ADVANCED PORTFOLIO REPLICATION STRATEGY

Giacomo Giovanni Papa, Alessandro Torazzi, Emma Hyckenberg Dalin, Seyedehyekta Kamaneh

INTRODUCTION

Objective: Develop and evaluate systematic strategies to **replicate the performance of a black box index.**

We tested multiple machine learning models and found that the best performing approach was the **Kalman Filter initialized with Ridge regression**. Subsequently, we applied various strategies to **enhance robustness.**

Model selection

Among all the models tested, the best performing one was the Kalman Filter with Ridge initialization.

Robustness Enhancements

To improve stability and realism, we introduced:

- Beta scaling to adjust exposure intensity.
- Buffering to reduce overtrading.
- Transaction costs to penalize

Tuning

We applied grid search over key hyperparameters and used rolling cross-validation to evaluate model performance over time, ensuring both robustness and generalization.

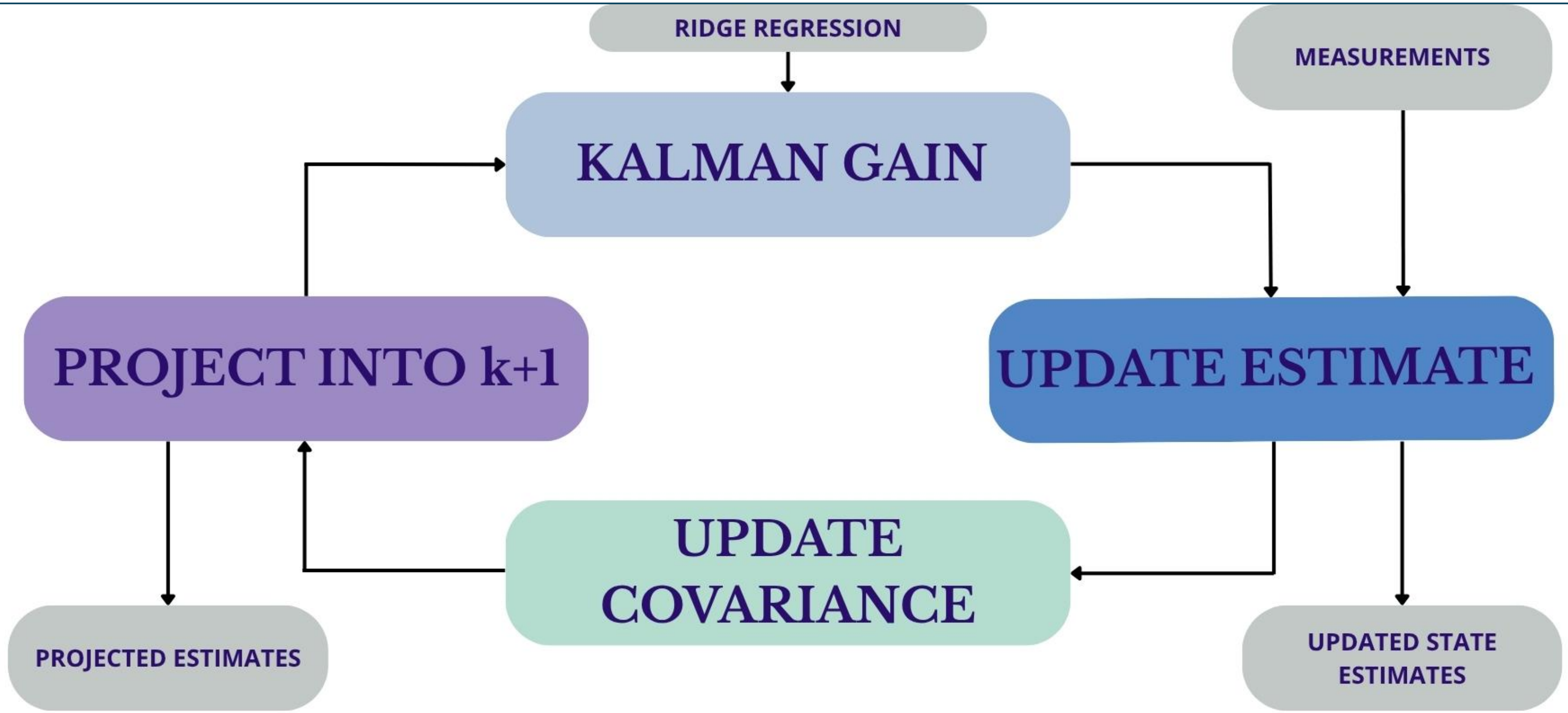
Ensemblinbg

We combined multiple linear models (OLS, Ridge, Lasso, ElasticNet) through equal weight averaging to create a more stable and diversified replica. This approach reduces model specific

Trading strategy

We exploited the divergence between two Kalman filter replicas tracking the same target. By taking opposite positions when their correlation drops.

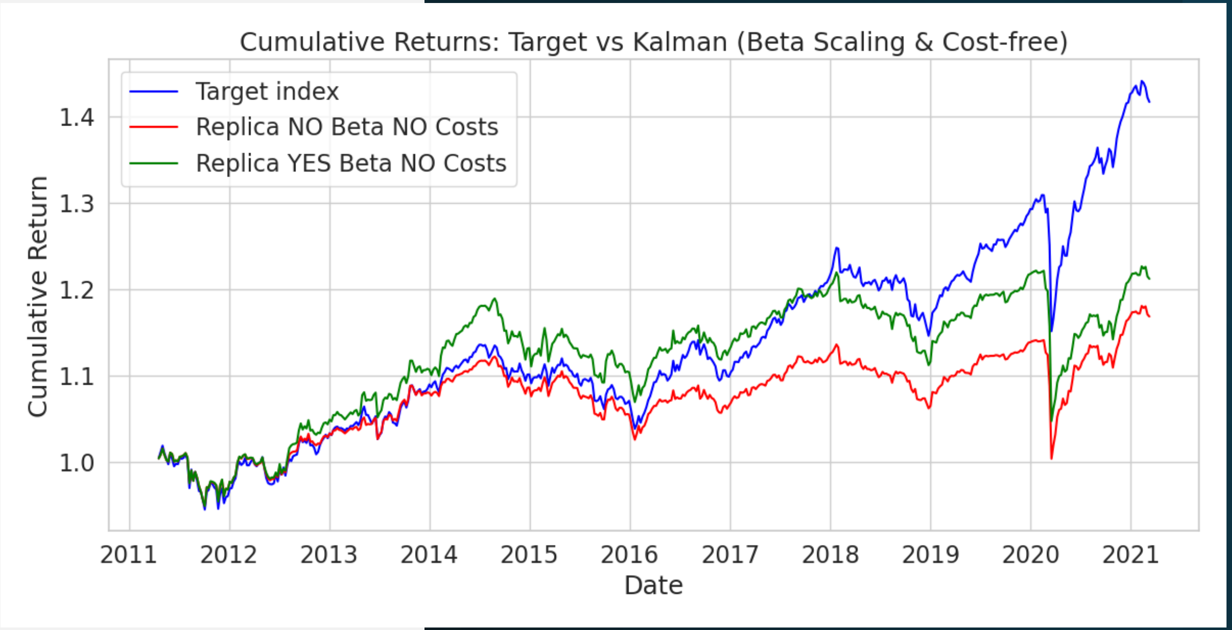
THE MODEL



UPDATE ESTIMATE	$\hat{x}_k = \hat{x}'_k + K_k\left(z_k - H\hat{x}'_k\right)$	PROJECT INTO k+1	$\hat{x}'_{k+1} = \Phi\hat{x}_k$ $\mathrm{P}_{k+1} = \Phi P_k \Phi^T + Q$
UPDATE COVARIANCE	$\mathrm{P}_k = \left(I - K_k H\right) P'_k$	KALMAN GAIN	$\mathrm{K}_k = P'_k H^T\left(H P'_k H^T + R\right)^{-1}$

