assets in portfolio j, which maps $q_{i,t}$ to their within-portfolio weight $q_{i,t}\phi_{j,t}(\boldsymbol{z}_{i,t-1})$, with $\sum_{i=1}^{N_t^j} q_{i,t}\phi_{j,t}(\boldsymbol{z}_{i,t-1}) = 1$. An investor allocates weight w_j to portfolio j. Her holding of asset i in portfolio j is

$$h_{i,t} := w_j q_{i,t}\phi_{j,t}(\boldsymbol{z}_{i,t-1}). \tag{17}$$

For asset i in portfolio j, we define $f_{m,t}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w}) := w_j\phi_{j,t}(\boldsymbol{z}_{i,t-1})$. The basis-portfolio returns and strategy returns are

$$\begin{aligned} R_{j,t}(g) &:= \sum_{i=1}^{N_t^j} q_{i,t}\phi_{j,t}(\boldsymbol{z}_{i,t-1})r_{i,t}, \\ R_t(g, \boldsymbol{w}) &= \sum_{j=1}^{m}\sum_{i=1}^{N_t^j} q_{i,t} f_{m,t}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w})r_{i,t} = \sum_{j=1}^{m} w_j R_{j,t}(g). \end{aligned} \tag{18}$$

The portfolio-based $f_{m,t}$ serves as an approximation to f^*.

We assume the strategy return under the optimal individual asset-level allocation is $R_t^{\text{asset},*} = (\boldsymbol{h}_t^*)^\top \boldsymbol{r}_t$, where $\boldsymbol{r}_t$ is the vector of individual-asset excess returns. Define

$$\begin{aligned} \mathcal{U}_0^{\text{basis},*} &:= \max_{g,\boldsymbol{w}} \mathcal{U}_0(g, \boldsymbol{w}), \\ \mathcal{U}_0^{\text{asset},*} &:= \max_{\boldsymbol{h}} \left\{ \mathbb{E}[\boldsymbol{h}_t^\top \boldsymbol{r}_t] - \frac{\delta}{2}\operatorname{Var}(\boldsymbol{h}_t^\top \boldsymbol{r}_t) \right\}. \end{aligned} \tag{19}$$

Theorem 1. *Under the feasibility and optimality conditions in Internet Appendix B.1.1, the utility loss from investing through characteristic-managed basis-portfolios satisfies*

$$\mathcal{U}_0^{\text{asset},*} - \mathcal{U}_0^{\text{basis},*} = \frac{\delta}{2}\min_{g,\boldsymbol{w}} \operatorname{Var}\left[R_t^{\text{asset},*} - R_t(g, \boldsymbol{w})\right] = \frac{\delta}{2}\rho_m^2, \tag{20}$$

where ρ_m^2 denotes the minimum return-difference variance across m basis-portfolios. If characteristic groups become finer as $m \to \infty$ and their fixed allocations satisfy the within-group approximation and return-moment conditions in the Internet Appendix Lemma 1, the corresponding strategy returns converge in mean square to $R_t^{\text{asset},*}$, $\rho_m \to 0$, and $\mathcal{U}_0^{\text{asset},*} - \mathcal{U}_0^{\text{basis},*} \longrightarrow 0$.

The identity measures the utility cost of basis-portfolio investment by the smallest attainable variance of the return difference from the optimal individual asset-level strategy. Under the stated conditions, the utility loss relative to the optimal individual asset-level strategy vanishes as characteristic groups become finer.

2.6.2 Panel Tree recovery of optimal basis-portfolios

We first consider basis-portfolios generated by general iterative trees. For a generation rule g, write $\mathcal{U}_0(g) := \max_{\boldsymbol{w}} \mathcal{U}_0(g, \boldsymbol{w})$ for its optimal mean–variance utility. For this comparison, $\mathcal{U}_0^{\text{basis},*} := \sup_{m \geq 1} \mathcal{U}_0^{\text{basis},*}(m)$ denotes the optimal value over all feasible numbers of basis-portfolios, where $\mathcal{U}_0^{\text{basis},*}(m)$ makes explicit the portfolio count suppressed in Equation (19).

A Panel Tree starts with the full set of individual assets. Each split selects an existing leaf and divides it into two using any admissible characteristic and threshold. It replaces one leaf with two, so L leaf basis-portfolios require $L - 1$ splits. The optimal mean–variance utility over feasible iterative trees with at most L leaves is

$$\mathcal{U}_0^{\text{tree},*}(L) := \max_{\mathcal{T}} \mathcal{U}_0(\mathcal{T}), \tag{21}$$

where $\mathcal{T}$ denotes an iterative-tree generation rule. For each finite L, the threshold grid is finite and the maximum is attained under the conditions in Internet Appendix B.2. We ask

whether the optimal utility from leaf basis-portfolios converges to that from characteristic-managed basis-portfolios as L increases. Their difference in investment utility is

$$A_L := \mathcal{U}_0^{\text{basis},*} - \mathcal{U}_0^{\text{tree},*}(L). \tag{22}$$

Let Θ_{L-1} denote the Panel Tree generated with at most L leaves, retaining the terminal tree through step $L-1$ if growth stops early. We compare the investment utility of its basis-portfolios with that of the optimal characteristic-managed basis-portfolios under the same limit of L portfolios. The difference in investment utility is

$$\Gamma_L := \mathcal{U}_0^{\text{tree},*}(L) - \mathcal{U}_0(\Theta_{L-1}). \tag{23}$$

Theorem 2. *The Panel Tree's utility loss satisfies*

$$0 \leq \mathcal{U}_0^{\text{basis},*} - \mathcal{U}_0(\Theta_{L-1}) = A_L + \Gamma_L. \tag{24}$$

Suppose $\mathcal{U}_0^{\text{basis},*}$ is finite and there are feasible basis-portfolio strategies with utility arbitrarily close to it that satisfy the conditions in Internet Appendix Lemma 2. These conditions require continuity of $f_{m,t}$, feasible fixed allocations within finer characteristic groups, and suitable iterative partitions and return moments. Then these strategies can be approximated by leaf basis-portfolios, and $A_L \to 0$ as $L \to \infty$. If feasible splits also continue to deliver sufficient utility gains under the conditions in Internet Appendix Lemma 3, then $\Gamma_L \to 0$ and

$$\mathcal{U}_0^{\text{basis},*} - \mathcal{U}_0(\Theta_{L-1}) \longrightarrow 0. \tag{25}$$

The construction connects optimal characteristic-managed basis-portfolios to leaf basis-portfolios formed from finer characteristic groups. The condition on split gains connects the optimal iterative tree to the Panel Tree selected under the global economic objective. The Internet Appendix Corollary 1 combines this result with Theorem 1 to establish convergence to the optimal investment utility from individual assets.

2.6.3 Transaction costs and basis-portfolio generation

We ask whether cost-aware basis-portfolio generation adds value beyond cost-aware reallocation on a cost-ignorant basis. Consider a common class $\mathbb{T}_L$ of admissible Panel Trees with at most L leaves. Both regimes allow the same characteristics and split thresholds under the same leaf feasibility rules. For each tree Θ, let $\mathcal{A}(\Theta)$ be the MVU allocation set in Section 2.4, whose dimension and norm bound depend on leaf count. The feasible domain is

$$\mathfrak{F}_L := \{(\Theta, \boldsymbol{w}) : \Theta \in \mathbb{T}_L,\ \boldsymbol{w} \in \mathcal{A}(\Theta)\}. \tag{26}$$

Both regimes share the within-leaf weighting convention, transaction-cost schedule, and initial stock holdings. Let $\lambda \in \{0, 1\}$ indicate whether expected transaction costs enter the objective. The optimized value is

$$\mathcal{U}_\lambda(\Theta) := \sup_{\boldsymbol{w} \in \mathcal{A}(\Theta)} \{\mathcal{U}_0(\Theta, \boldsymbol{w}) - \lambda \mathbb{E}[C_t(\Theta, \boldsymbol{w})]\}. \tag{27}$$

Here $\mathcal{U}_0(\Theta, \boldsymbol{w})$ is the gross-return mean–variance criterion in Equation (16), and C_t is the stock-level economic cost in Equation (6). The two optimal trees satisfy

$$\Theta^0 \in \arg\max_{\Theta \in \mathbb{T}_L} \mathcal{U}_0(\Theta), \qquad \Theta^1 \in \arg\max_{\Theta \in \mathbb{T}_L} \mathcal{U}_1(\Theta). \tag{28}$$

Both bases are evaluated with cost-aware allocation, so $\mathcal{U}_1(\Theta^0)$ includes optimal reallocation on the cost-ignorant basis.

For each feasible tree, define

$$\begin{aligned} \mathcal{D}(\Theta) &:= \mathcal{U}_0(\Theta^0) - \mathcal{U}_0(\Theta), \\ \mathcal{K}(\Theta) &:= \mathcal{U}_0(\Theta) - \mathcal{U}_1(\Theta). \end{aligned} \tag{29}$$

The first term measures the gross-return utility forgone by changing the basis. The second measures the loss from trading frictions after optimal reallocation on that basis.

Theorem 3. *Suppose the feasible domain $\mathfrak{F}_L$ is the same for $\lambda = 0$ and $\lambda = 1$, the relevant values are finite, the within-tree suprema and the two design maxima are attained, and stock-level costs are nonnegative and integrable. Then $\mathcal{D}(\Theta) \geq 0$ and $\mathcal{K}(\Theta) \geq 0$ for every feasible tree, and*

$$\begin{aligned} \mathcal{U}_1(\Theta^1) - \mathcal{U}_1(\Theta^0) &= \max_{\Theta \in \mathbb{T}_L} \{\mathcal{K}(\Theta^0) - \mathcal{K}(\Theta) - \mathcal{D}(\Theta)\} \\ &\geq 0. \end{aligned} \tag{30}$$

The gain is strictly positive if and only if there exists a feasible tree $\widetilde{\Theta}$ such that

$$\mathcal{K}(\Theta^0) - \mathcal{K}(\widetilde{\Theta}) > \mathcal{D}(\widetilde{\Theta}). \tag{31}$$

Equality holds if and only if the selected Θ^0 also maximizes $\mathcal{U}_1$ over $\mathbb{T}_L$.

Cost-aware redesign improves utility if and only if a feasible change in the basis reduces the friction loss by more than the gross-return utility forgone. Internet Appendix B.3 provides the proof, and Internet Appendix B.4 derives implications for split selection and a sufficient condition for strict improvement.

2.7 Portfolio Generation versus Allocation

We cross cost-ignorant (CI) and cost-aware (CA) rules for portfolio generation and allocation to distinguish gains from the two stages. Let $\widehat{\Theta}_\tau^{x_1}$ denote the tree generated under rule $x_1 \in \{\mathrm{CI}, \mathrm{CA}\}$, and let $\boldsymbol{w}_\tau^{x_1,x_2}$ denote its optimal basis-portfolio allocation under rule $x_2 \in \{\mathrm{CI}, \mathrm{CA}\}$. The CI rule omits transaction costs from the relevant objective, whereas the CA rule includes them as specified in Sections 2.4 and 2.5.

In each pair (x_1, x_2), the first entry denotes the portfolio-generation rule and the second denotes the allocation rule. Thus, $(\mathrm{CI}, \mathrm{CI})$ is the fully cost-ignorant benchmark, and $(\mathrm{CA}, \mathrm{CA})$ is the fully cost-aware Panel Tree. The hybrid $(\mathrm{CI}, \mathrm{CA})$ applies cost awareness only to allocation, whereas $(\mathrm{CA}, \mathrm{CI})$ applies it only to generation. Specifically, $(\mathrm{CI}, \mathrm{CA})$ reoptimizes basis-portfolio weights on the CI-generated tree under the cost-aware objective, whereas $(\mathrm{CA}, \mathrm{CI})$ reoptimizes basis-portfolio weights on the CA-generated tree without accounting for transaction costs. All four strategies are evaluated under the same realized transaction-cost schedule.

The population comparison in Theorem 3 allows basis-portfolio generation to change while applying cost-aware allocation to both bases. Its empirical counterpart is the comparison of CA–CA with CI–CA. Our experiment evaluates this comparison using estimated trees and realized out-of-sample net returns, whose performance is not guaranteed by the population result.

Let $\bar{R}_\tau^{\mathrm{net,OOS}}(x_1, x_2)$ and $s_\tau^{\mathrm{net,OOS}}(x_1, x_2)$ denote the sample mean and standard devia-

tion of monthly out-of-sample net excess returns. The reported performance is

$$\mathcal{U}_\tau^{\mathrm{OOS}}(x_1, x_2) = \begin{cases} \bar{R}_\tau^{\mathrm{net,OOS}}(x_1, x_2) - \dfrac{\delta}{2}\left(s_\tau^{\mathrm{net,OOS}}(x_1, x_2)\right)^2, & \tau = \mathrm{MVU}, \\ \dfrac{\bar{R}_\tau^{\mathrm{net,OOS}}(x_1, x_2)}{s_\tau^{\mathrm{net,OOS}}(x_1, x_2)}, & \tau = \mathrm{SR}. \end{cases} \tag{32}$$

We report OOS performance with net-return moments, including realized variation in costs. For ease of interpretation, we multiply MVU by 100 and annualize SR.

When $\mathcal{U}_\tau^{\mathrm{OOS}}(\mathrm{CA}, \mathrm{CA}) > \mathcal{U}_\tau^{\mathrm{OOS}}(\mathrm{CI}, \mathrm{CI})$, the percentage of the full cost-aware gain recovered by each hybrid is

$$100 \times \frac{\mathcal{U}_\tau^{\mathrm{OOS}}(x_1, x_2) - \mathcal{U}_\tau^{\mathrm{OOS}}(\mathrm{CI}, \mathrm{CI})}{\mathcal{U}_\tau^{\mathrm{OOS}}(\mathrm{CA}, \mathrm{CA}) - \mathcal{U}_\tau^{\mathrm{OOS}}(\mathrm{CI}, \mathrm{CI})}, \qquad (x_1, x_2) \in \{(\mathrm{CI}, \mathrm{CA}), (\mathrm{CA}, \mathrm{CI})\}. \tag{33}$$

3 Empirical Study

3.1 Data and Empirical Design

In the empirical study, we apply the framework to U.S. common stocks. We use the monthly panel data from 1974 to 2023 contributed by Jensen et al. (2023), containing stock returns and 153 cross-sectionally rank-standardized firm characteristics.[6] We re-estimate the models annually using a rolling window of the preceding 10 years for each of the 40 out-of-sample years $y \in \{1984, \ldots, 2023\}$. Once estimated, the models remain fixed throughout year y, while firm characteristics, leaf membership, and within-leaf weights

[6] We restrict the sample to common stocks listed on U.S. exchanges and require non-missing lagged market equity and returns, together with positive book equity and lagged market equity. Each characteristic is standardized cross-sectionally to lie in the $[0, 1]$ interval. Missing characteristic values are first imputed using industry medians and then, when necessary, cross-sectional medians. Internet Appendix D provides additional details.

are updated monthly. We evaluate both strategies with the proportional transaction-cost rate in (7), which varies with firm size and calendar time. Empirically, we report *gross* and *net* out-of-sample performance, along with turnover.

The proposed Panel Tree models have parameters in common with other tree models. We set the candidate thresholds to the quintiles $\mathcal{Q} = \{0.2, 0.4, 0.6, 0.8\}$. We require each candidate split to provide at least $n_{\min} = 20$ stocks in each child node in every month of the training sample. Each tree is allowed at most nine splits. For covariance estimation, we use a ridge-regularized covariance estimator with $\gamma = 10^{-4}$ as in Cong et al. (2025).

We compare cost-aware and cost-ignorant Panel Trees estimated on the same sample, under the same rolling-window protocol, and evaluated using the same transaction-cost schedule. These two specifications differ only in whether transaction costs enter the training objective. We study the maximum Sharpe ratio (SR) and mean–variance utility (MVU) with risk aversion $\delta = 2$ for demonstration.

The empirical analyses focus on three questions. First, how does incorporating transaction costs change the structure of the Panel Tree and the characteristics used to generate basis-portfolios? Second, how does cost awareness change the net performance, transaction costs, and turnover of Panel Tree strategies? Third, do the cost-aware strategy returns contain asset pricing information beyond conventional factor models and the cost-ignorant strategy returns? We examine these questions in Subsections 3.2, 3.3, and 3.4.

3.2 Panel Tree Structure

This subsection examines how incorporating transaction costs alters the structure of the estimated Panel Trees and the characteristics they select. We proceed at two levels.

First, using the latest estimation window (January 2013–December 2022), we compare the cost-ignorant and cost-aware Panel Tree diagrams and the specific characteristics each tree selects to form its $m = 10$ basis-portfolios under the maximum Sharpe-ratio objective. Second, we aggregate characteristic-selection patterns across all 40 rolling windows to verify that the structural differences are not specific to a single training sample but instead reflect a systematic reallocation of the tree's split capacity. [7] [8]

Panel Tree Diagrams. Figures 1 and 2 report the cost-ignorant and cost-aware Panel Trees under the maximum Sharpe-ratio objective with $m = 10$ leaves, estimated over the latest sample, January 2013–December 2022. Table 1 lists the characteristics selected in each Panel Tree, with the name, description, and category following Jensen et al. (2023).

In Figure 1, the root node of the cost-ignorant Panel Tree splits the cross section by whether a firm's highest 5-day return (`rmax5_21d`, Bali et al., 2011) is below or above 0.2 (20% quantile). On both sides of the split, the Panel Tree then applies the same second-stage screen based on Amihud illiquidity (`ami_126d`, Amihud, 2002) at 0.8: S2 refines the low-`rmax5_21d` branch, and S3 refines the high-`rmax5_21d` branch. From there the two sides diverge. On the S2 side, the split creates two terminal leaves. On the S3 side, the tree keeps refining the partition: S4 splits on equity net payout (`eqnpo_12m`, Daniel and Titman, 2006) at 0.2, and S5 splits on R&D capital scaled by assets (`rd5_at`, Li, 2011) at 0.6. Later splits use payout yield (`eqpo_me`), sales growth (`sale_gr1`), zero trades (`zero_trades_21d`), and Dimson beta (`beta_dimson_21d`). After nine splits, the cost-ignorant Panel Tree generates 10 leaf basis-portfolios.

[7] Split capacity is the number of splits made by a Panel Tree subject to a prespecified limit on tree growth. In our empirical setting, each tree is allowed at most nine splits. Its allocation across characteristics reflects which characteristics are selected and how often they are used to split the asset universe.

[8] The Internet Appendix E reports the corresponding structural evidence for MVU with $\delta = 2$.

