Economics Bulletin

# Volume 45, Issue 4

# Can NASDAQ-100 derivatives ETF portfolio beat QQQ?

Chi-Sheng Lo
*Institute of Information Science, Academia Sinica*

## Abstract

Portfolio optimization in derivative ETF markets presents complex challenges in balancing competing objectives across instruments with fundamentally different risk-return profiles. This paper constructs a portfolio strategy to optimize NASDAQ-100 derivative ETF allocations by balancing tracking error minimization relative to the Invesco QQQ Trust (QQQ) with excess return maximization. The approach dynamically allocates investments across three specialized ETFs: a short-position fund (YQQQ), an income-focused covered-call fund (QYLD), and a triple-leveraged fund (TQQQ). Using a deep reinforcement learning (DRL) framework, the strategy applies anomaly detection to optimize rebalancing timing, incorporating dividend payments to enhance returns. The approach achieves positive excess returns across all evaluation periods, though risk-adjusted performance progressively deteriorates from substantial outperformance during training to underperformance during testing. This progression reveals both the potential and limitations of reinforcement learning approaches for multi-objective portfolio optimization when encountering evolving market conditions.

**Citation:** Chi-Sheng Lo, (2025) "Can NASDAQ-100 derivatives ETF portfolio beat QQQ?", *Economics Bulletin*, Volume 45, Issue 4, pages 1636-1648
**Contact:** Chi-Sheng Lo - sigma.vantage@gmail.com.
**Submitted:** October 17, 2025. **Published:** December 30, 2025.

# 1. Introduction

## 1.1 Objective

This research aims to formulate a portfolio strategy for retail and institutional investors, optimizing allocations of NASDAQ-100 derivative ETFs to balance tracking error (TE) minimization relative to the QQQ benchmark with excess return (ER) maximization. The portfolio dynamically allocates capital across three specialized ETFs: YQQQ (short-position fund), QYLD (income-focused covered-call fund), and TQQQ (triple-leveraged fund), achieving strategic diversification across inverse, income-generating, and leveraged instruments. To enable this strategy, the framework applies Deep Reinforcement Learning through Proximal Policy Optimization with actor-critic architecture, anomaly detection via Isolation Forest algorithms, and multi-objective optimization techniques to facilitate adaptive portfolio rebalancing across varying market regimes.

## 1.2 Literature Review

The literature relevant to this study encompasses three primary areas: derivative investment strategies, portfolio optimization techniques, and machine learning applications in financial markets.

Although Covered Call ETFs represent a relatively recent innovation, the underlying Covered Call strategy enjoys longstanding popularity. Israelov and Nielsen (2018) examined the S&P 500 Covered Call strategy's performance from April 1996 to December 2013, discovering it produced 5% annualized ER while surpassing both long equity and short straddle approaches.

Constructing portfolios that effectively combine multiple investment approaches requires sophisticated optimization frameworks. Markowitz's (1952) pioneering work in modern portfolio theory catalyzed the emergence of numerous linear programming techniques designed for optimal portfolio allocation. Rudolf et al. (1999) found linear deviation best represents investor risk profiles, while Crama and Schyns (2003) applied simulated annealing algorithms for near-global portfolio optimization. Maringer and Kellerer (2003) introduced a hybrid metaheuristic combining simulated annealing and evolutionary strategies for efficient portfolio optimization with limited assets and non-negative weights, enabling multi-objective strategies. Subsequently, Wu et al. (2007) developed a novel goal programming framework with dual-objectives tailored for Taiwan market. Canakgoz and Beasley (2009) implemented regression-based methodologies to establish three-stage index tracking alongside two-stage enhanced indexation models. Guastaroba and Speranza (2012) introduced the Kernel Search methodology to reduce computational demands in MILP applications for index-tracking problems. De Paulo et al. (2016) developed a computationally efficient Lagrange multiplier framework for single-period enhanced indexation that balances TE and ER under cardinality constraints.