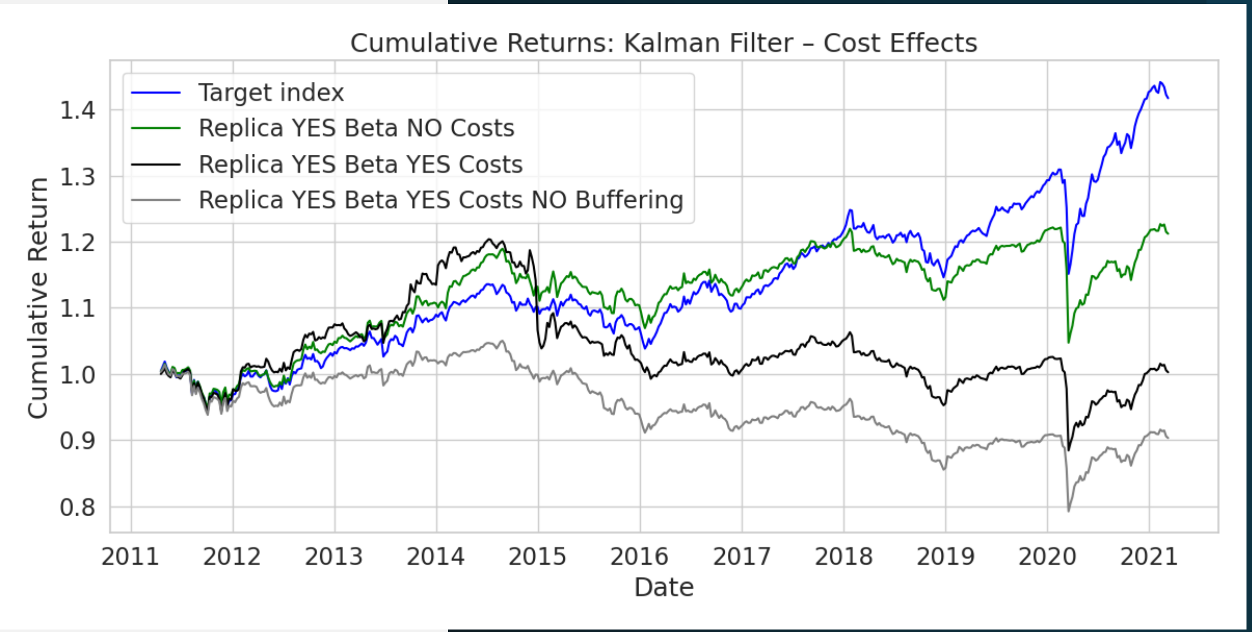
Beta Scaling

- **Purpose:** adjusting the futures weights to match the Beta of the replica with that of the target index
- **Methodology:**
    - Rolling beta calculation (26-week window)
    - Beta between replica and target returns
    - Weights scaling factor: 1/β
- **Outcome:** Enhanced tracking accuracy and stability