In Figure 2, the root node of the cost-aware Panel Tree splits on sales-per-employee growth (sale_emp_gr1, Abarbanell and Bushee, 1998) at 0.8, and most subsequent refinement occurs on the high-sale_emp_gr1 branch. Split S2 further partitions that branch on nonannual seasonality (seas_1_1na, Heston and Sadka, 2008) at 0.6. Subsequent splits on the high-sale_emp_gr1 branch rely primarily on accounting-fundamental signals, including operating accruals (oaccruals_ni), consecutive earnings increases (ni_inc8q), EBITDA relative to enterprise value (ebitda_mev), changes in long-term investments (lti_gr1a), and cash-to-assets (cash_at). At S8, the tree returns to the low-sale_emp_gr1 branch and splits it on dollar trading volume (dolvol_126d) at 0.2. The final split again refines the high-sale_emp_gr1 branch using market equity (market_equity) at 0.2. After nine splits, the cost-aware Panel Tree also yields 10 basis-portfolios.

The cost-ignorant and cost-aware Panel Trees produce noticeably different basis-portfolios. The contrast starts at the root and carries through to the overall mix of selected characteristics. In Figure 1, the cost-ignorant Panel Tree splits first on rmax5_21d, a price-volume signal associated with lottery demand and speculative trading (Bali et al., 2011). In Figure 2, the cost-aware Panel Tree instead splits first on sales-per-employee growth (sale_emp_gr1), a slower-moving accounting-fundamental signal related to labor force efficiency (Abarbanell and Bushee, 1998). Since the root creates the coarsest partition of the cross section, replacing a fast-moving price-volume signal with a slower-moving accounting-fundamental signal tends to make the downstream partitions less sensitive to monthly characteristic variations.

Table 1 shows that cost-aware training shifts characteristic selection toward accounting-fundamental signals. In the cost-ignorant Panel Tree (Panel A), five of the nine selected characteristics are price-volume signals, and the first three upper-level splits use short-

horizon trading or illiquidity measures. In the cost-aware Panel Tree (Panel B), six of the nine selected characteristics are accounting-fundamental signals, including `sale_emp_gr1`, `oaccruals_ni`, and `ebitda_mev`. Because split characteristics determine leaf assignment, this shift toward slower-moving signals tends to generate more persistent leaf membership and less stock-level turnover.

The cost-aware tree also retains nonannual seasonality (`seas_1_1na`), dollar trading volume (`dolvol_126d`), and market equity (`market_equity`). Conditional on the root partition, `dolvol_126d` further separates stocks in the low-`sale_emp_gr1` region and improves the optimized cost-aware training criterion. This criterion evaluates the split through the joint expected returns, covariances, and transaction costs of the resulting multi-characteristic leaf basis-portfolios. The resulting conditional portfolio contribution is distinct from the characteristic's unconditional return premium. This interpretation is consistent with the multivariate, transaction-cost perspective of DeMiguel et al. (2020), under which characteristics can create portfolio value through their interactions in the joint investment opportunity set.

Characteristic Selection. The single-window comparison above documents a structural shift from price-volume signals toward accounting-fundamental signals under cost-aware training. We now ask whether this shift persists across the full out-of-sample exercise or is specific to the latest estimation window. Because each Panel Tree uses exactly nine splits, we can track how a fixed split budget is allocated across characteristic categories over 40 rolling windows. We summarize this allocation using split frequency and a split-order-weighted score that assigns 9 to the first split, 8 to the second, and so on down to 1 for the ninth split, so that upper-level splits receive higher weights.

Table 2 reports split frequency and score aggregated by meta-category (Panel A) and disaggregated by characteristic category (Panel B), for both the cost-ignorant (CI) and cost-aware (CA) Panel Trees across 40 rolling windows. Panel A shows that cost-aware training tilts the split budget toward accounting-fundamental signals. The accounting-fundamental score share rises from 50.89% under the cost-ignorant Panel Tree to 61.72% under the cost-aware Panel Tree, and its frequency share rises from 54.17% to 64.72%. Because the score measure weights upper-level splits more heavily, the 10.83 percentage-point increase in score share implies that slower-moving characteristics are selected earlier in the sequential tree splits.

Panel B shows that the reallocation is concentrated in specific categories. The largest score increases occur in Quality ($+4.18$), followed by Low Risk ($+1.48$), Investment ($+1.15$), and Accruals ($+0.63$). These are characteristic categories whose underlying signals are mostly updated at quarterly or annual frequency, so they generate relatively stable leaf membership. The largest declines occur in Short-Term Reversal (-1.90), Momentum (-1.80), Profit Growth (-1.45), and Size (-1.45). Short-Term Reversal is rarely selected under cost-aware training. This finding is consistent with the economic feature importance results of Jensen et al. (2026), who show that short-term reversal is far less important for a large investor because its predictive ability does not compensate for the turnover it requires. In our setting, the mechanism operates at the level of basis-portfolio generation: the cost-aware Panel Tree does not merely downweight short-horizon signals in the final allocation but substantially reduces their importance in defining the basis-portfolios.

Figure 3 reports the average proportion of total splits assigned to each of the 13 characteristic categories following Jensen et al. (2023) for the cost-ignorant and cost-aware Panel Trees, respectively. It confirms that the reallocation is selective. Cost-aware train-

ing increases the split shares of Quality, Investment, Value, and Low Risk, each of which accounts for more than 10% of total splits. By contrast, it reduces the shares of Momentum and Short-Term Reversal to below 2.5%. Some price-volume categories, such as Low Risk, are retained because they define comparatively stable partitions of the cross section. The decline of Profit Growth shows that cost-aware training does not mechanically favor every accounting-fundamental category. Instead, the reallocation is selective and is consistent with a joint trade-off between portfolio contribution and transaction costs. This pattern is consistent with the theoretical result of Jensen et al. (2026), who show that feature importance depends on two terms: the value the feature adds to the frictionless Markowitz portfolio and the transaction cost it generates through portfolio turnover. In our context, accounting-fundamental signals that score high on the first term and low on the second term are promoted, while price-volume signals with the opposite profile are demoted.

Figure 4 demonstrates that the shift in characteristic selection persists throughout the majority of the sample period. The heatmap illustrates the difference in split scores (cost-aware minus cost-ignorant) for each characteristic category across the 40 out-of-sample windows, revealing predominantly positive differentials for Quality, Investment, and Low Risk, alongside negative differentials for Momentum and Short-Term Reversal.

In summary, these results deliver a clear message. Incorporating transaction costs in the Panel Tree objective shifts basis-portfolio generation structurally, redirecting the model away from fast-moving price-volume signals and toward slower-moving accounting-fundamental signals. The next subsection examines whether this structural shift translates into improved net-of-cost performance out of sample.

3.3 Out-of-sample Net Performance

Baseline Comparison. Table 3 reports the *net-of-cost* out-of-sample performance for the cost-ignorant (CI) and cost-aware (CA) Panel Trees. Panels A and B present the maximum-SR and MVU objectives, respectively, and Panel C reports equal-weighted (EW) and value-weighted (VW) market-portfolio benchmarks. For each specification, we report gross and net return, volatility, and objective value, as well as average monthly transaction cost and turnover. The benchmark objectives in Panel C are annualized Sharpe ratios and therefore provide a direct comparison for Panel A.

Table 3 yields three findings. First, the cost-aware Panel Tree trades less than the cost-ignorant counterpart. Under the Sharpe ratio objective, average monthly transaction costs fall from 0.56% for CI to 0.34% for CA, and turnover falls from 85.87% to 53.16%. Under mean–variance utility with $\delta = 2$, average monthly transaction costs fall from 2.31% to 1.38%, while turnover declines from 351.92% to 220.86%. Once transaction costs enter the split decision, the Panel Tree incurs much lower turnover at the stock level.

Second, the lower-turnover profile is accompanied by higher net performance. Under the cost-aware Panel Tree, the net Sharpe ratio rises from 0.17 under CI to 0.98 under CA, and the net mean–variance utility with $\delta = 2$ rises from 0.62 to 1.40. Further, the net Sharpe ratio also exceeds the EW and VW benchmark values of 0.22 and 0.55 in Panel C.

Third, the improvement in net performance comes from higher net mean returns rather than lower volatility. Under the Sharpe-ratio objective, the monthly net return rises from 0.07% for the cost-ignorant Panel Tree to 0.46% for the cost-aware Panel Tree, even as net volatility increases from 1.32% to 1.63%. Under mean–variance utility with $\delta = 2$, the monthly net return increases from 1.69% to 2.51%, while net volatility rises from 10.33% to 10.58%. The sources of the higher net return differ across the two objectives.

Under the Sharpe-ratio objective, cost-aware training combines a higher gross return with lower transaction costs. Under mean–variance utility, the gross return is slightly lower, but the reduction in transaction costs more than offsets this difference. Thus, the higher net objective reflects the preservation of more average return after transaction costs, despite higher realized volatility. This pattern is consistent with the regularization view of portfolio choice with transaction costs in Olivares-Nadal and DeMiguel (2018) and with the structural reallocation documented in Section 3.2.

Sources of the Net Performance Gain. The baseline comparison does not reveal whether the higher net objective comes mainly from generating a different set of basis-portfolios or from allocating differently across a given set. We address this question using the 2×2 basis-portfolio generation versus allocation experiment defined in Section 2.7. For every annual estimation window, we pair each set of leaf basis-portfolios generated by the estimated cost-ignorant (CI) or cost-aware (CA) Panel Tree with each CI or CA allocation rule. Changing the allocation rule leaves the estimated tree structure and basis-portfolio definition unchanged, while stocks continue to migrate across leaves and within-leaf stock weights continue to update during the out-of-sample year. These four cells use the same stock universe, covariance estimator, rolling-window protocol, and realized proportional-cost schedule.

Table 4 reports the results. Panel A presents the maximum-SR objective, and Panel B presents MVU with $\delta = 2$. The first element of each strategy label identifies the tree-generation rule, and the second identifies the allocation rule. For instance, CI–CA means a strategy with cost-ignorant basis-portfolios and cost-aware asset allocation. Within each panel, the CI–CI strategy is the zero benchmark and the CA–CA strategy is the 100% end-

point. The third and fourth columns report the increase in the net objective relative to CI–CI and the percentage of the total CI–CI-to-CA–CA improvement recovered by each specification, respectively. The percentages compare each hybrid separately with the two endpoints and therefore are not additive. Following DeMiguel et al. (2024), the final column reports bootstrap-based one-sided p-values for the null that CA–CA does not deliver a higher net objective than the strategy in the corresponding row.

First, most of the net performance improvement originates from basis-portfolio generation. Under maximum SR, CA–CI delivers a net objective of 0.91, compared with 0.17 for CI–CI and 0.98 for CA–CA, and therefore recovers 91.77% of the total improvement. Although cost-aware allocation accounts for only the remaining 8.23% of the gain, CA–CA still outperforms CA–CI with a p-value of 0.003. Under MVU, CA–CI raises the net objective from 0.62 to 1.40, marginally above the CA–CA point estimate before rounding, and thus recovers 100.87% of the total improvement. The p-value of 0.620 provides no evidence that CA–CA outperforms CA–CI. Their nearly identical net objectives indicate that, conditional on the CA tree, CA allocation has little incremental effect under MVU. These results attribute the bulk of the cost-aware performance gain to basis-portfolio generation rather than to the subsequent allocation decision.

Second, the complementary CI–CA comparison shows that downstream allocation accounts for a smaller share of the improvement. Applying the CA allocation rule to the original CI basis-portfolios raises the net objective to 0.43 under maximum SR and to 0.83 under MVU, recovering 32.26% and 26.13% of the respective endpoint gains. Cost-aware allocation improves the point estimate for a fixed CI basis-portfolio set, but it cannot reproduce the gains achieved when transaction costs shape the basis-portfolios themselves. Taken together, the two hybrid comparisons indicate that basis-portfolio generation is the

primary source of the CA gain, while allocation operates through a secondary, partially overlapping margin.

Time-Series Analysis. We next examine whether the gains from cost-aware Panel Trees are persistent over time or driven by a few episodes. Figure 5 plots the 10-year rolling out-of-sample Sharpe ratios of the two strategies, complementing the full-sample averages in Table 3. Panel (a) is for the cost-ignorant Panel Tree and Panel (b) for the cost-aware Panel Tree. In each panel, the gross and net series depict the 10-year rolling out-of-sample Sharpe ratios before and after transaction costs, respectively, and the horizontal dashed and dotted lines mark the corresponding OOS average Sharpe ratios. Figure 6 presents the same rolling comparison for mean–variance utility with $\delta = 2$.

In Figure 5, the OOS average net Sharpe ratio rises from 0.17 for the cost-ignorant strategy to 0.98 for the cost-aware strategy. The cost-ignorant specification leaves a large wedge between gross and net performance for much of the sample, whereas the corresponding wedge is visibly narrower under the cost-aware Panel Tree. The cost-aware net series also remains above its cost-ignorant counterpart across most of the rolling evaluation period, indicating that the improvement is not confined to a few favorable episodes.

Figure 6 shows a similar time-series pattern under mean–variance utility. The OOS average net utility rises from 0.62 to 1.40 for $\delta = 2$. The time-series plots show a narrower gross-net wedge and a higher net series for the cost-aware Panel Tree than the cost-ignorant model over most of the sample, with the contrast especially clear around the 2008 global financial crisis. Together, the time-series plots show that cost-aware training reduces the strategy's sensitivity to trading frictions throughout the evaluation period.

Cumulative Wealth. Figure 7 reports the cumulative net excess returns of the cost-ignorant and cost-aware Panel Trees under the maximum-Sharpe objective over the 1984–2023 out-of-sample period. The shaded region highlights the cumulative performance gap between the two strategies and shows how the monthly net-performance differences accumulate over time. The cumulative net excess return of the cost-aware strategy begins to separate from that of the cost-ignorant strategy in the late 1980s. The gap widens further through the 2000s and 2010s, and remains substantial at the end of the sample. By 2023, the cost-aware Panel Tree reaches 742.3% in cumulative net excess return, compared with 31.8% for the cost-ignorant Panel Tree.

3.4 Asset Pricing Implications

The preceding subsections establish that cost-aware training improves net-of-cost performance and alters the composition of basis-portfolios. These findings raise a deeper question: does incorporating transaction costs in Panel Trees impact the pricing information in the generated basis-portfolios? Since each Panel Tree generates a set of tradable leaf basis-portfolios, its maximum-SR allocation defines a tangency portfolio that serves as a pricing-kernel proxy; see Cochrane (2009). Therefore, incorporating transaction costs may therefore change the pricing-kernel proxy implied by the Panel Tree. We address this question in two steps. First, we examine whether the net-of-cost returns generated by the cost-aware and cost-ignorant Panel Trees are spanned by standard asset pricing factor models. Second, we run cross-strategy spanning regressions on the gross and net strategy variants to examine whether the cost-aware Panel Tree shifts the payoff space in economically implementable directions.

Factor Spanning Tests. To assess the asset pricing content of the two Panel Tree strategies, we regress each strategy's gross and net returns on a comprehensive set of published factor models and test whether the intercept α in the time-series regression $R_t = \alpha + \boldsymbol{\beta}' \boldsymbol{F_t} + e_t$ is significantly different from zero. A failure to reject a zero alpha is consistent with the factor set spanning the strategy's average return, whereas a significantly positive alpha indicates incremental reward that the factors cannot explain.

Table 5 reports the monthly intercepts (%) from time-series regressions of the cost-ignorant (CI) and cost-aware (CA) strategy returns on six factor models. The specifications are CAPM (Sharpe, 1964), FF3 (Fama and French, 1992), FF5 (Fama and French, 2015), FF5 augmented with momentum, short-term reversal, and liquidity (Pástor and Stambaugh, 2003), Q5 (Hou et al., 2015), and ALL. The ALL specification combines the augmented FF5 factors with the Q5 factors, for 12 factors in total. Panel A reports results under the maximum-SR objective, and Panel B reports results under MVU with $\delta = 2$. Within each panel, the gross columns report the factor-adjusted returns before transaction costs, whereas the net columns show the factor-adjusted returns after transaction costs. All t-statistics use Newey–West standard errors (Newey and West, 1987).

Panel A reveals a sharp gross-to-net contrast under the maximum-SR objective. Before costs, both strategies earn significantly positive alphas in every specification. The CI gross alpha ranges from 0.39% to 0.60% per month, while the CA gross alpha ranges from 0.60% to 0.77%. After realized transaction costs, however, the CI intercept ranges from -0.16% to 0.04%. It is close to zero under CAPM and FF3, turns negative as more sophisticated factors are added, and is significantly negative at the 5% level under Q5 and ALL. By contrast, the CA net alpha remains positive and statistically significant at the 1% level among these specifications. It declines from 0.43% $(t = 5.01)$ under CAPM to

0.26% ($t = 3.02$) under the ALL specification, indicating that the broader factor set cannot absorb all of the cost-aware strategy's net return. Thus, a high gross alpha does not by itself imply implementable factor-adjusted performance: only the CA strategy retains a reliably positive intercept after transaction costs.

Panel B shows a similar contrast under MVU ($\delta = 2$). Before costs, the CI and CA alphas range from 2.02% to 3.69% and from 2.53% to 3.51%, respectively, and all are statistically significant at the 1% level. After costs, the CI alpha ranges from -0.17% to 1.38%. It remains significantly positive under CAPM and FF3 but becomes statistically insignificant from zero once more factors are included. The CA net alpha, in contrast, remains between 1.17% and 2.15%. For example, it equals 1.98% ($t = 3.59$) under FF5, 1.69% ($t = 2.98$) under the augmented FF5 model, 1.17% ($t = 2.02$) under Q5, and 1.30% ($t = 2.35$) under the ALL specification.

Across the six benchmark factor models, the CA net alpha remains positive and statistically significant under both portfolio objectives. These models include market, size, value, profitability, investment, momentum, short-term reversal, liquidity, return on equity, and expected growth factors. The CA strategy's *annualized* net alphas range from 3.1% to 5.2% under the maximum-SR objective and from 14.0% to 25.8% under MVU. The CI net alphas become negative under richer models in Panel A and remain significantly positive only under CAPM and FF3 in Panel B.

These findings add an implementation dimension to Cong et al. (2025), who show that Panel Tree test assets expand the mean–variance frontier and generate returns unexplained by benchmark pricing models. Before costs, both the CI and CA strategies display the factor-adjusted performance associated with Panel Tree basis-portfolios. Cost-aware training preserves positive factor-adjusted performance after transaction costs across the

six benchmark models.

Cross-Strategy Spanning Tests. The factor model tests treat observable factors as the benchmark. Next, we ask whether the cost-aware and cost-ignorant strategies span each other. If the two Panel Trees simply recover the same underlying payoff space, with cost awareness only reducing turnover, then each strategy's return should be spanned by the other, and the cross-strategy alphas should be zero. By contrast, if the basis-portfolios generated by the cost-aware Panel Tree expand the mean–variance frontier of the cost-ignorant Panel Tree, the spanning pattern will be asymmetric.

To formalize this test, we construct four strategy return series: Cost-Ignorant Gross, Cost-Ignorant Net, Cost-Aware Gross, and Cost-Aware Net. Table 6 reports cross-strategy spanning regressions across these four variants. Each cell corresponds to a time-series regression $Y_t = \alpha + \beta X_t + e_t$, where the row identifies the test asset Y_t and the column identifies the spanning factor X_t. The failure to reject the null hypothesis of zero α is consistent with the column strategy explaining the expected returns of the row strategy, whereas a significant α rejects the null.