Time series is used as part of the feature engineering process to enhance the input state for the DRL model. Anomaly detection, alongside Vector Autoregression (VAR), plays a critical role in my portfolio optimization framework. Sims (1980) established VAR methodology for modeling multiple time series interactions, while Liu et al. (2008) developed the isolation forest algorithm, which offers efficient anomaly detection in large financial datasets.

Moreover, machine learning's expanding role in finance has driven adoption of DRL. Neuneier (1995) spearheaded reinforcement learning applications for asset allocation in German equity markets. Yu et al. (2019), Wu et al. (2020), and Chaouki et al. (2020) utilized DRL architectures to enable dynamic allocation strategies. Su et al. (2024) deployed multi-module DRL systems targeting risk management and profit enhancement. Cui et al. (2024) proposed a DRL hyper-heuristic framework for multi-period portfolio optimization that searches for trading strategies rather than actions, demonstrating superior performance from 2012 to 2022.

To evaluate risk-adjusted performance and risk exposure in this study, several key metrics are employed. Sharpe (1966) introduced the Sharpe ratio, which measures portfolio performance by dividing the excess return over the risk-free rate by the total standard deviation of returns. Sortino and Price (1994) refined this approach by focusing solely on downside risk. This study also incorporates Value at Risk (VaR), popularized by J.P. Morgan's RiskMetrics framework (1994), and Conditional Value at Risk (CVaR), proposed by Rockafellar and Uryasev (2000). CVaR quantifies the expected severity of losses beyond the VaR threshold, offering a more comprehensive assessment of extreme risks.

### 1.3 Background on the benchmark

QQQ, a passively managed ETF launched by Invesco on March 10, 1999, tracks the Nasdaq-100 Index, encompassing 100 prominent Nasdaq-listed companies. Rebalanced quarterly and reconstituted annually, it ranks as the second-most traded ETF in the U.S. by average daily volume as of March 31, 2025.

### 1.4 Background on the portfolio constituents

QYLD, launched on December 12, 2013 by Global X, employs a "covered call" strategy. It invests at least 80% in Nasdaq-100 securities while selling one-month at-the-money call options for premium income. Tracking the CBOE NASDAQ-100 BuyWrite V2 (BXNT) Index, it prioritizes yield over capital appreciation, limiting upside in bull markets. YQQQ, launched on August 14, 2024 by YieldMax, is an actively managed ETF that generates monthly income by selling Nasdaq-100 call options, buying put options, and holding US treasuries. It provides inverse exposure to the Nasdaq-100 with capped gains. TQQQ, launched on February 9, 2010 by ProShares, is a leveraged ETF targeting three times (3X) the daily Nasdaq-100 Index performance using derivatives. It is designed for short-term traders seeking amplified exposure to the Nasdaq-100.

## 2. Setting the Stage

### 2.1 Computational Environment

The main computation runs on a server with eight NVIDIA GeForce RTX 3090 GPUs. The portfolio optimization framework is implemented using several Python packages. Data manipulation utilizes `pandas` for ETF price data handling and log return computation, and `numpy` for numerical operations. The DRL framework employs `PyTorch` to implement the PPO agent with Actor-Critic architecture and ridge-regularized VAR models. Anomaly detection leverages the `IsolationForest` algorithm from `scikit-learn` to identify market regime changes that trigger dynamic portfolio rebalancing. Visualization is provided by `matplotlib` for plotting portfolio allocations and performance results.

## 2.2 Methodology

The PPO-based DRL algorithm with VAR and Isolation Forest optimizes NASDAQ 100 derivative ETF portfolios by generating dynamic, uneven weights that account for each ETF's distinct characteristics. PPO handles nonlinear market dynamics through its policy gradient framework without restrictive distributional assumptions, while VAR models capture temporal dependencies and cross-asset correlations, and Isolation Forest detects market shifts to trigger timely rebalancing during volatile periods. This integrated approach outperforms alternatives like Mean Variance Optimization with its Gaussian return assumptions that fail during market stress, Mixed Integer Linear Programming with its linear formulation limitations, autoencoders designed for feature extraction, and traditional Q-learning methods unsuitable for continuous action spaces. The methodology dynamically optimizes risk-return trade-offs through continuous adaptation rather than relying on static distributions or simplistic market assumptions.

