

Methodology: Walk-Forward Factor Selection with Meta-Selection

JKP Characteristics Trading Factor (CTF) Challenge Submission

1. Overview

This strategy constructs a long-short US equity portfolio by selecting and weighting stock characteristics ("factors") using a four-stage walk-forward process. As documented by Jensen, Kelly and Pedersen (2023), many published financial anomalies fail to replicate or show substantial out-of-sample decay due to post-publication crowding. To resolve this "replication crisis," this strategy implements an adaptive walk-forward ensembling framework inspired by Kelly, Kuznetsov, Malamud and Xu (2025). Rather than committing to a single configuration ex-ante, the strategy maintains a menu of 72 candidate configurations and selects among them dynamically based solely on their trailing 20-year out-of-sample performance, ensuring zero configuration selection bias.

2. The Four Stages

Stage 1: Factor Portfolio Returns

For each of ~ 400 stock characteristics and each month t , cross-sectionally percentile-rank all stocks on the characteristic. Assign weight $w_i = \text{rank}_i - 0.5$ (long the high-ranked, short the low-ranked). Compute the factor's long-short return: $LS_t = \sum(w_i * r_i) / \sum(|w_i|)$, using ret_exc_lead1m (the stock's excess return realized over month $t+1$). Require ≥ 20 stocks with valid ranks; drop factors with fewer than 60 valid months. This directly follows the univariate sorting protocol established by Jensen, Kelly and Pedersen (2023) to construct standardized, investable anomaly portfolios.

Look-ahead safeguard: The LS value stored at month t is the return earned over $t+1$. It becomes known only at the end of $t+1$. All subsequent stages index their lookback windows accordingly: at rebalance date rp , they use LS rows $\leq rp - 1$ month.

Stage 2: Walk-Forward Configuration Simulation

A fixed, symmetric menu of 72 configurations is defined as the full cross of: lookback window {10y, 20y, expanding} x number of factors {5, 10} x selection rule {Sharpe, ClusterSharpe, Sortino} x weighting rule {EW, RankSharpe, InvVol, MeanVar}.

Every 6 months (rebalance date rp), each configuration selects factors and computes weights using only LS rows $\leq rp - 1$ month, then earns the realized LS returns of its chosen factors until the next rebalance. Its resulting out-of-sample return series is recorded.

Selection rules: Sharpe selects top- N factors by annualized Sharpe. Sortino selects top- N factors by Sortino ratio. ClusterSharpe hierarchically clusters factors by return correlation (average linkage), then takes the best-Sharpe factor from each of the top N clusters.

Weighting rules: EW (equal weights), RankSharpe (weights proportional to Sharpe rank), InvVol (weights proportional to inverse volatility), MeanVar (mean-variance optimization with 30% diagonal shrinkage, non-negative weights). In the MeanVar sleeve, we apply a 30% diagonal shrinkage to the covariance matrix. This implementation is guided by Ao, Li and Zheng (2019), who prove that standard sample covariance matrices are unstable in high dimensions (where $N > T$) and that linear shrinkage asymptotically approaches mean-variance efficiency.

Look-ahead safeguard: Factor selection and weighting at rp use only LS rows through $rp - 1$ month. The LS value at $rp - 1$ represents the return realized over month rp , which is known at the end of month rp . Rebalances with under 24 months of history are skipped.

Stage 3: Meta-Selection

At each rebalance rp , compute each configuration's annualized Sharpe ratio over its own out-of-sample returns up to $rp - 1$

month using a trailing 20-year window. Among configurations with ≥ 120 months of OOS history, blend the top 3 equally. If no configuration yet has 120 months of history, blend ALL active configurations equally (the "EW prior"). Following Kelly, Kuznetsov, Malamud and Xu (2025), blending the top-performing configurations over a rolling window (as opposed to hard-selecting a single model) dynamically adapts the portfolio to the prevailing market regime while reducing overfit.

Look-ahead safeguard: The meta-selection criterion uses only each configuration's own out-of-sample history up to $rp - 1$ month. The methodology live at any date depends only on data before that date - this removes configuration selection bias.

Stage 4: Stock-Level Weights

For each month t , identify the governing rebalance's chosen configurations. For each, map factor weights to individual stocks: stock i receives weight $fw_k * (rank_{i,k} - 0.5) / \sum_j (|rank_{j,k} - 0.5|)$, summed over its factors k . Average equally across chosen configurations. Output all non-zero weights as (id, eom, w) . This linear rank-based stock weight mapping is a non-parametric implementation of the Parametric Portfolio Policies (PPP) framework of Brandt, Santa-Clara and Valkanov (2009), which scales stock positions linearly with their standardized characteristics.

Look-ahead safeguard: Stock ranks use only month- t characteristics (known at month-end t). Factor weights and selections were determined at the governing rebalance using only past data.

3. Temporal Integrity

The strategy enforces a strict three-layer no-look-ahead discipline:

- Factor selection at rebalance rp uses LS rows $\leq rp - 1$ month.
- Factor weighting at rp uses the same lookback window.
- Meta-selection at rp uses only each configuration's OOS returns through $rp - 1$ month.

The variable `ret_exc_lead1m` is used ONLY to compute the factor LS returns in Stage 1 (it is the prediction target). It is never used as a characteristic or feature. Its one-month-forward nature is accounted for by the convention that the LS value at month t is realized over $t+1$, and all lookback windows end one month before the rebalance date.

4. Design Constants (Fixed A Priori)

Parameter	Value	Purpose
Rebalance frequency	6 months	Meta-selection / factor refresh cycle
Meta top-K	3	Number of configs blended after meta-selection
Meta min observations	120 months	OOS history required before meta-selection
Meta window size	20 years	Trailing window for meta-selection OOS Sharpe
Min factor observations	60 months	Factor must have this many valid LS months
Min stocks per side	10	Minimum stocks for a valid factor portfolio
Covariance shrinkage	30%	Diagonal shrinkage in MeanVar weighting
Cluster multiplier	3x	$n_clusters = 3 * n_factors$ in ClusterSharpe

5. Determinism and Reproducibility

The strategy uses no randomness: no random seeds, no stochastic optimization, no sampling. Two consecutive runs on identical data produce byte-identical output. The only dependencies are numpy, pandas, and scipy - all deterministic for the operations used.

6. Verification

- V2 (Truncation test): Running on data truncated at any date produces identical factor selections and meta-choices for all prior rebalances, confirming no future data leaks into past decisions.
- V5 (Determinism): Two consecutive runs produce identical weights, confirming no stochastic elements.

7. References and Citations

Ao, M., Li, Y. and Zheng, X. (2019). "Approaching mean-variance efficiency for large portfolios" ***The Review of Financial Studies***, **32(7)**, pp.2890-2919.

Brandt, M.W., Santa-Clara, P. and Valkanov, R. (2009). "Parametric portfolio policies: Estimation with a large number of assets" ***The Review of Financial Studies***, **22(9)**, pp.3411-3447.

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

Kelly, B., Kuznetsov, B., Malamud, S. and Xu, A. (2025). "Artificial Intelligence Asset Pricing Models" ***National Bureau of Economic Research Working Paper Series***, No. w33351.