The results have three implications. First, the gross-versus-net comparison of the same strategy quantifies the implementation wedge attributable to transaction costs. Under maximum SR, the intercept is 0.57% ($t = 18.10$) for CI Net explaining CI Gross, and 0.36% ($t = 19.92$) for CA Net explaining CA Gross. Under MVU, the corresponding estimates are 2.32% ($t = 17.29$) and 1.41% ($t = 20.65$). The reverse regressions produce intercepts of similar magnitude with the opposite sign. This pattern is consistent with lower transaction costs for the cost-aware strategies.

Second, the two gross strategies fail to span one another. Under maximum SR, re-

gressing CI Gross on CA Gross yields an intercept of 0.37% ($t = 5.59$), while the reverse regression yields 0.48% ($t = 5.76$). Under MVU, the corresponding intercepts are 2.14% ($t = 3.39$) and 1.94% ($t = 3.19$). The bidirectional rejections indicate that cost awareness changes the composition of the underlying basis-portfolios. This finding corroborates the structural evidence in Section 3.2: the cost-aware and cost-ignorant Panel Trees select different characteristics and partition the cross section differently, leading to distinct payoff spaces even before transaction costs are deducted.

Third, the net-space comparison reveals an asymmetry. Under maximum SR, CI Net regressed on CA Net produces an intercept of -0.09% ($t = -1.49$), whereas CA Net regressed on CI Net produces 0.42% ($t = 5.41$). Under MVU, the corresponding estimates are 0.49% ($t = 1.00$) and 1.67% ($t = 3.71$). Thus, when CI is treated as a return factor, CA earns a positive and precisely estimated alpha under both objectives; when CA is treated as a return factor, the CI intercept is insignificant from zero. It implies that the CA strategy contains incremental pricing information that is not spanned by the CI strategy.

Summary. Across the six benchmark factor models, the CI net alpha is not robust: it becomes negative under sophisticated factor models for maximum SR and remains significantly positive only under CAPM and FF3 for MVU. By contrast, the CA net alpha remains positive and statistically significant across all six models under both portfolio objectives. The cross-strategy regressions show a similar pattern: the CA net return earns a positive intercept when the CI net return serves as the explanatory factor, while the reverse intercept is statistically insignificant from zero. Together, these tests indicate that the CA strategy contains incremental pricing information not captured by either the benchmark factor models or the CI strategy returns.

4 Conclusion

This paper develops a cost-aware Panel Tree model that learns tradable basis-portfolios from high-dimensional and unbalanced panel data. By allowing individual asset-level transaction costs to enter the economic objective, the framework integrates the two-step basis-portfolio generation and allocation problem into a unified implementable portfolio selection problem.

We establish three theoretical results under the mild conditions. First, characteristic-managed basis-portfolios approach the optimal investment utility of individual assets, as characteristic groups become finer. Second, the Panel Tree approaches the investment utility of optimal basis-portfolios as the tree grows. Third, we characterize the population value of cost-aware basis-portfolio redesign after allocation has already adjusted for costs, and identify conditions under which this value is strictly positive.

Using U.S. equities from 1974 to 2023 and evaluating performance out of sample from 1984 to 2023, we find that the cost-aware Panel Tree strategy lowers transaction costs and turnover, and improves net performance under both the maximum Sharpe-ratio and mean–variance objectives. The generation–allocation decomposition shows that basis-portfolio generation accounts for more than 90% of the maximum-SR improvement and essentially all of the MVU improvement, whereas cost-aware allocation over the original cost-ignorant basis-portfolio set recovers a substantially smaller share. The cost-aware Panel Tree also selectively reallocates split capacity toward accounting-fundamental signals while retaining price-volume signals that survive the cost-aware criterion, producing more persistent leaf membership and lower stock-level turnover. The cost-aware strategy retains positive and statistically significant net alphas across six benchmark factor models, whereas the cost-ignorant strategy's net alpha is not robust to sophisticated factor

specifications. Cross-strategy spanning tests likewise show that the cost-aware net return contains incremental pricing information not spanned by its cost-ignorant counterpart.

We suggest that future research study how cost-aware basis-portfolio generation interacts with broader investor objectives and market structure. For example, how do the learned basis-portfolios change with investor size and price-impact costs, and how does the set of economically relevant characteristics evolve when multiple arbitrageurs compete over the same implementable opportunity set? Beyond U.S. equities, the framework extends naturally to international markets, bonds, and other asset classes in which implementability disciplines which characteristic-based signals are worth trading.

References

Abarbanell, J. S. and B. J. Bushee (1998). Abnormal returns to a fundamental analysis strategy. *The Accounting Review*, 19–45.

Ahn, D.-H., J. Conrad, and R. F. Dittmar (2009). Basis assets. *Review of Financial Studies 22*(12), 5133–5174.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*(1), 31–56.

Ao, M., L. Yingying, and X. Zheng (2019). Approaching mean-variance efficiency for large portfolios. *Review of Financial Studies 32*(7), 2890–2919.

Avramov, D., S. Cheng, and L. Metzker (2023). Machine learning vs. economic restrictions: Evidence from stock return predictability. *Management Science 69*(5), 2587–2619.

Babier, A., T. C. Y. Chan, A. Diamant, and R. Mahmood (2025). Learning to optimize contextually constrained problems for real-time decision generation. *Management Science 71*(2), 1165–1186.

Bali, T. G., N. Cakici, and R. F. Whitelaw (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics 99*(2), 427–446.

Bertsimas, D. and R. Cory-Wright (2022). A scalable algorithm for sparse portfolio selection. *INFORMS Journal on Computing 34*(3), 1489–1511.

Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies 22*(9), 3411–3447.

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance 54*(2), 655–671.

Bryzgalova, S., M. Pelger, and J. Zhu (2025). Forest through the trees: Building cross-sections of stock returns. *The Journal of Finance 80*(5), 2447–2506.

Chen, L., M. Sim, X. Zhang, L. Zhao, and M. Zhou (2026). Robust actionable prescriptive analytics. *Operations Research 74*(1), 550–571.

Cochrane, J. H. (2009). *Asset pricing: Revised edition*. Princeton University Press.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance 66*(4), 1047–1108.

Cong, L. W., G. Feng, J. He, and X. He (2025). Growing the efficient frontier on panel trees. *Journal of Financial Economics 167*, 104024.

Daniel, K., L. Mota, S. Rottke, and T. Santos (2020, 05). The cross-section of risk and returns. *The Review of Financial Studies 33*(5), 1927–1979.

Daniel, K. and S. Titman (2006). Market reactions to tangible and intangible information. *The Journal of Finance 61*(4), 1605–1643.

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science 55*(5), 798–812.

DeMiguel, V., L. Garlappi, and R. Uppal (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies 22*(5), 1915–1953.

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal (2020). A transaction-cost perspective on the multitude of firm characteristics. *Review of Financial Studies 33*(5), 2180–2222.

DeMiguel, V., A. Martín-Utrera, and R. Uppal (2024). A multifactor perspective on volatility-managed portfolios. *The Journal of Finance 79*(6), 3859–3891.

DeMiguel, V. and F. J. Nogales (2009). Portfolio selection with robust estimation. *Operations Research 57*(3), 560–577.

Ding, Y., Y. Li, and X. Zheng (2021). High dimensional minimum variance portfolio estimation under statistical factor models. *Journal of Econometrics 222*(1), 502–515.

Elmachtoub, A. N. and P. Grigas (2022). Smart "predict, then optimize". *Management Science 68*(1), 9–26.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *The Journal of Finance 47*(2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Fan, Q., M. C. Medeiros, H. Yang, and S. Yang (2024). Cost-aware portfolios in a large universe of assets. *arXiv preprint arXiv:2412.11575*.

Gao, J. and D. Li (2013). Optimal cardinality constrained portfolio selection. *Operations Research 61*(3), 745–761.

Giglio, S., D. Xiu, and D. Zhang (2025). Test assets and weak factors. *The Journal of Finance 80*(1),

259–319.

Heston, S. L. and R. Sadka (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics 87*(2), 418–445.

Hou, K., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment approach. *Review of Financial Studies 28*(3), 650–705.

Jacobs, B. I., K. N. Levy, and H. M. Markowitz (2005). Portfolio optimization with factors, scenarios, and realistic short positions. *Operations Research 53*(4), 586–599.

Jensen, T. I., B. Kelly, S. Malamud, and L. H. Pedersen (2026). Machine learning and the implementable efficient frontier. *Review of Financial Studies*, hhag022.

Jensen, T. I., B. Kelly, and L. H. Pedersen (2023). Is there a replication crisis in finance? *The Journal of Finance 78*(5), 2465–2518.

Ledoit, O. and M. Wolf (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis 88*(2), 365–411.

Ledoit, O. and M. Wolf (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *Review of Financial Studies 30*(12), 4349–4388.

Ledoit, O. and M. Wolf (2020). Analytical nonlinear shrinkage of large-dimensional covariance matrices. *The Annals of Statistics 48*(5), 3043–3065.

Lewellen, J., S. Nagel, and J. Shanken (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics 96*(2), 175–194.

Li, F. (2011). Earnings quality based on corporate investment decisions. *Journal of Accounting Research 49*(3), 721–752.

Li, S., V. DeMiguel, and A. Martin-Utrera (2024). Comparing factor models with price-impact costs. *Journal of Financial Economics 162*, 103949.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Moreira Costa, C., D. Kreber, and M. Schmidt (2022). An alternating method for cardinality-constrained optimization: A computational study for the best subset selection and sparse portfolio problems. *INFORMS Journal on Computing 34*(6), 2968–2988.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Novy-Marx, R. and M. Velikov (2016). A taxonomy of anomalies and their trading costs. *Review of Financial Studies 29*(1), 104–147.

Olivares-Nadal, A. V. and V. DeMiguel (2018). A robust perspective on transaction costs in portfolio optimization. *Operations Research 66*(3), 733–739.

Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*(3), 425–442.

Sharpe, W. F. (1971). A linear programming approximation for the general portfolio analysis prob-

lem. *Journal of Financial and Quantitative Analysis 6*(5), 1263–1275.

Simon, F., S. Weibels, and T. Zimmermann (2026). Deep parametric portfolio policies. *Management Science, fothcoming*.

Xu, W., J. Tang, K. F. C. Yiu, and J. W. Peng (2024). An efficient global optimal method for cardinality constrained portfolio optimization. *INFORMS Journal on Computing 36*(2), 690–704.

Figure 1: **Cost-Ignorant Panel Tree Diagram**

Notes. This figure reports the cost-ignorant Panel Tree diagram with the objective of maximizing the Sharpe ratio and the restriction of no more than 10 leaves. The model is estimated using the latest rolling window, from January 2013 to December 2022. The figure shows the selected characteristics, the cutoff values for each parent node, and the associated node labels and split order. Nodes are labeled N# (index) and S# (split order). Terminal leaf nodes display the median monthly number of assets (#) and portfolio weights (w).

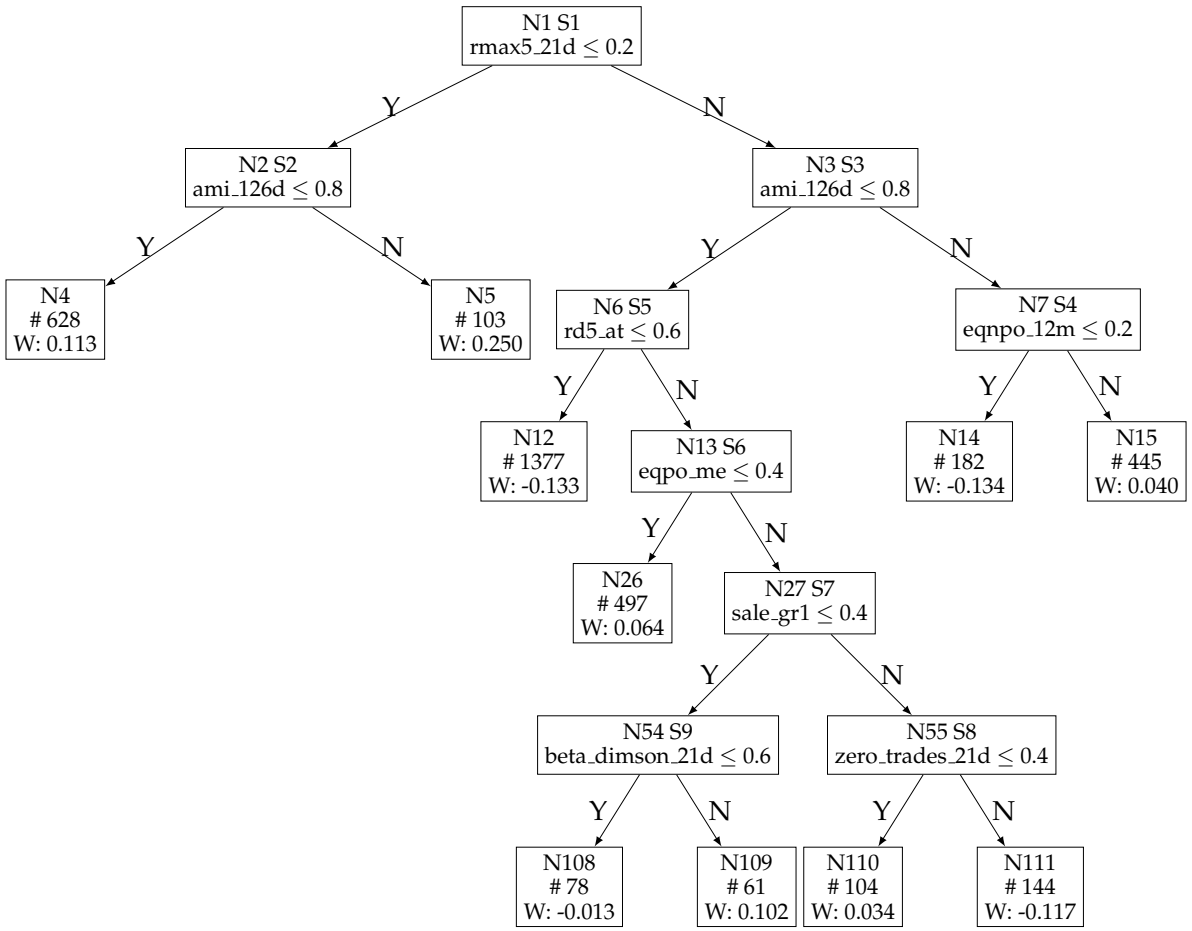


Figure 2: **Cost-Aware Panel Tree Diagram**

Notes. This figure reports the cost-aware Panel Tree diagram for the cost-aware Sharpe-ratio objective and the restriction of no more than 10 leaves. The model is estimated using the latest rolling window, from January 2013 to December 2022. The figure shows the selected characteristics, the cutoff values for each parent node, and the associated node labels and split order. Nodes are labeled N# (index) and S# (split order). Terminal leaf nodes display the median monthly number of assets (#) and portfolio weights (w).

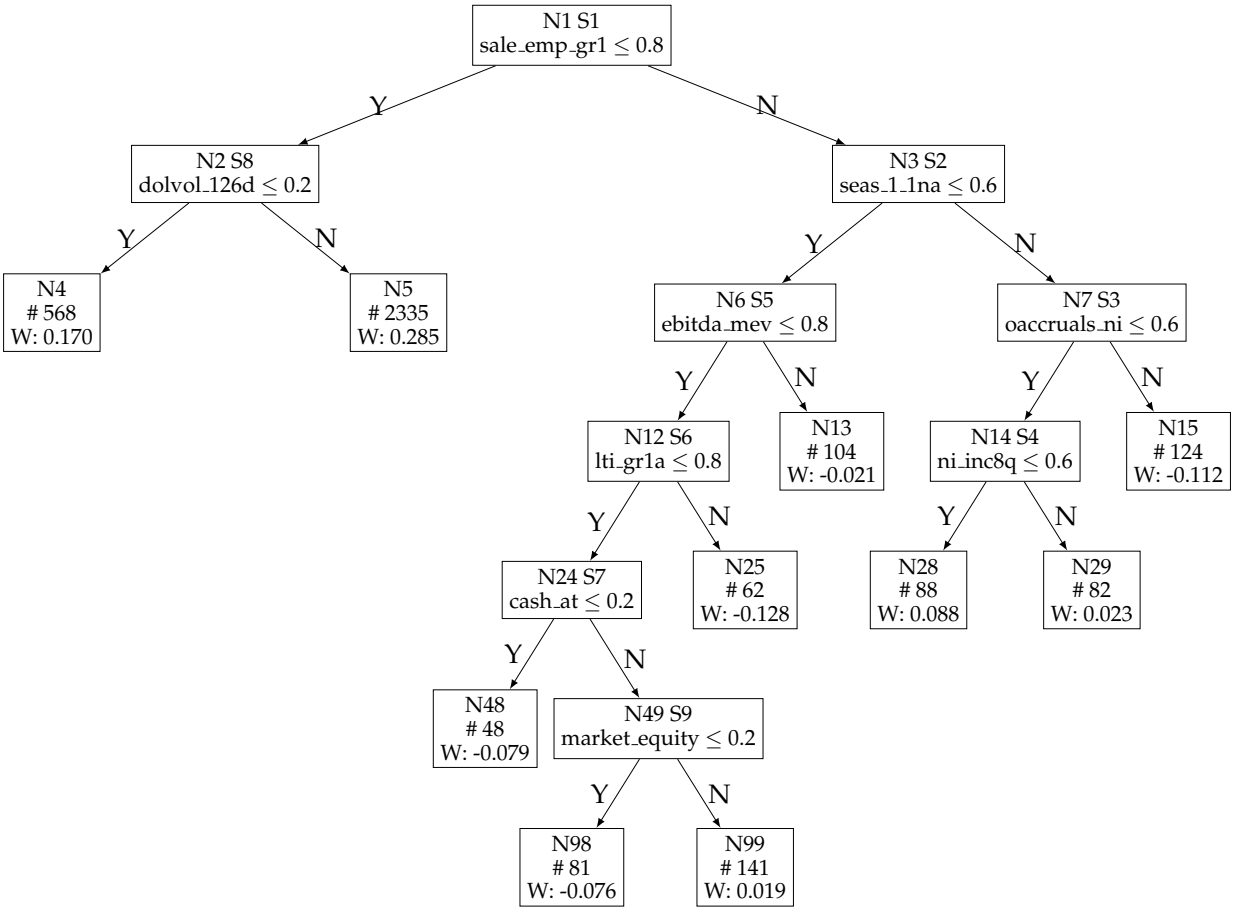


Figure 3: **Average Split Frequency by Characteristic Category: Cost-aware versus Cost-ignorant Panel Trees**

Notes. This figure reports the average proportion of total splits assigned to each characteristic category under the cost-aware (CA) and cost-ignorant (CI) Panel Trees. The averages are computed across 40 rolling windows, and each tree contains nine splits per window. The figure provides an intuitive category-level view of how fixed split capacity is reallocated once transaction costs are internalized.