Transaction Costs and Buffering

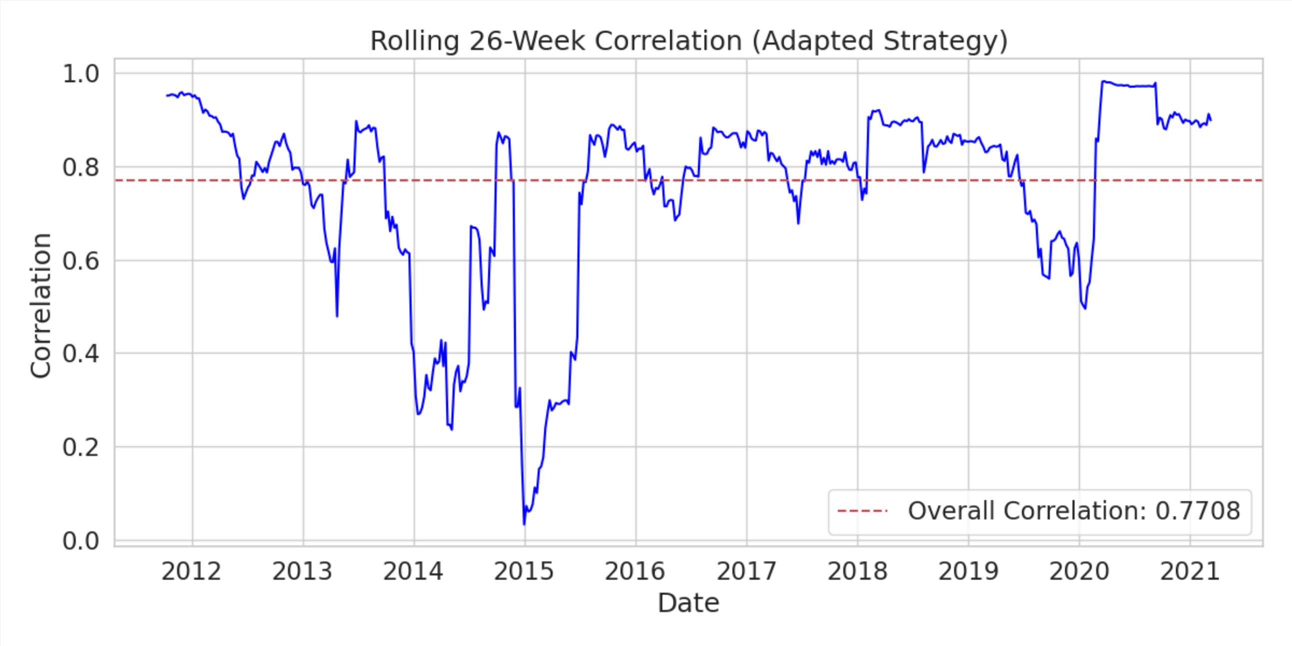
- **Purpose:** adding transaction costs allow to have more realistic results, while the Buffering System minimizes the amount of transaction performed to reduce total costs
- **Methodology:**
    - 5 bps (weight scaled) for each transaction
    - Buffering: each week, the weights are updated only if the weight change is greater than 5%
- **Outcome:** More realistic framework, with Buffering System enhancing performance