Initially, to address incomplete data, an OLS regression model was employed to estimate missing returns using QQQ as the benchmark index, ensuring that all ETFs have return data for analysis starting from January 5, 2010. The analysis employs a sequential time-based split: 2010-2018 data for model training, 2019-2023 for validation and parameter selection, and 2024 through June 2025 for out-of-sample testing. Subsequently, VAR modeling generates predictive features that encapsulate temporal dependencies within financial datasets, thereby strengthening the DRL agent's decision-making capabilities. The model determines the optimal number of past periods to consider by balancing accuracy and complexity using the Bayesian Information Criterion (BIC) by Schwarz (1978), with regularization applied to prevent overfitting. These prediction errors and their trends serve as key features for market anomaly detection.

Complementing the predictive modeling, Isolation Forest anomaly detection identifies critical market shifts by analyzing multiple features, including market volatility and prediction errors, to detect significant deviations from normal market behavior. This method triggers portfolio adjustments at key dates alongside regular monthly rebalancing, enhancing the portfolio's responsiveness to changing market conditions.

Building upon these analytical components, the PPO-based DRL agent synthesizes predictive signals and market anomaly insights to achieve optimal risk-return equilibrium. The portfolio optimization aims to balance TE minimization and ER maximization relative to the QQQ benchmark. The formal mathematical formulation of this optimization problem is detailed in Section 2.3. The objective function optimized by the PPO agent is formulated as $(1 - \lambda) \times \mathrm{ER} - \lambda \times \mathrm{TE}$, where the tracking weight parameter $\lambda$ controls this trade-off, allowing flexible prioritization between minimizing TE and maximizing ER. Values of $\lambda$ closer to 1 prioritize tight benchmark tracking, while lower values emphasize outperformance, enabling adaptation to different investor risk preferences.

Finally, the reinforcement learning approach generates probabilistic portfolio allocations based on market features, with actions sampled to produce valid portfolio weights that comply with constraints. The algorithm iteratively refines these allocations by optimizing a loss function that balances performance, stability, and exploration. The tracking weight parameter $\lambda$ is dynamically adjusted using validation data to optimize performance, with early stopping employed when improvements stall to prevent overfitting. The complete algorithmic implementation of this PPO-based portfolio optimization framework is detailed in Section 2.4.

## 2.3 Main Optimization Problem Formulation

The main optimization problem seeks to maximize utility $U(w)$ defined as:

$$\max_{w} U(w) = (1-\lambda)ER(w) - \lambda TE(w) \tag{1}$$

Subject to:

$$\text{Full investment:} \quad \sum_{i=1}^{n} w_i = 1 \tag{2}$$

$$\text{Cardinality constraint:} \quad K_{\min} \leq \sum_{i=1}^{n} z_i \leq K_{\max} \tag{3}$$

$$\text{Bounded weights:} \quad \alpha_i z_i \leq w_i \leq b_i z_i, \quad i = 1, \ldots, N \tag{4}$$

Where:

$$TE(w) = \sqrt{\frac{D}{P} \|R_p - R_b\|_2^2} \tag{5}$$

$$ER(w) = \frac{1}{P}(R_p - R_b)^T e \tag{6}$$

Parameters:

- $\lambda \in [0, 1]$: Scalar balancing trade-off between ER and TE
- $e$: Vector of ones; Size $P \times 1$ used to sum excess returns across all periods
- $D$: Number of trading days
- $P$: Number of periods in optimization window
- $K_{\min}, K_{\max}$: Bounds on number of selected assets
- $\alpha_i$: Minimum weight for asset $i$
- $w_i$: Weight of asset $i$
- $b_i$: Maximum weight for asset $i$
- $z_i \in [0, 1]$: Binary selection variable indicating whether asset $i$ is selected