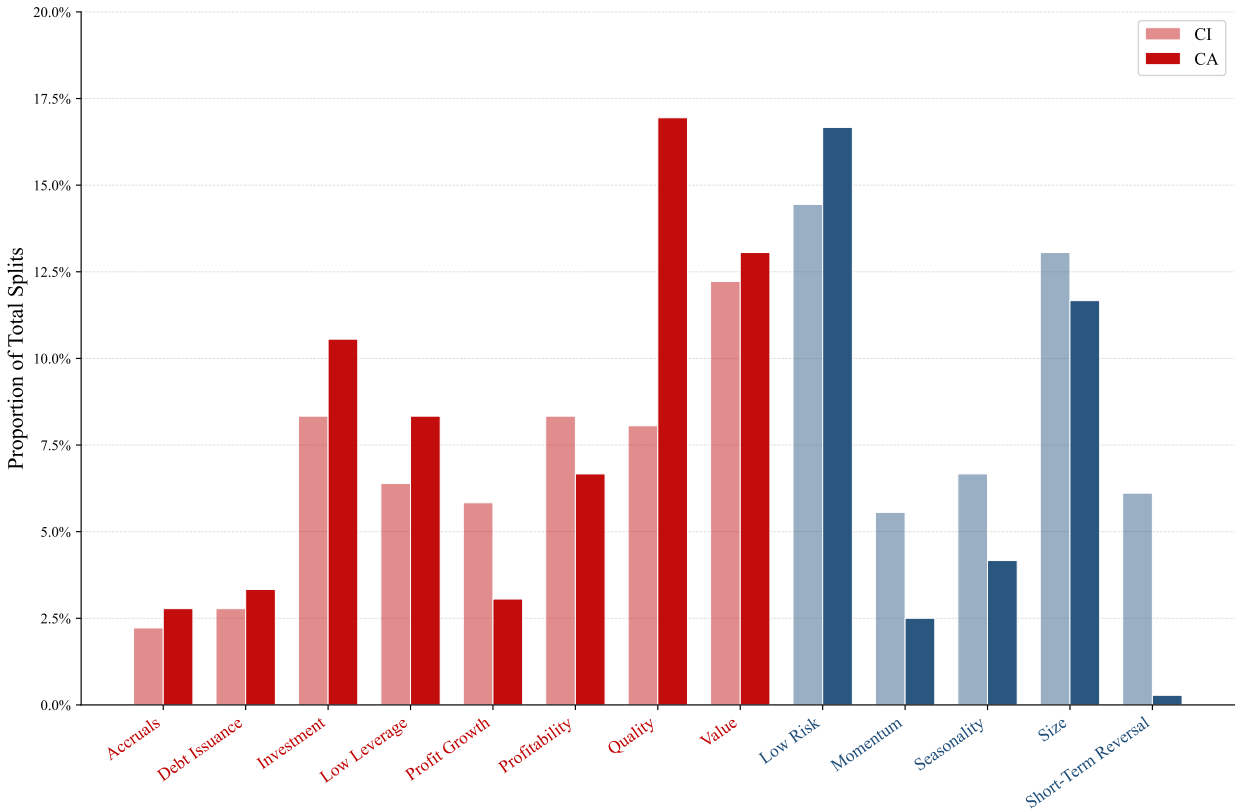


Figure 4: **Difference in Score-weighted Split Importance across Rolling Windows: CA minus CI**

Notes. This figure reports, for each rolling window and characteristic category, the difference in split scores between the cost-aware and cost-ignorant Panel Trees. Within each rolling window, the split score is 9 for the first split, 8 for the second, and so on down to 1 for the ninth split. Positive values indicate that the category becomes more important under cost-aware training. Because score is a split-order-weighted measure of split importance, the figure captures not only whether a category is selected but also whether it is used higher in the tree.

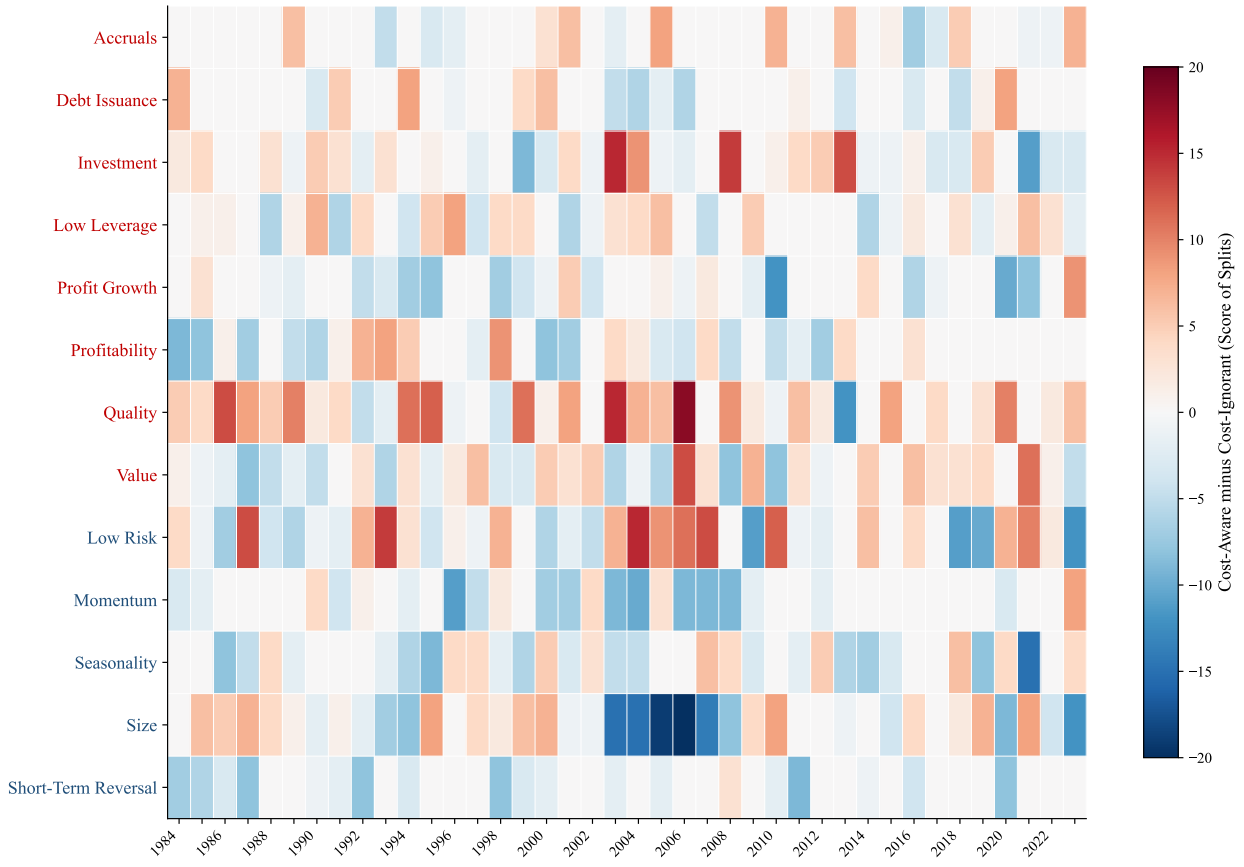
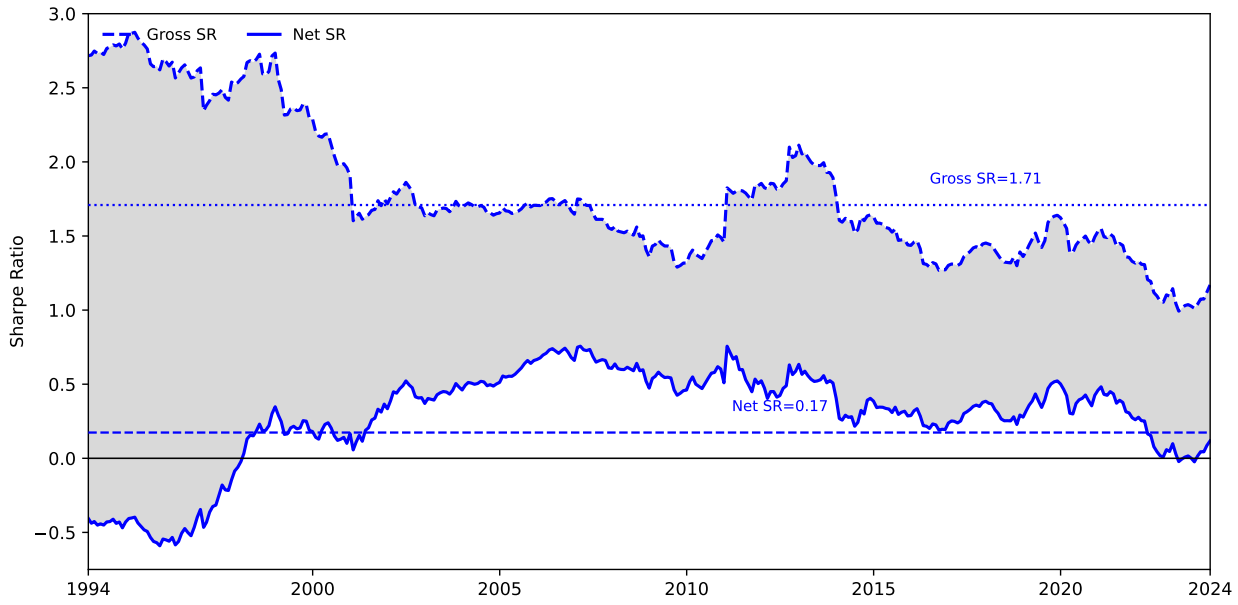
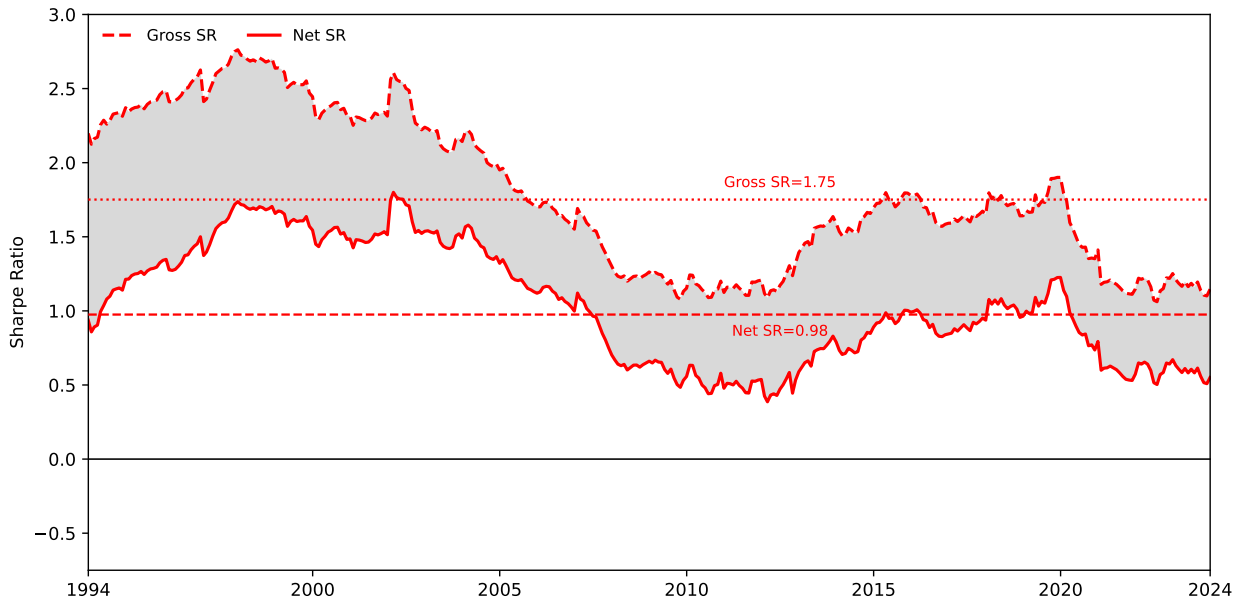


Figure 5: **Rolling Sharpe Ratio Curve Comparison**

Notes. This figure shows the out-of-sample Sharpe ratio of strategies with and without transaction costs. Panel (a) presents the results of cost-ignorant Panel Trees, and Panel (b) illustrates cost-aware Panel Trees. In each panel, the dash-dotted line plots the 10-year rolling gross Sharpe ratio and the solid line plots the corresponding 10-year rolling net Sharpe ratio. The horizontal dotted and dashed lines indicate the OOS average sample gross and net Sharpe ratios, respectively. The gray shaded area highlights the difference between the gross and net rolling Sharpe ratios, illustrating the impact of transaction costs. The models are updated annually with a 10-year rolling window.



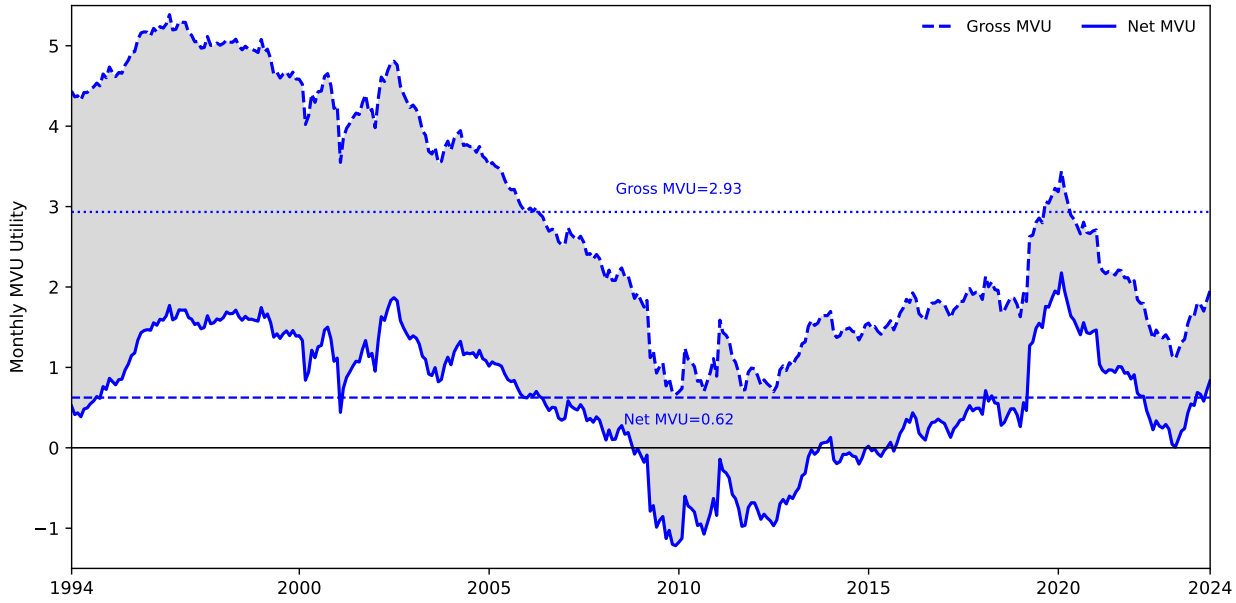
(a) Cost-Ignorant Panel Trees



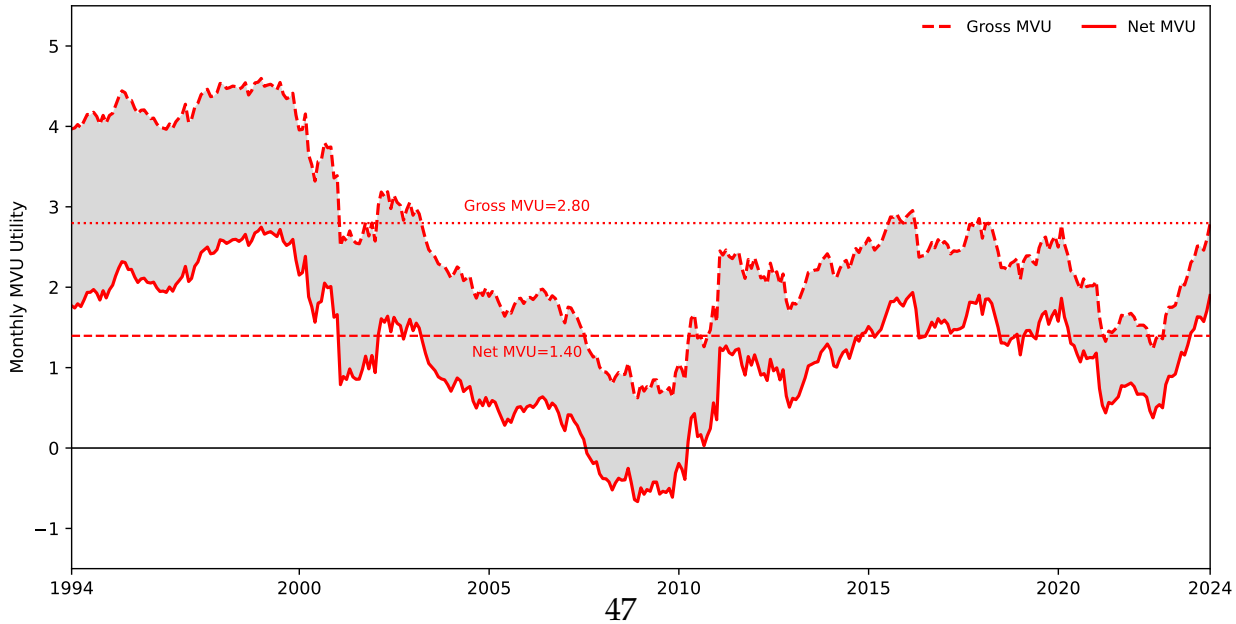
(b) Cost-Aware Panel Trees

Figure 6: **Mean–Variance Utility Curve Comparison**

Notes. This figure shows the out-of-sample mean–variance utility of strategies with and without transaction costs. Panel (a) presents the results of cost-ignorant Panel Trees, and Panel (b) illustrates cost-aware Panel Trees. In each panel, the dash-dotted line plots the 10-year rolling gross mean–variance utility and the solid line plots the corresponding 10-year rolling net mean–variance utility. The horizontal dotted and dashed lines indicate the OOS average gross and net mean–variance utility, respectively. The gray shaded area highlights the difference between the gross and net mean–variance utility, illustrating the impact of transaction costs. The models are updated annually with a 10-year rolling window.



(a) Cost-ignorant Panel Trees



(b) Cost-aware Panel Trees

Figure 7: **Cumulative Net Excess Returns**

Notes. This figure reports the out-of-sample cumulative net excess returns of two strategies. The blue line is for the cost-ignorant (CI) Panel Tree, while the red line is for the cost-aware (CA) Panel Tree. The out-of-sample period spans from 1984 to 2023. The gray shaded area highlights the difference in performance between the two strategies over time, illustrating the value added by internalizing transaction costs into the objective function.