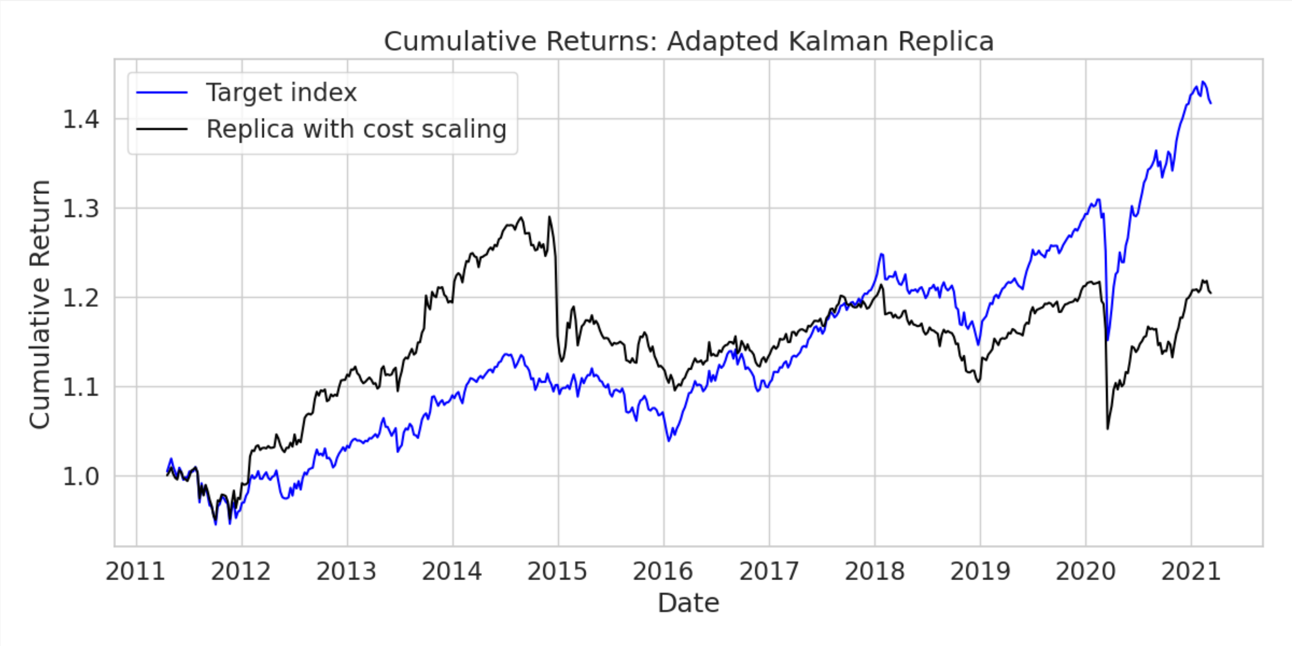


Rolling Validation and Hyperparameters tuning

- **Purpose:** tuning hyperparameters evaluating rolling out-of-sample performance
- **Methodology:**
  - Rolling Training window (variable, 52 weeks maximum) and Validation window (future 52 weeks)
  - Grid search for regularization ratio, Rolling Training window, Beta Scaling window and Kalman Filter parameters
- **Outcome:** Robust hyperparameters tuning, without reducing model's generalization capability