## 2.4 Algorithm

**Algorithm 1** Deep Reinforcement Learning Portfolio Optimization with Dynamic Rebalancing

**Require:** Asset returns $R \in \mathbb{R}^{T \times n}$

Train VAR(p) model on $R$; compute prediction errors $\varepsilon_t$
Apply Isolation Forest on market features; identify anomaly dates $D_{\text{anomaly}}$
Generate rebalancing schedule $T = \{\text{monthly dates} \cup D_{\text{anomaly}}\}$, min days apart
Initialize actor $\pi_\theta : \mathbb{R}^d \to \mathbb{R}^n$, critic $V_\phi : \mathbb{R}^d \to \mathbb{R}$ networks
**for** episode $e = 0, 1, \ldots, E$ **do**
  Initialize experience buffer $\mathcal{B} = \emptyset$
  **for** each rebalancing date $t \in T$ **do**
    Construct state $s_t = [R_{\text{window}}, \text{market\_indicators}, \varepsilon_t] \in \mathbb{R}^d$
    $\mu_t, \sigma_t, V_t \leftarrow \pi_\theta(s_t), V_\phi(s_t)$
    Sample $w_{\text{raw}} \sim \mathcal{N}(\mu_t, \sigma_t)$; compute $\log \pi_t \leftarrow \log \pi_\theta(w_{\text{raw}} | s_t)$
    $w_t \leftarrow \text{Project Constraints}(w_{\text{raw}}, \text{max\_weight}, \text{min\_stocks}, \text{max\_stocks})$
    $r_t \leftarrow (1 - \lambda) \times \text{Excess Return}(w_t, R, B) - \lambda \times \text{Tracking Error}(w_t, B)$
    **if** $t \le T_{\text{train}}$ **then**
      Store $(s_t, w_t, \log \pi_t, r_t, V_t)$ in buffer $\mathcal{B}$
    **end if**
  **end for**
  Compute advantages $A_t \leftarrow r_t - V_t$; normalize $A_t$
  **for** update $u = 1, \ldots, 10$ **do**
    Sample mini-batch from $\mathcal{B}$; compute $\rho_t \leftarrow \pi_\theta(w_t|s_t)/\pi_{\theta_{\text{old}}}(w_t|s_t)$
    $\mathcal{L}_{\text{actor}} \leftarrow -\mathbb{E}[\min(\rho_t A_t, \text{clip}(\rho_t, 1 - \varepsilon, 1 + \varepsilon) A_t)]$
    $\mathcal{L}_{\text{critic}} \leftarrow \mathbb{E}[(V_t - r_t)^2]$; $\mathcal{H} \leftarrow -\mathbb{E}[\pi_\theta \log \pi_\theta]$
    Update $\theta, \phi$ with $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{actor}} + 0.5 \times \mathcal{L}_{\text{critic}} - 0.05 \times \mathcal{H}$
  **end for**
  **for** $\lambda \in \{\lambda_k\}$ **do**
    Evaluate on validation set $T_{\text{val}}$; compute $R_{\text{val}}(\lambda)$
  **end for**
  $\lambda_{\text{best}} \leftarrow \arg\max_\lambda R_{\text{val}}(\lambda)$; store best weights if improved
  **if** no improvement for 20 episodes **then**
    **break**
  **end if**
**end for**
$W^* \leftarrow$ best stored weights
Apply final constraints: $W^*[\text{inactive}] \leftarrow 0.01$; renormalize to sum = 1
**if** $\|W^*\|_0 < \text{min\_stocks}$ **then**
  force select additional stocks
**end if**
**if** $\|W^*\|_0 > \text{max\_stocks}$ **then**
  trim lowest weighted positions
**end if**
**return** optimized portfolio weights $W^*$