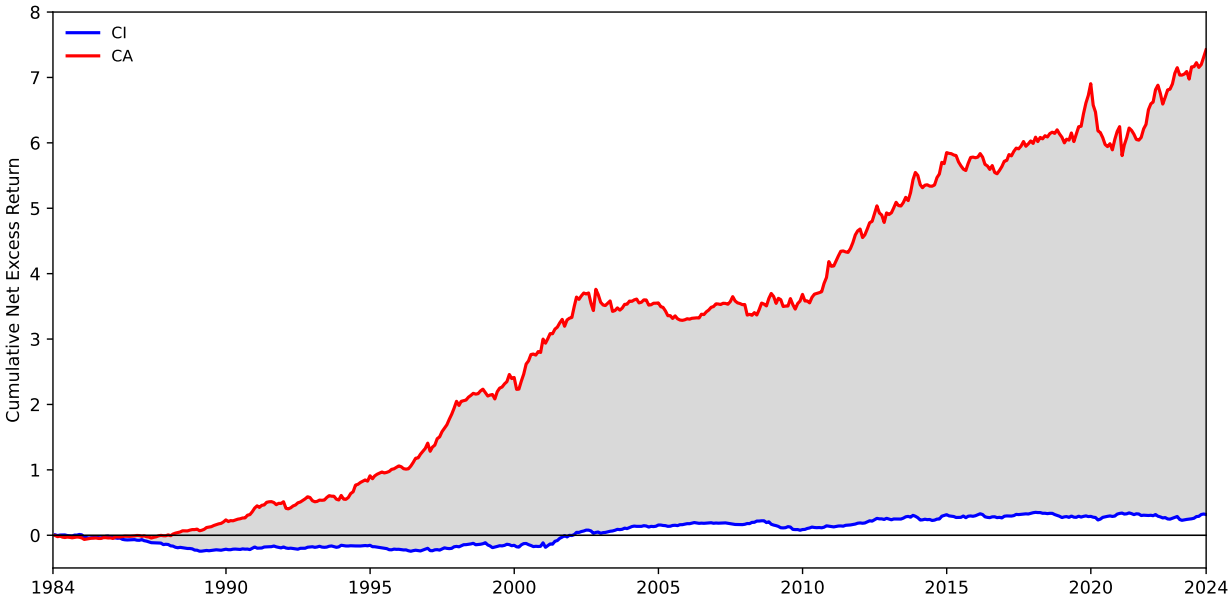


Table 1: **Characteristics Selected by Panel Trees**

Notes. This table reports the firm characteristics selected by the Panel Tree models. Panel A lists the characteristics used in the cost-ignorant Panel Tree (corresponding to Figure 1). Panel B details the characteristics selected by the cost-aware Panel Tree (corresponding to Figure 2). For each characteristic, the table reports its name, a brief description, and its category. More information on these characteristics can be found in Jensen et al. (2023).

No.	Characteristic Name	Description	Category
		Panel A: Characteristics in **Cost-Ignorant** Panel Tree	
1	rmax5_21d	Highest 5 days of return	Low Risk
2	ami_126d	Amihud Measure	Size
3	ami_126d	Amihud Measure	Size
4	eqnpo_12m	Equity net payout	Value
5	rd5_at	R&D capital-to-book assets	Low Leverage
6	eqpo_me	Payout yield	Value
7	sale_gr1	Sales Growth (1 year)	Investment
8	zero_trades_21d	Number of zero trades with turnover as tiebreaker (1 month)	Low Risk
9	beta_dimson_21d	Dimson beta	Low Risk
		Panel B: Characteristics in **Cost-Aware** Panel Tree	
1	sale_emp_gr1	Labor force efficiency	Profit Growth
2	seas_1_1na	Year 1-lagged return, nonannual	Momentum
3	oaccruals_ni	Percent operating accruals	Accruals
4	ni_inc8q	Number of consecutive quarters with earnings increases	Quality
5	ebitda_mev	EBITDA-to-market enterprise value	Value
6	lti_gr1a	Change in long-term investments	Seasonality
7	cash_at	Cash-to-assets	Low Leverage
8	dolvol_126d	Dollar trading volume	Size
9	market_equity	Market Equity	Size

Table 2: **Reallocation of Split Capacity across Rolling Windows**

Notes. This table summarizes how the nine splits per tree are allocated across characteristic categories under the cost-ignorant (CI) and cost-aware (CA) Panel Trees, averaged over 40 rolling windows. Panel A aggregates splits into two meta-categories: accounting-fundamental and price-volume. Split share is the average split frequency divided by nine. Within each rolling window, the split score is 9 for the first split, 8 for the second, and so on down to 1 for the ninth split; larger values indicate that a category is used more prominently in the tree, organizing a larger share of the cross section. Panel B reports the largest category-level changes (CA minus CI) in split frequency and score.

Panel A: **Aggregate Split Allocation by Meta-category**

	Avg. frequency		Frequency share (%)		Avg. score		Score share (%)	
Meta-category	CI	CA	CI	CA	CI	CA	CI	CA
Accounting-fundamental	4.88	5.83	54.17	64.72	22.90	27.78	50.89	61.72
Price-volume	4.12	3.18	45.83	35.28	22.10	17.23	49.11	38.28

Panel B: **Largest Category-level Changes (CA minus CI)**

Largest increases	Δ Freq.	Δ Score	Largest decreases	Δ Freq.	Δ Score
Quality	+0.80	+4.18	Short-Term Reversal	-0.53	-1.90
Low Risk	+0.20	+1.48	Momentum	-0.28	-1.80
Investment	+0.20	+1.15	Profit Growth	-0.25	-1.45
Accruals	+0.05	+0.63	Size	-0.13	-1.45

Table 3: **Out-of-Sample Performance under Proportional Transaction Costs**

Notes. This table reports out-of-sample performance under proportional transaction costs for the cost-ignorant (CI) and cost-aware (CA) Panel Trees and two classical benchmarks. Panel A reports results under the *maximum-Sharpe Ratio* objective, Panel B reports results under *Mean–Variance Utility* with $\delta = 2$, and Panel C reports results for benchmarks. EW denotes the equal-weighted market portfolio, and VW denotes the value-weighted market portfolio. Return, volatility, transaction cost (TC), and turnover are monthly and expressed in percent. The gross and net objective values in Panel C are annualized Sharpe ratios. For readability, the reported values for *Mean–Variance Utility* are multiplied by 100.

Strategy	Gross return	Net return	Gross vol.	Net vol.	Gross obj.	**Net obj.**	TC	Turnover
Panel A: **Maximum Sharpe Ratio**								
CI	0.63	0.07	1.28	1.32	1.71	**0.17**	0.56	85.87
CA	0.80	0.46	1.57	1.63	1.75	**0.98**	0.34	53.16
Panel B: **Mean–Variance Utility**								
CI	4.00	1.69	10.34	10.33	2.93	**0.62**	2.31	351.92
CA	3.89	2.51	10.47	10.58	2.80	**1.40**	1.38	220.86
Panel C: **Benchmarks**								
EW	0.44	0.36	5.79	5.79	0.26	**0.22**	0.08	11.42
VW	0.72	0.71	4.52	4.52	0.55	**0.55**	0.01	1.21

Table 4: **Basis-Portfolio Generation versus Allocation**

Notes. This table reports the 2×2 comparison of cost-ignorant (CI) and cost-aware (CA) Panel Tree basis-portfolio generation and allocation rules. The first element of each strategy label identifies the basis-portfolio generation rule, and the second identifies the allocation rule. All specifications use the same 480 out-of-sample months from January 1984 through December 2023 and are evaluated using the same transaction-cost schedule. Net objective is the annualized net Sharpe ratio in Panel A and net mean–variance utility multiplied by 100 in Panel B. Gain from CI–CI is the difference between each net objective and the CI–CI net objective within the same panel. Total gain recovered divides this difference by the CA–CA-minus-CI–CI difference and is calculated using unrounded objective values. The two hybrid percentages are separate counterfactual comparisons and need not sum to 100. Following DeMiguel et al. (2024), the final column reports bootstrap-based one-sided p-values for the null that CA–CA does not deliver a higher net objective than the strategy in the corresponding row. We jointly resample the monthly net returns of all four strategies 10,000 times using a stationary block bootstrap with an average block length of five months and recompute the relevant objective in every bootstrap sample.

Strategy	Net objective	Gain from CI–CI	Total gain recovered (%)	p-value
		Panel A: **Maximum Sharpe Ratio**		
CI–CI	0.17	–	–	< 0.001
CI–CA	0.43	0.26	32.26	0.003
CA–CI	0.91	0.74	91.77	0.003
CA–CA	0.98	0.80	100.00	–
		Panel B: **Mean–Variance Utility**		
CI–CI	0.62	–	–	0.058
CI–CA	0.83	0.20	26.13	0.126
CA–CI	1.40	0.78	100.87	0.620
CA–CA	1.40	0.77	100.00	–

Table 5: **Factor-Model Regressions for Panel Tree Returns**

Notes. This table reports the monthly alphas (%) from time-series regressions of gross and net Panel Tree strategy returns on benchmark asset pricing factors. CI and CA denote cost-ignorant and cost-aware Panel Trees. Gross and Net denote gross and net strategy returns. Panel A reports results under the maximum Sharpe ratio, and Panel B reports results under mean–variance utility. In each panel, rows 1–5 use, respectively, CAPM; the Fama–French three-factor model (FF3); the Fama–French five-factor model (FF5); FF5 augmented with momentum, short-term reversal, and liquidity; Q5, which contains the q-factor market, size, investment, return-on-equity, and expected-growth factors. The last row ALL includes all 12 factors. Newey–West t-statistics are in parentheses. All specifications use 480 months from January 1984 through December 2023. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Factor model	CI Gross	CA Gross	CI Net	CA Net
Panel A: **Maximum Sharpe Ratio**				
CAPM	0.60***	0.77***	0.04	0.43***
	(9.31)	(8.71)	(0.62)	(5.01)
FF3	0.58***	0.74***	0.02	0.41***
	(9.55)	(8.67)	(0.26)	(4.87)
FF5	0.48***	0.65***	-0.08	0.31***
	(8.32)	(7.91)	(-1.37)	(3.87)
FF5 Augmented	0.45***	0.64***	-0.12*	0.29***
	(7.72)	(7.60)	(-1.95)	(3.61)
Q5	0.39***	0.60***	-0.16**	0.26***
	(5.98)	(6.99)	(-2.31)	(3.14)
ALL	0.39***	0.60***	-0.15**	0.26***
	(6.43)	(6.64)	(-2.40)	(3.02)
Panel B: **Mean–Variance Utility ($\delta = 2$)**				
CAPM	3.69***	3.44***	1.38***	2.07***
	(7.46)	(7.70)	(2.89)	(4.59)
FF3	3.61***	3.51***	1.31***	2.15***
	(7.47)	(7.44)	(2.77)	(4.51)
FF5	2.91***	3.37***	0.60	1.98***
	(6.05)	(6.15)	(1.27)	(3.59)
FF5 Augmented	2.47***	3.08***	0.17	1.69***
	(5.05)	(5.46)	(0.35)	(2.98)
Q5	2.33***	2.53***	0.12	1.17**
	(4.76)	(4.35)	(0.25)	(2.02)
ALL	2.02***	2.65***	-0.17	1.30**
	(4.19)	(4.74)	(-0.36)	(2.35)

Table 6: **Cross-Strategy Spanning Tests**

Notes. This table reports pairwise spanning regressions across four strategy return series: gross and net returns of the cost-ignorant Panel Tree (CI Gross and CI Net) and gross and net returns of the cost-aware Panel Tree (CA Gross and CA Net). Panel A reports results under the maximum Sharpe ratio, and Panel B reports results under mean–variance utility. In each regression, the row strategy is the dependent variable and the column strategy is the explanatory return factor. Each off-diagonal entry reports the monthly intercept (in percent), followed by its Newey–West t-statistic in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Explanatory Strategy (factor)			
Dependent strategy	CI Gross	CI Net	CA Gross	CA Net
		Panel A: **Maximum Sharpe Ratio**		
CI Gross	–	0.57***	0.37***	0.48***
		(18.10)	(5.59)	(8.15)
CI Net	-0.56***	–	-0.20***	-0.09
	(-16.70)		(-2.81)	(-1.49)
CA Gross	0.48***	0.76***	–	0.36***
	(5.76)	(9.38)		(19.92)
CA Net	0.13	0.42***	-0.35***	–
	(1.59)	(5.41)	(-18.44)	
		Panel B: **Mean–Variance Utility**		
CI Gross	–	2.32***	2.14***	2.81***
		(17.29)	(3.39)	(5.53)
CI Net	-2.27***	–	-0.16	0.49
	(-17.27)		(-0.25)	(1.00)
CA Gross	1.94***	3.07***	–	1.41***
	(3.19)	(6.84)		(20.65)
CA Net	0.53	1.67***	-1.40***	–
	(0.87)	(3.71)	(-20.62)	

Algorithm 1 Cost-aware Panel Tree growth

Require: Update date T, training window $\mathcal{W}_T$, candidate rule $\mathcal{C}(\cdot)$, split budget S, and objective τ

Initialize the root tree Θ_0, solve for $\boldsymbol{w}^*_{\tau,T}(\Theta_0)$, and set $s \leftarrow 1$
If Θ_0 is infeasible, **stop**
Evaluate $\mathcal{U}_{\tau,T}(\Theta_0)$
while $s \leq S$ **do**
    Construct the admissible split set $\mathcal{C}(\Theta_{s-1})$
    If $\mathcal{C}(\Theta_{s-1}) = \varnothing$, **break**
    for all $\xi \in \mathcal{C}(\Theta_{s-1})$ **do**
        Form $\Theta' = \Theta_{s-1} \oplus \xi$, construct its leaf returns, and estimate their moments
        Compute $\boldsymbol{w}^{\mathrm{CI},*}_{\tau,T}(\Theta')$, fix the denominators in $\widetilde{h}^{+}_{i,t-1}$, and construct $\widehat{C}_T(\boldsymbol{w};\Theta')$
        Solve for $\boldsymbol{w}^*_{\tau,T}(\Theta')$ and evaluate its cost-aware gain, setting the gain to $-\infty$ if infeasible
    end for
    Choose $\xi^*_s \in \arg\max_{\xi\in\mathcal{C}(\Theta_{s-1})} \Delta\mathcal{U}_{\tau,T}(\xi;\Theta_{s-1})$
    If $\Delta\mathcal{U}_{\tau,T}(\xi^*_s;\Theta_{s-1}) \leq 0$, **break**
    $\Theta_s \leftarrow \Theta_{s-1} \oplus \xi^*_s$ and $s \leftarrow s+1$
end while
$s \leftarrow s-1$
Output: Θ_s and $\boldsymbol{w}^*_{\tau,T}(\Theta_s)$

Internet Appendix

(Not for Publication)

A Notation

A.1 Assets and Characteristics

$\mathcal{I}_t$ **:** Trading universe in month t.

N_t **:** The number of individual assets in month t.

$r_{i,t}$ **:** Excess return of individual asset i in month t.

$\boldsymbol{r}_t$ **:** The vector of asset excess returns ($N_t \times 1$).

$\boldsymbol{z}_{i,t-1}$ **:** The vector of K lagged, rank-standardized individual asset characteristics ($K \times 1$).

A.2 Basis-Portfolio Generation

g **:** General characteristic-based portfolio generation rule.

Θ **:** Characteristic-based portfolio generation rule defined by a Panel Tree.

$\mathcal{L}(\Theta)$ **:** Set of leaf basis-portfolios indexed by $j = 1, \ldots, m$.

m **:** The number of leaf basis-portfolios.

$\mathcal{I}_t^j$ **:** Set of individual assets in basis-portfolio j in month t.

N_t^j **:** The number of individual assets in basis-portfolio j in month t.

L **:** Maximum number of leaves in a Panel Tree.

$\mathbb{T}_L$ **:** Set of feasible Panel Trees with at most L leaves.

A.3 Portfolio Weights

$q_{i,t}$ **:** Market-portfolio benchmark weight of asset i in month t.

$\pi_{i,t}$ **:** Within-portfolio value weight of asset i in month t.

$\phi_{j,t}$ **:** Within-portfolio weight multiplier of basis-portfolio j in month t.

$\boldsymbol{w}, w_j$ **:** Vector of basis-portfolio allocation weights ($m \times 1$) and its j-th element.

$h_{i,t}$ **:** Portfolio weight of asset i in month t.

f^* **:** Weight-multiplier under the optimal individual asset-level strategy.

$f_{m,t}$ **:** Weight-multiplier by m basis-portfolios in month t.

A.4 Returns, Costs, and Allocation

$R_{j,t}, \mathbf{R}_t$ **:** Basis-portfolio excess return and vector of basis-portfolio excess returns ($m \times 1$).

$R_t(g, \boldsymbol{w})$ **:** Gross strategy excess return under generation rule g and allocation $\boldsymbol{w}$.

$C_t, \widetilde{C}_t$ **:** Exact and approximate transaction costs in month t, respectively.

$\widehat{C}_T$ **:** Sample average of $\widetilde{C}_t$.

$\widetilde{\Delta h}_{i,t}$ **:** Approximate trade in asset i in month t.

$\widehat{\boldsymbol{\mu}}_T$ **:** Sample mean vector ($m \times 1$) of basis-portfolio returns.

$\widehat{\boldsymbol{\Sigma}}_T$ **:** Covariance matrix with shrinkage ($m \times m$) of basis-portfolio returns.

δ **:** Mean–variance risk aversion parameter.

$\mathcal{A}_\tau$ **:** Feasible allocation set for objective $\tau \in \{\mathrm{MVU}, \mathrm{SR}\}$.

$\mathcal{U}_{\tau,T}(\Theta)$ **:** Optimal objective value for Panel Tree Θ under allocation objective τ.

A.5 Population Investment Opportunities

$\mathcal{U}_0(g, \boldsymbol{w})$ **:** Gross mean–variance utility under generation rule g and allocation $\boldsymbol{w}$.

$\mathcal{U}_0(g)$ **:** Optimal gross mean–variance utility under generation rule g.

$\mathcal{U}_0^{\text{asset},*}$ **:** Optimal utility from individual assets.

$\mathcal{U}_0^{\text{basis},*}$ **:** Optimal utility from basis-portfolios.

$\mathcal{U}_0^{\text{tree},*}(L)$ **:** Optimal utility over feasible iterative trees with at most L leaves.

ρ_m^2 **:** Minimum variance of the returns of the spread portfolio that longs the optimal individual asset-level strategy and shorts an m-basis-portfolio strategy.

A_L **:** Utility gap between the optimal basis-portfolio strategy and the optimal iterative-tree strategy with at most L leaves.

Γ_L **:** Utility gap between the optimal iterative-tree strategy and the Panel Tree strategy, both with at most L leaves.

A.6 Cost-Aware Basis Design and Comparisons

$\mathcal{A}(\Theta)$ **:** Feasible allocation set on tree Θ under mean–variance utility.

$\mathfrak{F}_L$ **:** Common domain of feasible pairs $(\Theta, \boldsymbol{w})$ with $\Theta \in \mathbb{T}_L$ and $\boldsymbol{w} \in \mathcal{A}(\Theta)$.