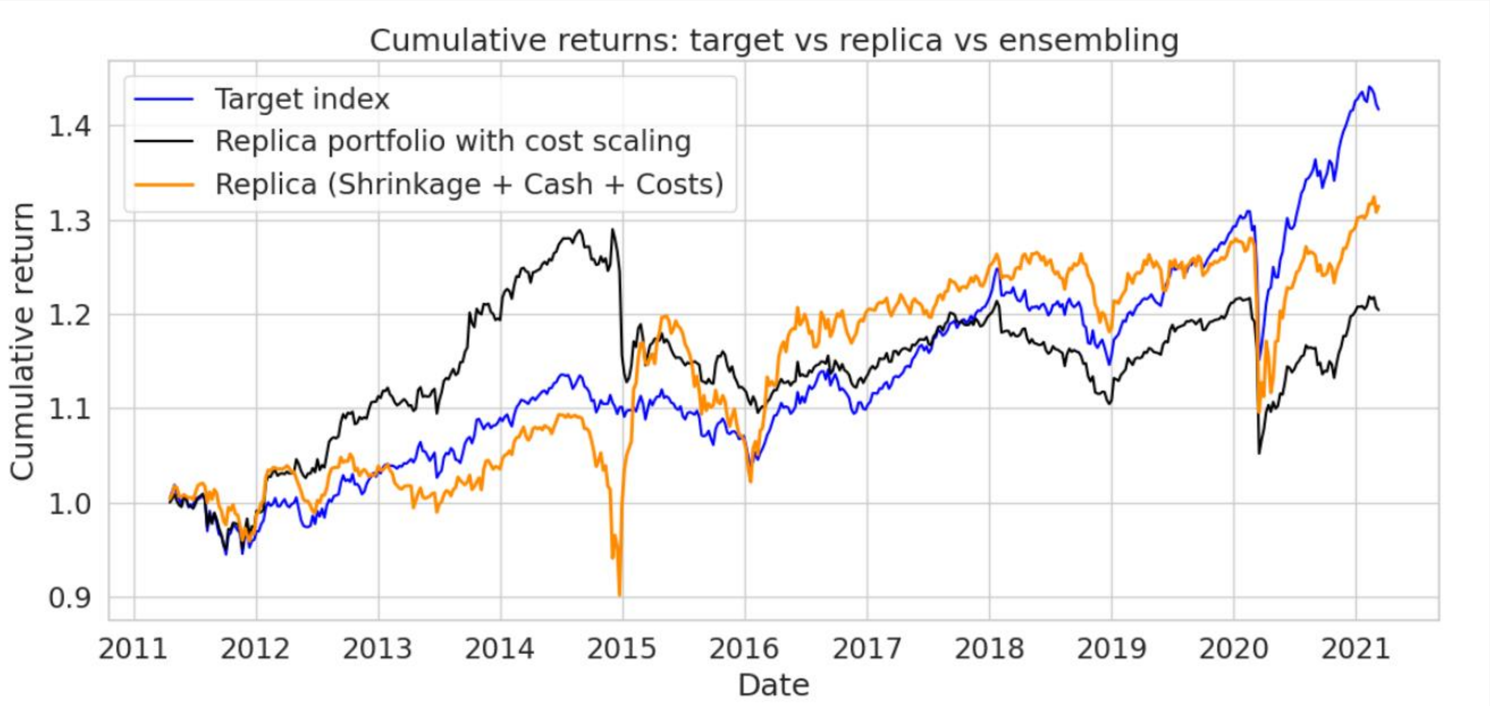
Final Kalman Portfolio Performance





Ensembling Model, Robust Replication Framework

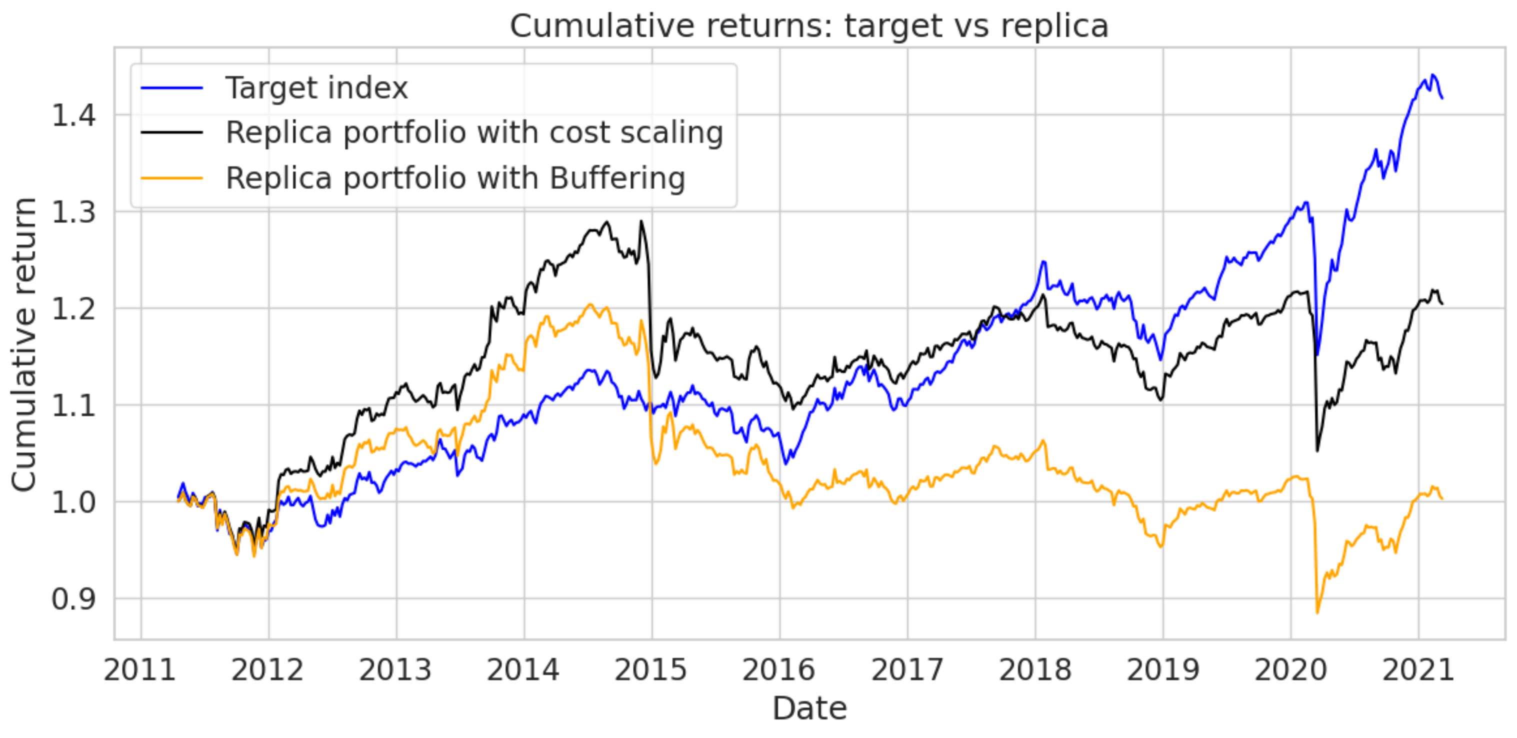
- **Adaptive Ensemble Weights**
  Combine multiple linear models (OLS, Ridge, Lasso, ElasticNet) trained in rolling windows.
- **Shrinkage Toward Equal Weights**
  Apply James–Stein shrinkage (30%) to dampen instability and avoid overfitting:



Residual Cash Allocation

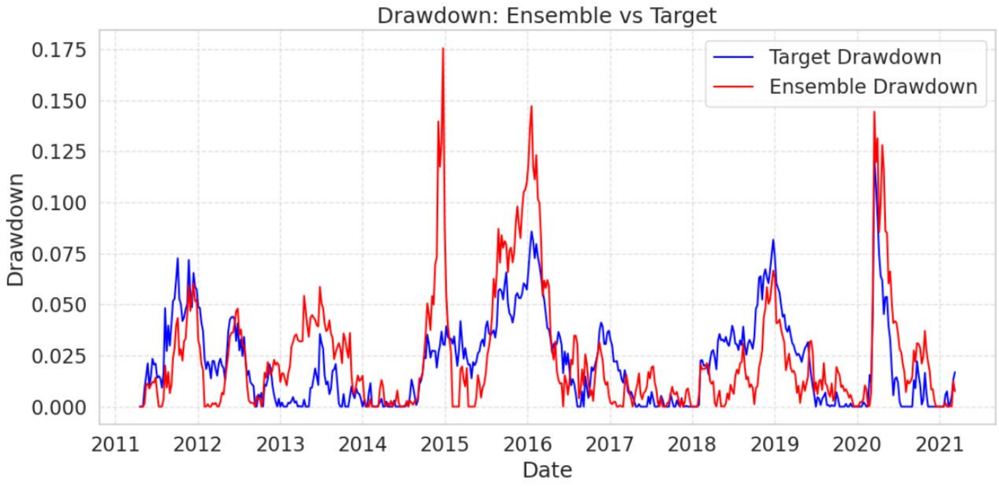
- Captures any leftover capital not invested in risky assets (e.g., due to shrinkage or low exposure)
- **Ensures full capital allocation, including cash**
- **Cash can be held (earning the risk-free rate**) or shorted (**earning the borrow rate**) when overexposed
- Contributes to overall return **without increasing portfolio risk**.

Cost Adjusted Kalman Filter



- **Idea:**
  To account for the loss in performance due to the inclusion of transaction costs, it is possible to directly account for costs in Kalman filter updates
- **Methodology:**
  The observed return y at time t is increased by the transaction costs from t-1, aligning the replica's expected return with net of costs dynamics

- **Outcome:**
  Forcing the model to internalize transaction costs improved significantly the replica's performance and information ratio, at cost of minimal losses in terms of correlation and tracking error



Annualized return	**3.13%**
Annualized Volatility	8.66%
Sharpe Ratio	0.36
Tracking Error	7.49%
Information Ratio	-0.07
Max drawdown	17.54%
Correlation	0.5174

The ensemble achieves **smooth and stable drawdowns** thanks to shrinkage and cash allocation, moreover it replicates them very well.
Including a **cash position** helps improve returns but also reduces correlation **(≈ 0.52)** and the **information ratio** (–0.07).
However, the model still delivers a reasonable Sharpe ratio (0.36) and a controlled drawdown (17.5%),
making it a robust and implementable strategy.