# 3. Results

## 3.1 Impact of parameters on ER and TE

Figures 1 through 5 illustrate how adjusting key parameters of contamination, window size, max stocks, min stocks, tracking weight, and minimum days for rebalancing affects the portfolio's ability to outperform QQQ while minimizing deviation relative to QQQ across training, validation, and testing periods. The baseline configuration serves as the analytical foundation, featuring a tracking weight of 0.25 that balances ER and TE, a maximum stock weight of 0.7, a minimum stock weight of 0.01, a contamination level of 0.01 for anomaly detection, a 30-day look-back window, a maximum of 4 assets, a minimum of 1 asset, 1000 training episodes, and a minimum rebalancing interval of 30 days. Each parameter reveals a distinct performance pattern when compared to the baseline configuration.

Figure 1 depicts the impact of tracking weights on the portfolio's ER and TE relative to QQQ. Tracking weights shape the balance between achieving higher ER and minimizing TE by adjusting the emphasis on returns versus deviation. In the training period, extremely low tracking weights achieved notably strong ER, while the validation period showed improved performance with higher tracking weights. Out-of-sample testing indicated that moderate tracking weights were most effective for both ER and TE management.

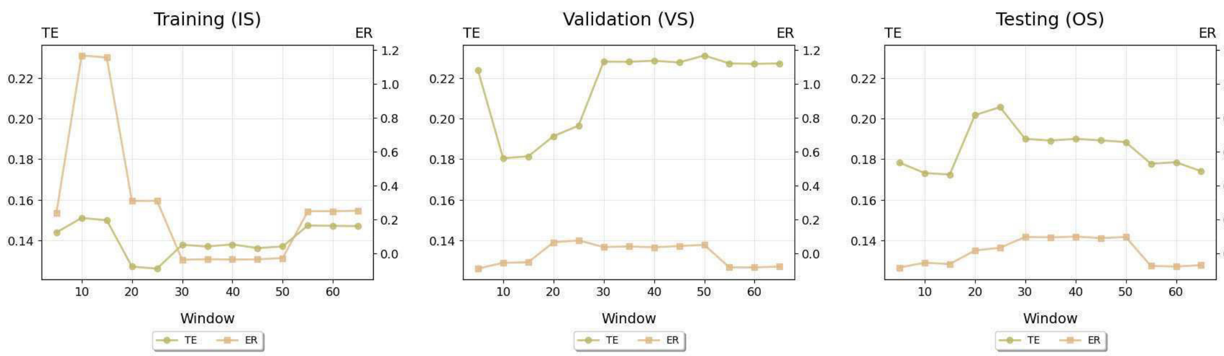


Figure 1: The effect of tracking weight on ER and TE

Figure 2 unveils the impact of contamination levels on the portfolio's ER and TE relative to QQQ. Contamination levels shape the balance between achieving higher ER and minimizing TE by adjusting sensitivity to market anomalies. The analysis reveals notable variations in optimal contamination settings across different time periods. In the training period, higher contamination levels generally achieved stronger ER, while the validation period showed improved performance with lower contamination settings. Out-of-sample testing indicated that the highest contamination level was most effective, achieving both superior ER and optimal TE control.