λ **:** Cost indicator, costs excluded at $\lambda = 0$ and included at $\lambda = 1$.

$\mathcal{U}_\lambda(\Theta)$ **:** Optimal cost-aware utility from tree Θ.

Θ^0, Θ^1 **:** Selected global CI- and CA-optimal generation rules, respectively.

$\boldsymbol{w}_\Theta^{\text{CI}}, \boldsymbol{w}_\Theta^{\text{CA}}$ **:** Gross-utility-optimal and cost-aware-optimal allocations on tree Θ.

$\mathcal{D}(\Theta)$ **:** Gross opportunity loss relative to the selected gross-optimal tree Θ^0.

$\mathcal{K}(\Theta)$ **:** Friction loss on tree Θ after allocation adjusts, including gross utility sacrificed and residual expected cost.

$\mathcal{N}(\Theta)$ **:** Common menu of the parent tree and its admissible one-split children.

$x, y \in \{0, 1\}$ **:** Dummy indices for generation and allocation, respectively, with 0 for CI and 1 for CA.

$\boldsymbol{w}^{x,y}$ **:** Optimal allocation under regime y on tree Θ^x.

J_{xy} **:** Population crossed value evaluated under the same cost-aware criterion in every cell.

B Theoretical Foundations of Basis-Portfolio Generation

This section presents the proofs for Theorems 1-3.

B.1 Proof of Theorem 1

B.1.1 Mean–Variance Utility Projection

Let R^* denote the optimal individual asset-level asset allocation return. For any feasible basis-portfolio generation rule g and allocation rule $\boldsymbol{w}$ with return R, we set $D = R - R^*$. For each R, assume there is an $\eta > 0$ such that $R^* + \varepsilon D$ can be generated by a feasible individual asset-level asset allocation rule for every $|\varepsilon| < \eta$. For each finite number m of characteristic-managed basis-portfolios, we assume that an optimal feasible generation rule and allocation exist.

Proof of the utility-loss identity in Theorem 1. For any asset return vector X, write $\mathcal{U}_0(X) := \mathbb{E}[X] - (\delta/2)\,\mathrm{Var}(X)$. Let R^* be the return under the optimal individual asset-level allocation and let R be the return generated by any basis-portfolio generation rule and allocation rule. Set $D = R - R^*$. By assumption, $R^* + \varepsilon D$ is feasible for all sufficiently small

positive and negative ε. Optimality therefore implies:

$$\frac{d}{d\varepsilon}\mathcal{U}_0(R^* + \varepsilon D)\Big|_{\varepsilon=0} = \mathbb{E}[D] - \delta\,\mathrm{Cov}(R^*, D) = 0. \tag{B.1}$$

Using $R = R^* + D$, the utility difference is

$$\begin{aligned}\mathcal{U}_0(R^*) - \mathcal{U}_0(R) &= \mathbb{E}[R^*] - \mathbb{E}[R] + \frac{\delta}{2}\big\{\mathrm{Var}(R) - \mathrm{Var}(R^*)\big\} \\ &= \underbrace{-\mathbb{E}[D] + \delta\,\mathrm{Cov}(R^*, D)}_{=0,\text{ according to (B.1)}} + \frac{\delta}{2}\,\mathrm{Var}(D) \\ &= \frac{\delta}{2}\,\mathrm{Var}(R^* - R).\end{aligned} \tag{B.2}$$

Minimizing over feasible basis-portfolio strategies gives

$$\mathcal{U}_0^{\text{asset},*} - \mathcal{U}_0^{\text{basis},*} = \frac{\delta}{2}\min_{g,\boldsymbol{w}}\mathrm{Var}[R^* - R_t(g, \boldsymbol{w})] = \frac{\delta}{2}\rho_m^2. \tag{B.3}$$

□

B.1.2 Characteristic-Function Approximation

To establish the convergence conclusion in Theorem 1, define the approximation error in portfolio weight for asset i in portfolio j as

$$e_{i,j,t} := f_{m,t}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w}) - f^*(\boldsymbol{z}_{i,t-1}) = w_j\phi_{j,t}(\boldsymbol{z}_{i,t-1}) - f^*(\boldsymbol{z}_{i,t-1}).$$

Then $h_{i,t} - h^*_{i,t} = q_{i,t} e_{i,j,t}$. Since an individual asset belongs to only one portfolio at each date, these errors enter the strategy return through

$$R_t(g, \boldsymbol{w}) - R_t^{\text{asset},*} = \sum_{j=1}^{m} \sum_{i=1}^{N_t^j} q_{i,t} e_{i,j,t} r_{i,t}.$$

For each m, let a basis-portfolio generation rule g partition the compact characteristic space into disjoint regions $C_1, \ldots, C_m$. We assume that there exists a feasible allocation $\boldsymbol{w} \in \mathcal{A}$, fixed across evaluation dates, such that the multiplier applied to each stock's benchmark weight lies within the range of optimal multipliers in the region. Specifically,

$$\inf_{\boldsymbol{z} \in C_j} f^*(\boldsymbol{z}) \leq w_j \phi_{j,t}(\boldsymbol{z}_{i,t-1}) \leq \sup_{\boldsymbol{z} \in C_j} f^*(\boldsymbol{z}), \tag{B.4}$$

for every $j = 1, \ldots, m$, $i = 1, \ldots, N_t^j$, and evaluation date t.

Lemma 1 (Local function approximation and return convergence)**.** *Suppose f^* is continuous on the compact characteristic space, $\max_{1 \leq j \leq m} \operatorname{diam}(C_j) \to 0$ as $m \to \infty$, and Equation (B.4) holds for each m under a feasible allocation fixed across evaluation dates. Let $\mathcal{F}_{t-1}$ denote the information available at portfolio formation. Assume that portfolio membership, $q_{i,t}$, and $e_{i,j,t}$ are $\mathcal{F}_{t-1}$-measurable (predictability condition), and that $\mathbb{E}[r_{i,t}^2 \mid \mathcal{F}_{t-1}] \leq M < \infty$ uniformly in i, t (moment condition). Then $e_{i,j,t}$ converges uniformly to 0 over assets and evaluation dates, and*

$$\begin{aligned} 0 \leq \rho_m^2 &\leq \mathbb{E}\Big[\big(R_t(g, \boldsymbol{w}) - R_t^{\text{asset},*}\big)^2\Big] \\ &\leq M \max_{1 \leq j \leq m} \sup_{\boldsymbol{z}, \boldsymbol{z}' \in C_j} |f^*(\boldsymbol{z}) - f^*(\boldsymbol{z}')|^2 \longrightarrow 0. \end{aligned} \tag{B.5}$$

Here $\boldsymbol{z}$ and $\boldsymbol{z}'$ are any two characteristic vectors in the same region C_j, so the supremum measures the largest squared difference in values of f^* within that region.

Proof. By Equation (B.4),

$$|e_{i,j,t}| \leq \sup_{\boldsymbol{z},\boldsymbol{z}' \in C_j} |f^*(\boldsymbol{z}) - f^*(\boldsymbol{z}')|.$$

Because f^* is continuous on a compact characteristic space, sufficiently close characteristic vectors have arbitrarily close values of f^*, with the same distance threshold applying throughout the space. As the maximum region diameter tends to zero, the bound tends to zero uniformly over assets and evaluation dates.

Since $q_{i,t} \geq 0$ and $\sum_{j,i} q_{i,t} = 1$, Jensen's inequality gives

$$\left(\sum_{j=1}^{m} \sum_{i=1}^{N_t^j} q_{i,t} e_{i,j,t} r_{i,t} \right)^2 \leq \sum_{j=1}^{m} \sum_{i=1}^{N_t^j} q_{i,t} e_{i,j,t}^2 r_{i,t}^2.$$

Taking conditional expectations and using the second-moment bound therefore yields

$$\begin{aligned} \mathbb{E}\Big[\big(R_t(g, \boldsymbol{w}) - R_t^{\text{asset},*} \big)^2 \Big] &\leq M \, \mathbb{E}\left[\sum_{j=1}^{m} \sum_{i=1}^{N_t^j} q_{i,t} e_{i,j,t}^2 \right] \\ &\leq M \max_{1 \leq j \leq m} \sup_{\boldsymbol{z},\boldsymbol{z}' \in C_j} |f^*(\boldsymbol{z}) - f^*(\boldsymbol{z}')|^2 \longrightarrow 0. \end{aligned}$$

By feasibility,

$$\rho_m^2 \leq \operatorname{Var}\big[R_t(g, \boldsymbol{w}) - R_t^{\text{asset},*} \big] \leq \mathbb{E}\Big[\big(R_t(g, \boldsymbol{w}) - R_t^{\text{asset},*} \big)^2 \Big] \to 0.$$

Equation (20) then gives the utility convergence. □

B.2 Proof of Theorem 2

Assume that at least one tree is feasible and that the stock-level second-moment bound in Lemma 1 holds for all feasible trees. Each leaf portfolio uses nonnegative

weights that sum to one and are known at portfolio formation, so its return has finite mean and variance. For each tree, the mean–variance objective is therefore continuous in the allocation weights and reaches its maximum on the compact feasible allocation set. For any finite L, which is the number of leaves of a tree, the finite threshold grid gives only finitely many candidate trees. Selecting the tree with the highest optimized utility therefore attains $\mathcal{U}_0^{\text{tree},*}(L)$.

B.2.1 Approximation by leaf basis-portfolios

Lemma 2 (Approximation by leaf basis-portfolios)**.** *Fix a strategy $(g, \boldsymbol{w})$. Suppose $f_{m,t}$ is uniformly continuous on $[0,1]^K$ across evaluation dates. Under the predictability and moment conditions in Lemma 1, there exist feasible iterative trees with at most L leaves and allocation rules whose returns converge in mean square to $R_t(g, \boldsymbol{w})$. Their mean–variance utilities converge to $\mathcal{U}_0(g, \boldsymbol{w})$, as $L \to \infty$.*

Proof. We construct a sequence of iterative trees with increasing leaf limits L to approximate the given strategy $(g, \boldsymbol{w})$ with m basis-portfolios. It suffices to construct partitions whose maximum leaf diameter tends to zero.

Starting from $[0,1]^K$, we cycle through all K coordinates, splitting every current leaf at the midpoint of the selected coordinate. This construction refines every leaf in every coordinate, and it does not require general Panel Trees to use midpoint splits. After d complete cycles, the resulting tree has 2^{Kd} leaves, each with side length 2^{-d}. Set $d = \left\lfloor \frac{\log_2 L}{K} \right\rfloor$. Since the result concerns $L \to \infty$ for fixed K, we consider $2^{Kd} \leq L$. The grid and leaf-size assumptions ensure that the resulting tree $\widetilde{\mathcal{T}}_L$ is admissible.

Let $C_{j,L}$ denote the characteristic region of its jth leaf. These regions need not coincide with those of the target strategy's basis-portfolios. For any two characteristic vectors $\boldsymbol{z}, \boldsymbol{z}' \in C_{j,L}$, let z_p and z'_p denote their pth coordinates. Each coordinate interval has length

2^{-d}, so $|z_p - z'_p| \le 2^{-d}$ for $p = 1, \dots, K$. Hence

$$\max_j \operatorname{diam}(C_{j,L}) = \max_j \sup_{\boldsymbol{z},\boldsymbol{z}' \in C_{j,L}} \left(\sum_{p=1}^{K} (z_p - z'_p)^2 \right)^{1/2}$$
$$\le \left(\sum_{p=1}^{K} (2^{-d})^2 \right)^{1/2} = \sqrt{K}\, 2^{-d} \le 2\sqrt{K}\, L^{-1/K} \longrightarrow 0.$$

The last inequality follows from $d \ge (\log_2 L)/K - 1$, and the limit holds as $L \to \infty$ for fixed K.

We assign each leaf the target strategy's total holding in its region. Indexing the assets in leaf j at date t by $i = 1, \dots, N_t^j$, set

$$w_{j,L} := \sum_{i=1}^{N_t^j} h_{i,t}(g, \boldsymbol{w}) = \sum_{i=1}^{N_t^j} q_{i,t} f_{m,t}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w}).$$

With market-portfolio benchmark weight $q_{i,t}$, the individual asset-level holding in asset i is

$$\widetilde{h}_{i,t} = w_{j,L} \frac{q_{i,t}}{\sum_{u=1}^{N_t^j} q_{u,t}} = q_{i,t} \frac{w_{j,L}}{\sum_{u=1}^{N_t^j} q_{u,t}}.$$

For $\boldsymbol{z} \in C_{j,L}$, define

$$f_{L,t}^{\text{leaf}}(\boldsymbol{z}; \boldsymbol{w}_L) := \frac{w_{j,L}}{\sum_{u=1}^{N_t^j} q_{u,t}} = \sum_{u=1}^{N_t^j} \frac{q_{u,t}}{\sum_{v=1}^{N_t^j} q_{v,t}} f_{m,t}(\boldsymbol{z}_{u,t-1}; \boldsymbol{w}).$$

Thus, $f_{L,t}^{\text{leaf}}$ is the market-equity-weighted average of $f_{m,t}$ across assets within each leaf. For asset i in leaf j, define the approximation error

$$e_{i,j,t,L} := f_{L,t}^{\text{leaf}}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w}_L) - f_{m,t}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w}).$$

The averaging weights are nonnegative and sum to one. Subtracting $f_{m,t}$ and applying the triangle inequality gives

$$\begin{aligned}
|e_{i,j,t,L}| &= \left| \sum_{u=1}^{N_t^j} \frac{q_{u,t}}{\sum_{v=1}^{N_t^j} q_{v,t}} \left[f_{m,t}(\boldsymbol{z}_{u,t-1}; \boldsymbol{w}) - f_{m,t}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w}) \right] \right| \\
&\leq \sum_{u=1}^{N_t^j} \frac{q_{u,t}}{\sum_{v=1}^{N_t^j} q_{v,t}} \left| f_{m,t}(\boldsymbol{z}_{u,t-1}; \boldsymbol{w}) - f_{m,t}(\boldsymbol{z}_{i,t-1}; \boldsymbol{w}) \right| \\
&\leq \sup_{\boldsymbol{z}, \boldsymbol{z}' \in C_{j,L}} \left| f_{m,t}(\boldsymbol{z}; \boldsymbol{w}) - f_{m,t}(\boldsymbol{z}'; \boldsymbol{w}) \right|.
\end{aligned}$$

Uniform continuity of $f_{m,t}$ across evaluation dates and shrinking leaf diameters imply that $e_{i,j,t,L}$ converges uniformly to zero over assets and dates.

Since $\widetilde{h}_{i,t} - h_{i,t}(g, \boldsymbol{w}) = q_{i,t} e_{i,j,t,L}$, the return difference is

$$D_{t,L} := R_t(\widetilde{\mathcal{T}}_L, \boldsymbol{w}_L) - R_t(g, \boldsymbol{w}) = \sum_j \sum_{i=1}^{N_t^j} \left[\widetilde{h}_{i,t} - h_{i,t}(g, \boldsymbol{w}) \right] r_{i,t} = \sum_j \sum_{i=1}^{N_t^j} q_{i,t} e_{i,j,t,L} r_{i,t}.$$

Applying the weighted-square argument of Lemma 1, with $\sum_j \sum_{i=1}^{N_t^j} q_{i,t} = 1$, gives

$$\begin{aligned}
\mathbb{E}[D_{t,L}^2] &\leq \mathbb{E}\left[\sum_j \sum_{i=1}^{N_t^j} q_{i,t} e_{i,j,t,L}^2 r_{i,t}^2 \right] \\
&\leq M\, \mathbb{E}\left[\sum_j \sum_{i=1}^{N_t^j} q_{i,t} e_{i,j,t,L}^2 \right] \\
&\leq M \sup_t \max_j \sup_{\boldsymbol{z}, \boldsymbol{z}' \in C_{j,L}} \left| f_{m,t}(\boldsymbol{z}; \boldsymbol{w}) - f_{m,t}(\boldsymbol{z}'; \boldsymbol{w}) \right|^2 \longrightarrow 0.
\end{aligned}$$

The final convergence follows from the shrinking maximum leaf diameter and the uni-

form continuity of $f_{m,t}$ uniformly over evaluation dates. The utility difference satisfies

$$\begin{aligned}&\mathcal{U}_0(\widetilde{\mathcal{T}}_L, \boldsymbol{w}_L) - \mathcal{U}_0(g, \boldsymbol{w}) \\ &\quad = \mathbb{E}[D_{t,L}] - \delta \operatorname{Cov}(R_t(g, \boldsymbol{w}), D_{t,L}) - \frac{\delta}{2} \operatorname{Var}(D_{t,L}) \longrightarrow 0.\end{aligned}$$

Since $\mathbb{E}[D_{t,L}^2] \to 0$, both $\mathbb{E}[D_{t,L}]$ and $\operatorname{Var}(D_{t,L})$ converge to zero. By the Cauchy–Schwarz inequality, $\operatorname{Cov}(R_t(g, \boldsymbol{w}), D_{t,L})$ also converges to zero. □

B.2.2 Greedy splitting under the global economic objective

Fix L and its threshold grid $\mathcal{Q}$. Let Θ_s denote the Panel Tree after s decision steps. For each admissible candidate split ξ on this grid, define the utility gain of this split as

$$\Delta(\xi; \Theta_s) := \mathcal{U}_0(\Theta_s \oplus \xi) - \mathcal{U}_0(\Theta_s). \tag{B.6}$$

Each utility value is computed by optimizing the allocation across all leaf basis-portfolios of the Panel Tree. At each step, the algorithm evaluates all admissible candidate splits and selects the one with the largest strictly positive gain. If no positive-gain split remains, we keep the tree unchanged through step $L-1$ for the purpose of the proof. Write Θ_{L-1} for the terminal tree. Each intermediate tree is feasible and has at most L leaves. Define its utility gap relative to the optimal feasible tree as

$$G_{s,L} := \mathcal{U}_0^{\text{tree},*}(L) - \mathcal{U}_0(\Theta_s), \qquad s = 0, \ldots, L-1. \tag{B.7}$$

Lemma 3 (Utility loss from greedy splitting)**.** *Let $L \geq 2$. Suppose there exist a gain fraction $0 < \chi_L \leq 1$ and an additive tolerance $\epsilon_L \geq 0$ such that, at every step $s < L-1$ with $G_{s,L} > \epsilon_L/\chi_L$,*

at least one admissible split ξ satisfies

$$\Delta(\xi;\Theta_s) \geq \chi_L G_{s,L} - \epsilon_L > 0. \tag{B.8}$$

Then the greedy rule described above satisfies

$$0 \leq \Gamma_L \leq e^{-\chi_L(L-1)} G_{0,L} + \frac{\epsilon_L}{\chi_L}. \tag{B.9}$$

In particular, $\Gamma_L \to 0$ *if* $G_{0,L} = O(1)$, $\chi_L L \to \infty$, *and* $\epsilon_L/\chi_L \to 0$.