KALMAN FILTER AND ENSEMBLING MODEL COMPARISON

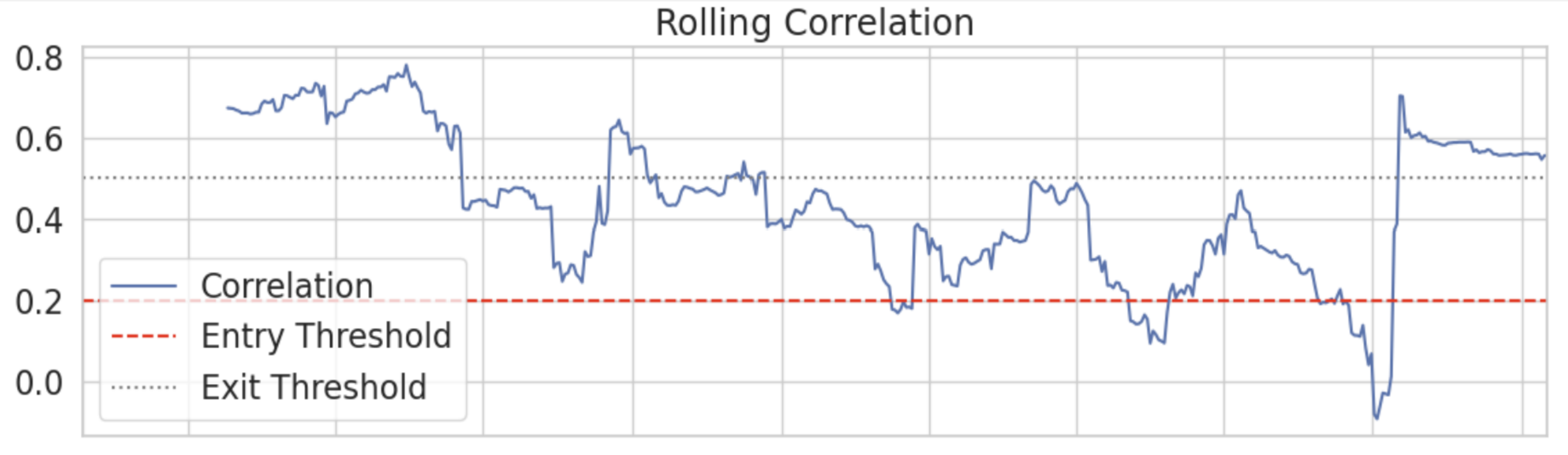
Metrics	Kalman filter with cost scaling	Ensemble (No Costs)
Annualized Return	2.08% 🔽	3.13% 🔺 Higher
Volatility	6.40% 🔽	8.66% 🔺 Higher risk
Tracking Error	4.13% 🔽	7.49% 🔺 Looser tracking
Correlation	0.77 Strong alignment	0.52 Weaker alignment
Information Ratio	–0.38	–0.07

The Adapted strategy stands out for its superior replication quality, it achieves **tight tracking error**, **strong correlation with the target**, and **controlled volatility**.

In contrast, the Ensemble strategy leverages the diversification of multiple linear models to achieve a higher annualized return. However, this comes at the expense of looser tracking and increased volatility. The residual cash position also significantly boosts the overall returns.

Pair Trading with Index Replicas

- **Purpose:** exploit our model to generate alpha
- **Methodology:**
  - Use two independent sets of futures to build separate replicas of the same target index
  - Monitor rolling correlation as a proxy for relative movements between the replicas
  - Enter in the trade when correlation is low and exit when correlation recovers
  - Trade against the spread between the two replicas, assuming correlation mean reversion



Pair Trading with Index Replicas

- **Methodology:**
  - Compute rolling Spread between the two replicas (KF1 andKF2) returns: KF1 - KF2
  - Monitor 52 week rolling correlation entering trade when it falls below 0.3 and exiting when it raises back above 0.5
  - Long spread if Kf1 has recently underperformed against K2 and short spread if it hasn't
- **Issues and Future directions:**
  - Weekly data is insufficient to properly backtest the strategy, daily data would be needed
  - The models hyperparameters could be tuned separately to improve performance