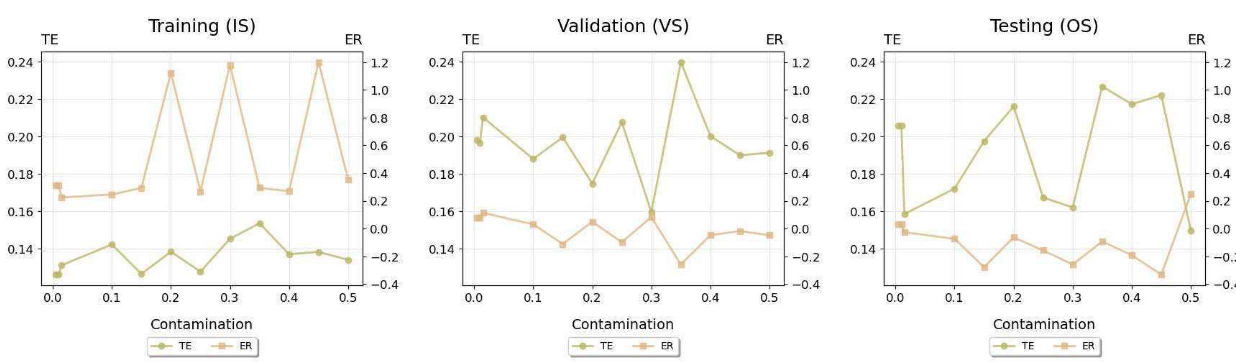


Figure 2: The effect of contamination on ER and TE

Figure 3 exhibits the impact of window sizes on the portfolio's ER and TE relative to QQQ. Window sizes shape the balance between achieving higher ER and minimizing TE by adjusting portfolio responsiveness to market trends. The analysis reveals notable differences in optimal window configurations across time periods. In the training period, smaller window sizes achieved substantially higher ER, while the validation and testing periods showed improved performance with moderate window sizes. Out-of-sample analysis confirmed that moderate window sizes were most effective, achieving superior ER compared to both smaller and larger windows.

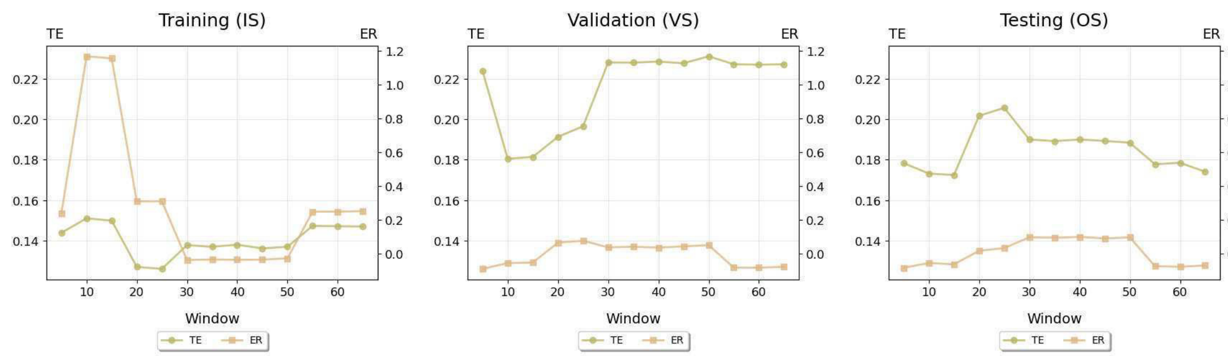


Figure 3: The effect of window size on ER and TE

Figure 4 displays the impact of maximum weights on the portfolio's ER and TE relative to QQQ. Maximum weights shape the balance between achieving higher ER and minimizing TE by constraining asset allocations. The analysis reveals complex relationships between weight constraints and performance across different periods. In the training period, lower maximum weights achieved substantially higher ER while also maintaining better TE control. The validation and testing periods showed different patterns, with higher maximum weights generally improving ER but at the cost of significantly elevated TE.