Proof. When $G_{s,L} > \epsilon_L/\chi_L$, the largest-gain split achieves at least the improvement in Equation (B.8). When $G_{s,L} \leq \epsilon_L/\chi_L$, accepting only positive gains or retaining the current tree gives $G_{s+1,L} \leq G_{s,L} \leq (1-\chi_L)G_{s,L} + \epsilon_L$. Thus, in both cases,

$$G_{s+1,L} = G_{s,L} - \Delta(\xi^*_{s+1};\Theta_s) \leq (1-\chi_L)G_{s,L} + \epsilon_L.$$

Iterating for $L-1$ decision steps gives

$$\Gamma_L = G_{L-1,L} \leq (1-\chi_L)^{L-1} G_{0,L} + \epsilon_L \sum_{s=0}^{L-2} (1-\chi_L)^s \leq e^{-\chi_L(L-1)} G_{0,L} + \frac{\epsilon_L}{\chi_L}.$$

The stated growth conditions make both terms vanish. □

Equation (B.8) additionally requires a positive-gain candidate whenever the remaining gap exceeds ϵ_L/χ_L, and quantifies the progress it delivers.

Proof of Theorem 2. Both A_L and Γ_L are nonnegative: the Panel Tree is a feasible iterative tree under the same leaf limit, and every feasible recursive-tree strategy is a feasible characteristic-managed basis-portfolio strategy.

To establish $A_L \to 0$, fix $\eta > 0$ and choose a feasible target strategy $(g, \boldsymbol{w})$ satisfying Lemma 2 and $\mathcal{U}_0(g, \boldsymbol{w}) > \mathcal{U}_0^{\text{basis},*} - \eta$. Such strategies exist by the theorem's assumptions. Hold this target fixed as L increases. The lemma supplies feasible tree strategies whose utilities converge to its utility. Since the optimal tree value is at least the utility of each constructed strategy,

$$\begin{aligned} 0 \le A_L &\le \mathcal{U}_0^{\text{basis},*} - \mathcal{U}_0(\widetilde{\mathcal{T}}_L, \boldsymbol{w}_L) \\ &\le \eta + \left| \mathcal{U}_0(g, \boldsymbol{w}) - \mathcal{U}_0(\widetilde{\mathcal{T}}_L, \boldsymbol{w}_L) \right|. \end{aligned}$$

Taking $L \to \infty$ gives $\limsup_{L\to\infty} A_L \le \eta$. Since $\eta > 0$ is arbitrary, $A_L \to 0$. Here η is a utility tolerance used only to approach the supremum; neither the target nor the constructed allocations need be the same for different tolerances.

The greedy lemma gives

$$0 \le \mathcal{U}_0^{\text{basis},*} - \mathcal{U}_0(\Theta_{L-1}) \le A_L + e^{-\chi_L(L-1)} G_{0,L} + \frac{\epsilon_L}{\chi_L}. \tag{B.10}$$

Since $A_L \to 0$, the conditions $G_{0,L} = O(1)$, $\chi_L L \to \infty$, and $\epsilon_L/\chi_L \to 0$ make the full bound vanish. This proves Equation (25). □

Corollary 1 (From individual assets to Panel Tree)**.** *Under the convergence conditions of Theorems 1 and 2,*

$$0 \le \mathcal{U}_0^{\text{asset},*} - \mathcal{U}_0(\Theta_{L-1}) = A_L + \Gamma_L \longrightarrow 0. \tag{B.11}$$

Proof. Under Theorem 1, feasible basis-portfolio strategies are admissible individual asset-level strategies and $\mathcal{U}_0^{\text{basis},*}(m) \to \mathcal{U}_0^{\text{asset},*}$. Taking the supremum over m therefore gives $\mathcal{U}_0^{\text{basis},*} = \mathcal{U}_0^{\text{asset},*}$. Theorem 2 independently establishes convergence to $\mathcal{U}_0^{\text{basis},*}$. Combining the two results proves the corollary. □

B.3 Proof of Theorem 3

Proof. For the fixed cost-ignorant-optimal tree, let

$$\boldsymbol{w}_{\Theta^0}^{\mathrm{CA}} \in \arg\max_{\boldsymbol{w}\in\mathcal{A}(\Theta^0)} \left\{\mathcal{U}_0(\Theta^0,\boldsymbol{w}) - \mathbb{E}[C_t(\Theta^0,\boldsymbol{w})]\right\}. \tag{B.12}$$

Thus the benchmark allocation is already optimized for costs. Its pair with Θ^0 remains feasible in the joint problem, so

$$\mathcal{U}_1(\Theta^1) = \max_{\Theta\in\mathbb{T}_L} \mathcal{U}_1(\Theta) \geq \mathcal{U}_1(\Theta^0). \tag{B.13}$$

Gross optimality of Θ^0 gives $\mathcal{D}(\Theta) \geq 0$. Nonnegative costs imply $\mathcal{U}_1(\Theta) \leq \mathcal{U}_0(\Theta)$, so $\mathcal{K}(\Theta) \geq 0$. Subtracting the optimized cost-aware criterion from the constant $\mathcal{U}_0(\Theta)$ gives

$$\mathcal{K}(\Theta) = \min_{\boldsymbol{w}\in\mathcal{A}(\Theta)} \left\{\mathcal{U}_0(\Theta) - \mathcal{U}_0(\Theta,\boldsymbol{w}) + \mathbb{E}[C_t(\Theta,\boldsymbol{w})]\right\}. \tag{B.14}$$

For every feasible tree,

$$\mathcal{U}_1(\Theta) = \mathcal{U}_0(\Theta^0) - \mathcal{D}(\Theta) - \mathcal{K}(\Theta). \tag{B.15}$$

Subtracting the value at Θ^0 and maximizing over Θ proves Equation (30). Equivalently,

$$\mathcal{U}_1(\Theta^1) - \mathcal{U}_1(\Theta^0) = \mathcal{K}(\Theta^0) - \min_{\Theta\in\mathbb{T}_L} \{\mathcal{D}(\Theta) + \mathcal{K}(\Theta)\}. \tag{B.16}$$

The maximum in Equation (30) is positive exactly when some feasible tree satisfies Equation (31). Equality means that $\mathcal{U}_1(\Theta^0)$ attains the global cost-aware optimum. This also shows that disjoint cost-ignorant-optimal and cost-aware-optimal sets are a sufficient, but unnecessary, condition for strictness relative to the selected Θ^0. □

B.4 Economic interpretation and split selection

Friction loss after allocation adjustment. For a fixed tree, let $\boldsymbol{w}_\Theta^{\mathrm{CA}}$ maximize its cost-aware criterion. Equation (B.14) gives

$$\begin{aligned}\mathcal{K}(\Theta) = \mathcal{U}_0(\Theta) - \mathcal{U}_0(\Theta, \boldsymbol{w}_\Theta^{\mathrm{CA}}) \\ + \mathbb{E}[C_t(\Theta, \boldsymbol{w}_\Theta^{\mathrm{CA}})].\end{aligned} \tag{B.17}$$

The friction loss combines the gross utility sacrificed during cost-aware reallocation and the expected trading cost that remains. If $\boldsymbol{w}_\Theta^{\mathrm{CI}}$ maximizes gross utility on the same tree, its feasibility in the cost-aware problem yields

$$0 \leq \mathbb{E}[C_t(\Theta, \boldsymbol{w}_\Theta^{\mathrm{CA}})] \leq \mathcal{K}(\Theta) \leq \mathbb{E}[C_t(\Theta, \boldsymbol{w}_\Theta^{\mathrm{CI}})]. \tag{B.18}$$

Theorem 3 therefore compares two losses: the gross opportunity lost by changing the basis and the friction loss remaining after allocation has adjusted on each basis.

Cost-aware split gains. For a parent tree Θ and an admissible split ξ, write $\Delta\mathcal{U}_\lambda(\xi;\Theta) = \mathcal{U}_\lambda(\Theta \oplus \xi) - \mathcal{U}_\lambda(\Theta)$ and $\Delta\mathcal{K}(\xi;\Theta) = \mathcal{K}(\Theta \oplus \xi) - \mathcal{K}(\Theta)$. Then

$$\Delta\mathcal{U}_1(\xi;\Theta) = \Delta\mathcal{U}_0(\xi;\Theta) - \Delta\mathcal{K}(\xi;\Theta). \tag{B.19}$$

A split improves the cost-aware criterion exactly when its gross utility gain exceeds the increase in friction loss after allocation adjusts. The increment in friction loss need not be positive.

Proposition 1 (Reversal of a candidate-split ranking)**.** *Let Θ_A and Θ_B be two admissible children of the same parent. Suppose $\mathcal{U}_0(\Theta_A) > \mathcal{U}_0(\Theta_B)$. The cost-aware criterion strictly prefers Θ_B if and only if*

$$\mathcal{K}(\Theta_A) - \mathcal{K}(\Theta_B) > \mathcal{U}_0(\Theta_A) - \mathcal{U}_0(\Theta_B). \tag{B.20}$$

Proof. Subtracting the two cost-aware values gives

$$\mathcal{U}_1(\Theta_B) - \mathcal{U}_1(\Theta_A) = \mathcal{K}(\Theta_A) - \mathcal{K}(\Theta_B) - [\mathcal{U}_0(\Theta_A) - \mathcal{U}_0(\Theta_B)].$$

The difference is positive precisely under Equation (B.20); equality gives a tie. □

The condition characterizes when differences in friction loss reverse the gross-utility ranking of two discrete candidate splits. Such reversals can change the selected characteristics and thresholds.

Corollary 2 (One-step comparison from a common parent)**.** *For a fixed training window, let*

$$\mathcal{N}(\Theta) = \{\Theta\} \cup \{\Theta \oplus \xi : \xi \text{ is an admissible split}\}.$$

Use the common optimized training score $\mathcal{U}_{\mathrm{MVU},T}$ in Equation (12), with the fixed-denominator cost convention. Suppose the cost-aware step maximizes this score over $\mathcal{N}(\Theta)$. If a cost-ignorant step selects a member Θ'_{CI} of the same menu, then

$$\mathcal{U}_{\mathrm{MVU},T}(\Theta'_{\mathrm{CA}}) \geq \max\{\mathcal{U}_{\mathrm{MVU},T}(\Theta), \mathcal{U}_{\mathrm{MVU},T}(\Theta'_{\mathrm{CI}})\}. \tag{B.21}$$

Proof. The parent and the cost-ignorant-selected candidate are both available in the cost-aware maximization. □

The corollary applies to a common parent and candidate menu. It does not order

the terminal trees produced by different greedy paths. Retaining the parent as an available choice does not require a refined tree to replicate all of its parent's fixed-allocation strategies.

A sufficient condition for strict redesign gains. The following proposition gives an additional sufficient condition for basis redesign to improve utility beyond optimal cost-aware allocation on the original basis.

Proposition 2 (A sufficient condition for strict improvement)**.** *Under the assumptions of Theorem 3, suppose there exist $a > 0$ and $b > 0$ such that every $\boldsymbol{w} \in \mathcal{A}(\Theta^0)$ satisfies*

$$\mathcal{U}_0(\Theta^0) - \mathcal{U}_0(\Theta^0, \boldsymbol{w}) < a \quad \Longrightarrow \quad \mathbb{E}[C_t(\Theta^0, \boldsymbol{w})] \geq b. \tag{B.22}$$

Suppose a feasible pair $(\widetilde{\Theta}, \widetilde{\boldsymbol{w}})$ satisfies

$$\begin{aligned} s := \mathcal{D}(\widetilde{\Theta}) + [\mathcal{U}_0(\widetilde{\Theta}) - \mathcal{U}_0(\widetilde{\Theta}, \widetilde{\boldsymbol{w}})] \\ + \mathbb{E}[C_t(\widetilde{\Theta}, \widetilde{\boldsymbol{w}})] < \min\{a, b\}. \end{aligned} \tag{B.23}$$

Then

$$\mathcal{U}_1(\Theta^1) - \mathcal{U}_1(\Theta^0) \geq \min\{a, b\} - s > 0. \tag{B.24}$$

Proof. For each allocation on Θ^0, either its gross-utility sacrifice is at least a, or Equation (B.22) gives an expected cost of at least b. Both terms are nonnegative. Taking the minimum of their sum over allocations therefore gives $\mathcal{K}(\Theta^0) \geq \min\{a, b\}$. The alternative allocation is feasible, so Equation (B.14) implies

$$\min_{\Theta \in \mathbb{T}_L} \{\mathcal{D}(\Theta) + \mathcal{K}(\Theta)\} \leq \mathcal{D}(\widetilde{\Theta}) + \mathcal{K}(\widetilde{\Theta}) \leq s.$$

Substituting these two bounds into Equation (B.16) proves the result. □

The condition makes the utility–cost tradeoff of the original basis explicit: allocations that sacrifice less than a in gross utility incur expected costs of at least b, and allocations with expected costs below b sacrifice at least a in gross utility. Any feasible alternative pair with total loss $s < \min\{a, b\}$ establishes strict improvement. The conclusion applies to the selected Θ^0. Applying this sufficient-condition argument to every gross-optimal rule requires checking the condition for each such rule.

Relation to the generation–allocation experiment. For this population comparison, let $x, y \in \{0, 1\}$ index generation and allocation, with zero denoting CI and one CA. Let $\boldsymbol{w}^{x,y}$ maximize $\mathcal{U}_0(\Theta^x, \boldsymbol{w}) - y\mathbb{E}[C_t(\Theta^x, \boldsymbol{w})]$ over $\mathcal{A}(\Theta^x)$. Evaluate all four cells under the same population cost-aware criterion:

$$J_{xy} := \mathcal{U}_0(\Theta^x, \boldsymbol{w}^{x,y}) - \mathbb{E}[C_t(\Theta^x, \boldsymbol{w}^{x,y})]. \tag{B.25}$$

Then $J_{01} = \mathcal{U}_1(\Theta^0)$ and $J_{11} = \mathcal{U}_1(\Theta^1)$, and

$$J_{11} - J_{00} = \underbrace{J_{01} - J_{00}}_{\text{allocation on original basis}} + \underbrace{J_{11} - J_{01}}_{\text{redesign after allocation adjusts}}. \tag{B.26}$$

Both increments are nonnegative under the stated population optimization assumptions. The first follows from within-tree cost-aware allocation, and the second from cost-aware design optimality. The second comparison corresponds to CA–CA versus CI–CA in Section 2.7. It allows allocation to reoptimize on the redesigned basis; it does not hold weights fixed. The result does not rank the magnitudes of the two increments. Changing the order of the generation and allocation adjustments can change their incremental

contributions, consistent with evaluating the empirical hybrid percentages separately.

Scope of the criterion and the guarantees. The population criterion penalizes gross-return variance and subtracts expected trading costs. For a feasible strategy, abbreviate its return and cost by R and C. MVU computed from net returns instead satisfies

$$\begin{aligned}\mathbb{E}[R-C]-\frac{\delta}{2}\operatorname{Var}(R-C)=\mathbb{E}[R]-\frac{\delta}{2}\operatorname{Var}(R)-\mathbb{E}[C]\\ +\delta\operatorname{Cov}(R,C)-\frac{\delta}{2}\operatorname{Var}(C).\end{aligned} \tag{B.27}$$

This identity requires the displayed second moments, including that of C, to be finite; that additional moment condition is not needed for the basic expected-cost dominance theorem. The empirical analysis evaluates realized net-return moments. Its objective values and signs are not guaranteed by the population inequalities above. Likewise, the training corollary concerns the specified fixed-denominator score, whose ordering need not persist under exact-cost reevaluation. None of these results establishes dominance between independently grown greedy terminal trees or their realized out-of-sample performance.

C Cost-Ignorant Allocations and Convex Implementation

C.1 Cost-Ignorant Allocations

Cost-Ignorant MVU Allocation The cost-ignorant MVU allocation is

$$\boldsymbol{w}_{\mathrm{MVU},T}^{\mathrm{CI},*}=\arg\max_{\boldsymbol{w}\in\mathcal{A}_{\mathrm{MVU}}}\quad\widehat{\boldsymbol{\mu}}_T^{\top}\boldsymbol{w}-\frac{\delta}{2}\boldsymbol{w}^{\top}\widehat{\boldsymbol{\Sigma}}_T\boldsymbol{w}. \tag{C.1}$$

Because $\widehat{\boldsymbol{\Sigma}}_T \succ 0$, the objective is strictly concave and the constrained quadratic program has a unique solution.

Cost-Ignorant maximum SR Allocation The cost-ignorant maximum SR allocation is

$$\boldsymbol{w}_{\mathrm{SR},T}^{\mathrm{CI},*} = \arg\max_{\boldsymbol{w}\in\mathcal{A}_{\mathrm{SR}}} \quad \frac{\widehat{\boldsymbol{\mu}}_T^\top \boldsymbol{w}}{\sqrt{\boldsymbol{w}^\top \widehat{\boldsymbol{\Sigma}}_T \boldsymbol{w}}}. \tag{C.2}$$

Because the Sharpe objective is scale invariant, imposing $\|\boldsymbol{w}\|_1 = 1$ gives

$$\boldsymbol{w}_{\mathrm{SR},T}^{\mathrm{CI},*} = \frac{\widehat{\boldsymbol{\Sigma}}_T^{-1}\widehat{\boldsymbol{\mu}}_T}{\|\widehat{\boldsymbol{\Sigma}}_T^{-1}\widehat{\boldsymbol{\mu}}_T\|_1}. \tag{C.3}$$

C.2 Convex Implementation

Stock-Level Trade Coefficients Let $\mathbf{e}_{i,t} \in \mathbb{R}^m$ denote the vector with a one in the position corresponding to the basis-portfolio containing stock i in month t and zeros elsewhere. Set $\mathbf{e}_{i,t} = \mathbf{0}$ when $i \notin \mathcal{I}_t$. Under the fixed-denominator convention $G_{t-1}^{\mathrm{CI}} = 1 + r_{f,t-1} + (\boldsymbol{w}_{\tau,T}^{\mathrm{CI},*})^\top \mathbf{R}_{t-1}$, define

$$\mathbf{a}_{i,t} = \mathbb{1}\{i \in \mathcal{I}_t\}\pi_{i,t}\mathbf{e}_{i,t} - \mathbb{1}\{i \in \mathcal{I}_{t-1}\}\frac{\pi_{i,t-1}(1 + r_{f,t-1} + r_{i,t-1})}{G_{t-1}^{\mathrm{CI}}}\mathbf{e}_{i,t-1}, \tag{C.4}$$

where $\mathbb{1}\{\cdot\}$ denotes the indicator function. The approximate stock trades and transaction-cost terms used in the optimization follow the fixed-denominator approximation:

$$\begin{aligned}\widetilde{\Delta h}_{i,t}(\boldsymbol{w}) &= \mathbf{a}_{i,t}^{\top}\boldsymbol{w},\\ \widetilde{C}_t(\boldsymbol{w}) &= \sum_{i=1}^{\widetilde{N}_t} \kappa_{i,t}\left|\mathbf{a}_{i,t}^{\top}\boldsymbol{w}\right|,\\ \widehat{C}_T(\boldsymbol{w}) &= \frac{1}{W_T}\sum_{t=1}^{W_T}\widetilde{C}_t(\boldsymbol{w}).\end{aligned} \tag{C.5}$$

Convex Programs Introduce auxiliary variables $u_{i,t} \geq |\mathbf{a}_{i,t}^{\top}\boldsymbol{w}|$. The MVU problem in Equation (9) is implemented as

$$\begin{aligned}\min_{\boldsymbol{w},\mathbf{u}} \quad & -\widehat{\boldsymbol{\mu}}_T^{\top}\boldsymbol{w} + \frac{\delta}{2}\boldsymbol{w}^{\top}\widehat{\boldsymbol{\Sigma}}_T\boldsymbol{w} + \frac{1}{W_T}\sum_{t=1}^{W_T}\sum_{i=1}^{\widetilde{N}_t}\kappa_{i,t}u_{i,t}\\ \text{s.t.} \quad & -u_{i,t} \leq \mathbf{a}_{i,t}^{\top}\boldsymbol{w} \leq u_{i,t}, \qquad u_{i,t} \geq 0, \quad t = 1,\ldots,W_T,\ i = 1,\ldots,\widetilde{N}_t,\\ & \mathbf{1}^{\top}\boldsymbol{w} = 1, \qquad \|\boldsymbol{w}\|_2^2 \leq 16/m.\end{aligned} \tag{C.6}$$

For the SR problem in Equation (10), use the volatility normalization $\mathbf{y} = \frac{\boldsymbol{w}}{\sqrt{\boldsymbol{w}^{\top}\widehat{\boldsymbol{\Sigma}}_T\boldsymbol{w}}}$. Positive homogeneity of $\widehat{C}_T$ gives the implementation

$$\begin{aligned}\max_{\mathbf{y},\mathbf{u}} \quad & \widehat{\boldsymbol{\mu}}_T^{\top}\mathbf{y} - \frac{1}{W_T}\sum_{t=1}^{W_T}\sum_{i=1}^{\widetilde{N}_t}\kappa_{i,t}u_{i,t}\\ \text{s.t.} \quad & \left\|\widehat{\boldsymbol{\Sigma}}_T^{1/2}\mathbf{y}\right\|_2 \leq 1,\\ & -u_{i,t} \leq \mathbf{a}_{i,t}^{\top}\mathbf{y} \leq u_{i,t}, \qquad u_{i,t} \geq 0, \quad t = 1,\ldots,W_T,\ i = 1,\ldots,\widetilde{N}_t.\end{aligned} \tag{C.7}$$

If an optimizer $\mathbf{y}^*$ attains a positive objective value, the volatility constraint binds, and

the allocation is

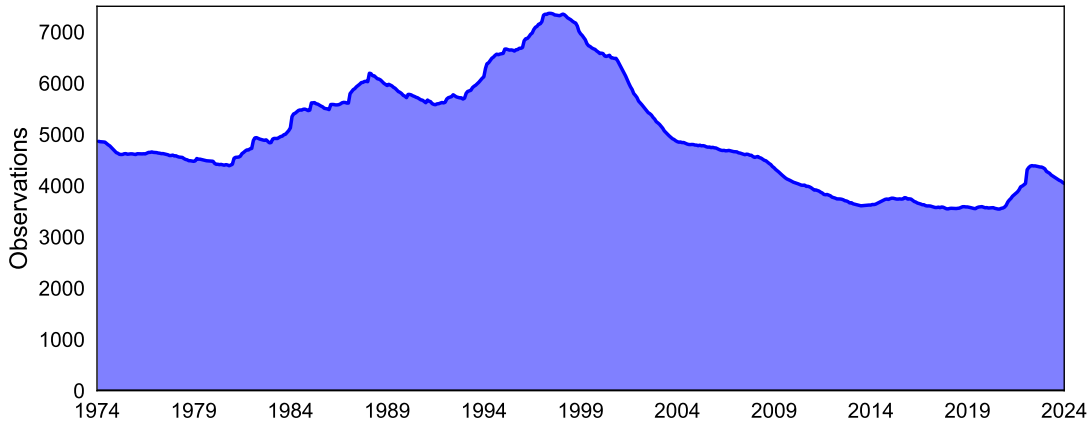
$$\boldsymbol{w}^*_{\mathrm{SR},T} = \frac{\mathbf{y}^*}{\|\mathbf{y}^*\|_1}.$$

D Stock Data

We use the publicly available dataset from Jensen et al. (2023), which combines return data from CRSP with accounting data from Compustat. We restrict our sample to common stocks (shrcd: 10 and 11) traded on USA exchanges (`excntry = USA`). To be included in our sample, stocks must have non-missing data on lagged market equity and returns, as well as positive book equity and lagged market equity.

Figure D.1: The Number of Stocks

Note: This figure shows the number of common stocks in our monthly sample from January 1974 to December 2023. The sample is constructed following Jensen et al. (2023) using CRSP–Compustat data and is restricted to U.S.-traded common stocks with non-missing lagged market equity and returns, and positive book equity and lagged market equity.



Our monthly sample starts in January 1974 and ends in December 2023. The number

of stocks in each month is shown in Figure D.1. The average and median numbers of monthly stock observations are 5,551 and 5,597 for the first 25 years, and 4,405 and 4,103 for the latter 25 years. We use the 153 characteristics studied in Jensen et al. (2023), and each characteristic is standardized cross-sectionally to the $[0, 1]$ interval. For missing characteristic values, we first impute using the industry median based on SIC, NAICS, and GICS classifications; if a value remains missing, we then impute using the cross-sectional median. We apply cross-sectional winsorization at the 1% and 99% levels to mitigate the impact of outliers.

E Tree Structure under MVU

This section examines whether the reallocation of split capacity documented under the maximum Sharpe-ratio objective also appears under mean–variance utility (MVU) with $\delta = 2$. The empirical implementation allows at most nine splits, but an MVU tree stops earlier when no admissible split improves the objective. Across the 40 rolling windows, the cost-ignorant (CI) and cost-aware (CA) trees average 8.18 and 8.25 realized splits, respectively. We therefore measure frequency shares relative to each strategy's average realized number of splits and assign a score of zero to unfilled split positions.

Table E.1 shows that cost-aware training tilts realized split capacity toward accounting-fundamental signals. Panel A reports that their frequency share rises from 58.72% under CI to 70.30% under CA, while their score share rises from 51.59% to 65.54%. Because the score measure places greater weight on earlier splits, the 13.95 percentage-point increase in score share, compared with the 11.58 percentage-point increase in frequency share, indicates that the shift occurs near the top of the tree as well as in the number of selections.

Panel B identifies the categories underlying this reallocation. The largest score in-

creases occur in Investment ($+2.58$), Quality ($+2.48$), Value ($+2.10$), and Low Leverage ($+1.00$), while the largest declines occur in Momentum (-2.23), Short-Term Reversal (-2.10), Profit Growth (-1.20), and Size (-1.03). Figure E.3 reports the corresponding category-level frequencies and shares, and Figure E.4 traces the CA-minus-CI score differences across rolling windows. The direction of the reallocation therefore parallels the maximum-SR evidence, although the category ranking varies with the portfolio objective.

Figure E.1: **Cost-Ignorant Panel Tree Diagram**

Notes. This figure reports the cost-ignorant Panel Tree diagram with the objective of mean–variance utility and the restriction of no more than 10 leaves. The model is estimated using the latest rolling window, from January 2013 to December 2022. The figure shows the selected characteristics, the cutoff values for each parent node, and the associated node labels and split order. Nodes are labeled N# (index) and S# (split order). Terminal leaf nodes display the median monthly number of assets (#) and portfolio weights (w).

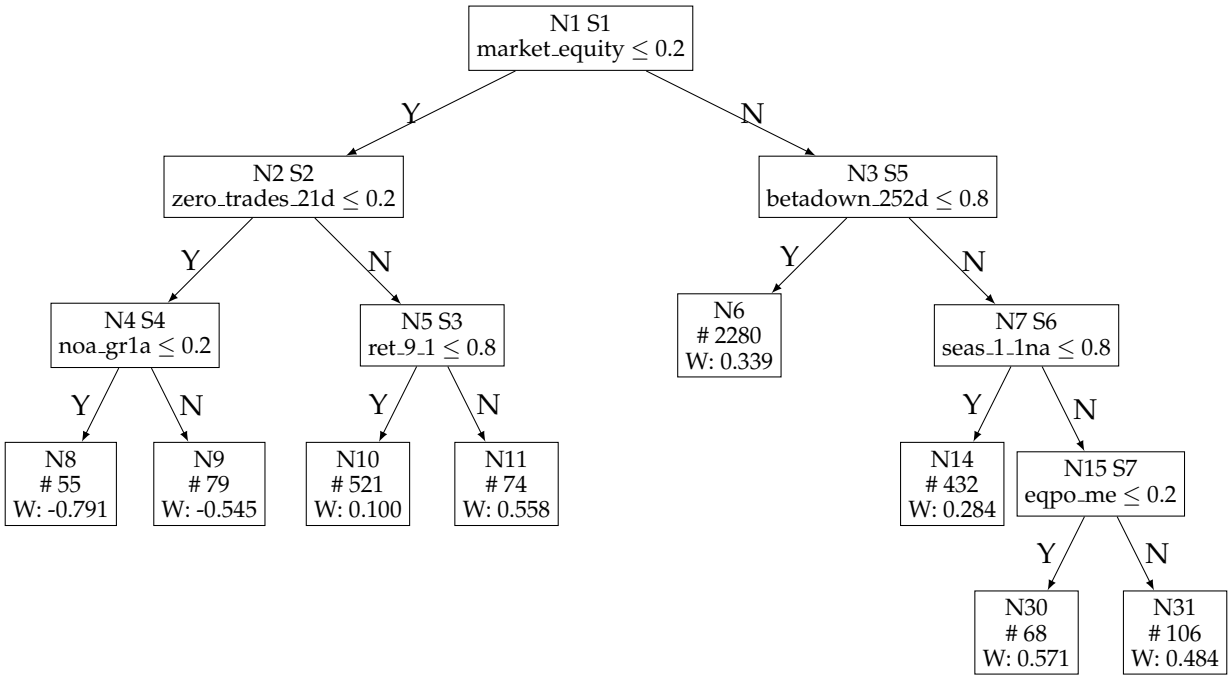


Figure E.2: **Cost-Aware Panel Tree Diagram**

Notes. This figure reports the cost-aware Panel Tree diagram with the objective of mean–variance utility and the restriction of no more than 10 leaves. The model is estimated using the latest rolling window, from January 2013 to December 2022. The figure shows the selected characteristics, the cutoff values for each parent node, and the associated node labels and split order. Nodes are labeled N# (index) and S# (split order). Terminal leaf nodes display the median monthly number of assets (#) and portfolio weights (w).

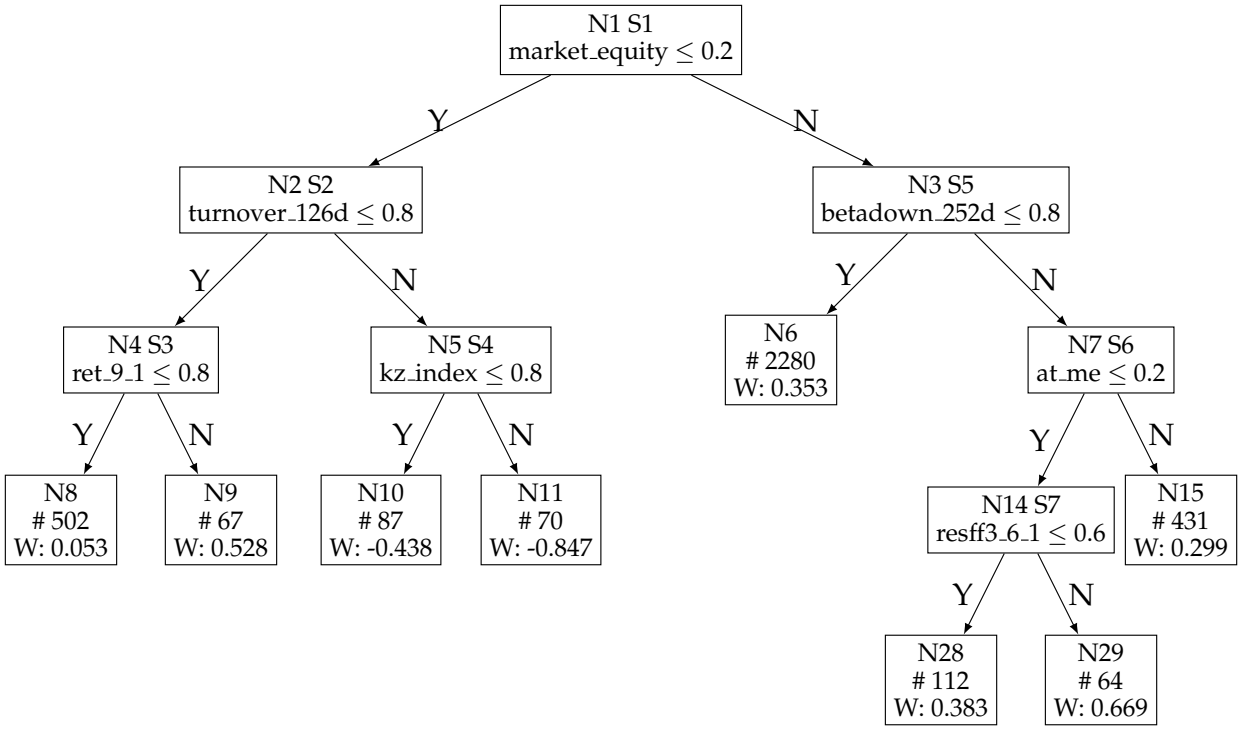


Table E.1: **Reallocation of MVU Split Capacity across Rolling Windows**

Notes. This table summarizes realized split allocation under the cost-ignorant (CI) and cost-aware (CA) MVU trees, averaged over 40 rolling windows. Panel A aggregates splits into two meta-categories: accounting-fundamental and price-volume. Frequency shares divide category-level average frequency by each strategy's average realized number of splits. Within each rolling window, the split score is 9 for the first split, 8 for the second, and so on down to 1 for the ninth split; absent splits receive zero. Panel B reports the four largest increases and decreases in average score (CA minus CI), together with the corresponding changes in average split frequency.

Panel A: **Aggregate Split Allocation by Meta-category**

	Avg. frequency		Frequency share (%)		Avg. score		Score share (%)	
Meta-category	CI	CA	CI	CA	CI	CA	CI	CA
Accounting-fundamental	4.80	5.80	58.72	70.30	22.38	28.63	51.59	65.54
Price-volume	3.38	2.45	41.28	29.70	21.00	15.05	48.41	34.46

Panel B: **Largest Category-level Changes (CA minus CI)**

Largest increases	Δ Freq.	Δ Score	Largest decreases	Δ Freq.	Δ Score
Investment	+0.55	+2.58	Momentum	-0.30	-2.23
Quality	+0.23	+2.48	Short-Term Reversal	-0.38	-2.10
Value	+0.65	+2.10	Profit Growth	-0.25	-1.20
Low Leverage	+0.08	+1.00	Size	-0.23	-1.03

Figure E.3: **MVU Split Frequency and Shares by Characteristic Category**

Notes. This figure reports the average proportion of total splits assigned to each characteristic category under the cost-aware (CA) and cost-ignorant (CI) Panel Trees. The averages are computed across 40 rolling windows, and each tree is allowed at most nine splits per window. The figure provides an intuitive category-level view of how fixed split capacity is reallocated once transaction costs are internalized.

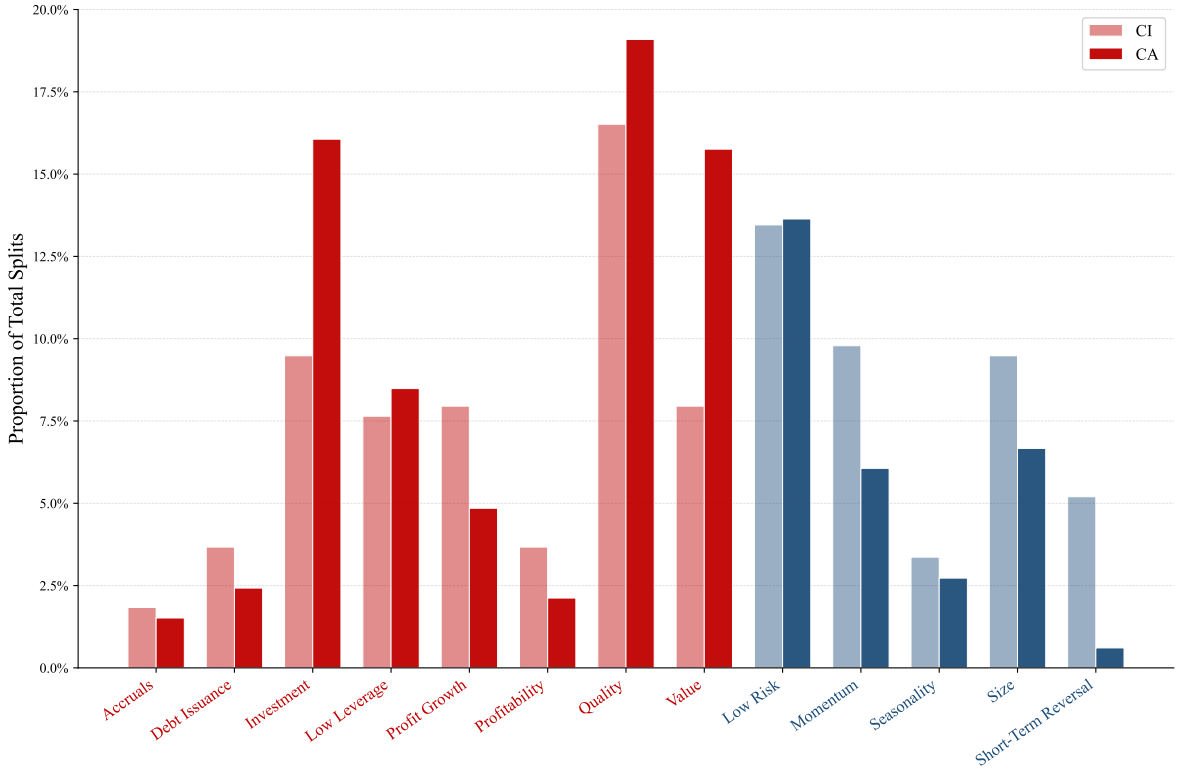


Figure E.4: **MVU Difference in Score-Weighted Split Importance: CA Minus CI**

Notes. This figure reports, for each rolling window and characteristic category, the difference in split scores between the cost-aware and cost-ignorant Panel Trees. Within each rolling window, the split score is 9 for the first split, 8 for the second, and so on down to 1 for the ninth split; absent splits receive zero. Positive values indicate that the category becomes more important under cost-aware training. Because score is a split-order-weighted measure of split importance, the figure captures not only whether a category is selected but also whether it is used higher in the tree.