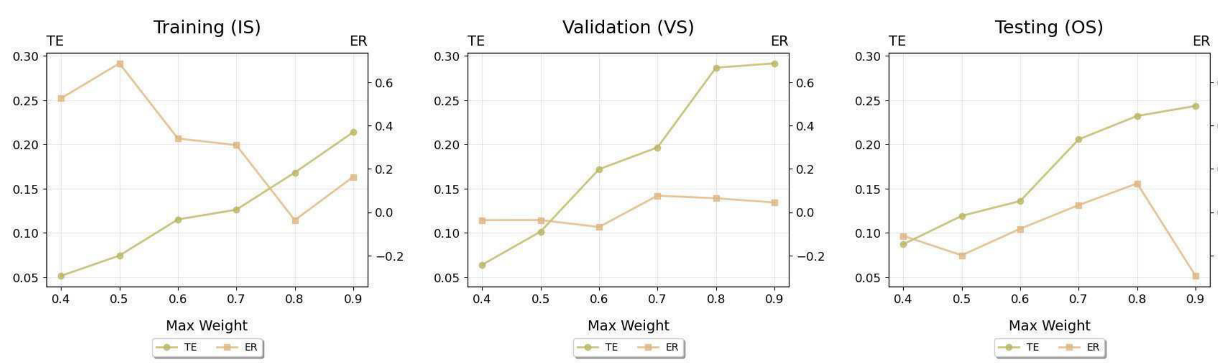


Figure 4: The effect of max weight on ER and TE

Figure 5 represents the impact of minimum days between rebalancing on the portfolio's ER and TE relative to QQQ. Minimum days shape the balance between achieving higher ER and minimizing TE by controlling rebalancing frequency. In the training period, different rebalancing frequencies showed varying performance patterns, with certain frequencies achieving favorable returns. The validation period exhibited different market dynamics with more modest performance across rebalancing frequencies. Out-of-sample testing indicated that moderate rebalancing frequencies provided reasonably balanced performance, achieving acceptable ER while maintaining suitable TE levels.

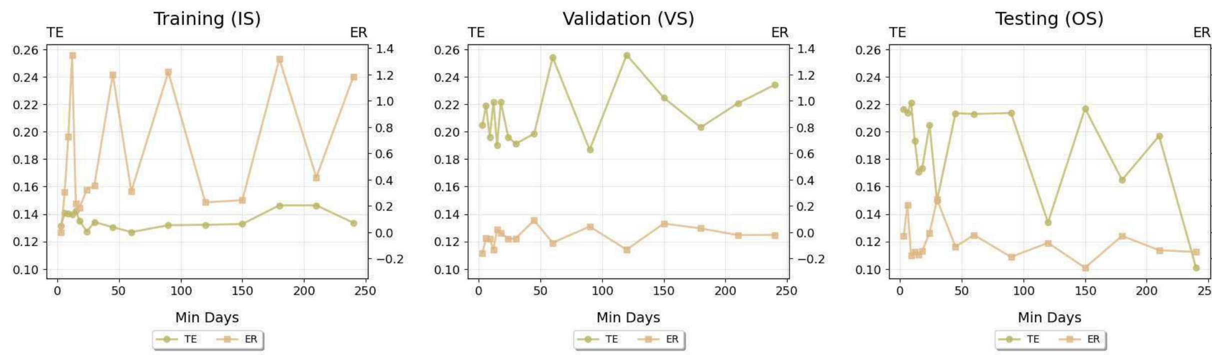


Figure 5: The effect of minimum days between rebalancing on ER and TE

### 3.2 Allocation

Figure 6 shows the allocation of the portfolio's investments across three assets, TQQQ, QYLD, and YQQQ, over time.

During the training period, QYLD dominates the portfolio with an average allocation of 40.0%, reflecting an income focused investment strategy during this foundational period. TQQQ maintains a substantial 35.8% average allocation, positioning the portfolio for growth opportunities. YQQQ secures a meaningful 24.2% allocation, demonstrating consistent diversification rather than minimal exposure.

During the validation period, the portfolio undergoes a strategic shift toward growth orientation. TQQQ becomes the dominant asset with 42.5% average allocation, overtaking QYLD as market conditions favor aggressive growth strategies. QYLD's allocation decreases to 33.8% but remains substantial, continuing to provide income stability. YQQQ maintains consistency at 23.7% average allocation, proving its value as a portfolio stabilizer.

During the testing period, the portfolio adopts a more balanced approach with TQQQ averaging 38.1%, QYLD 35.3%, and YQQQ 26.6%. This period exhibits notable tactical adjustments, with frequent 70% allocations to single assets during specific market conditions, indicating active management responses to opportunities. YQQQ achieves its highest relative importance during this period, reflecting refined portfolio optimization.

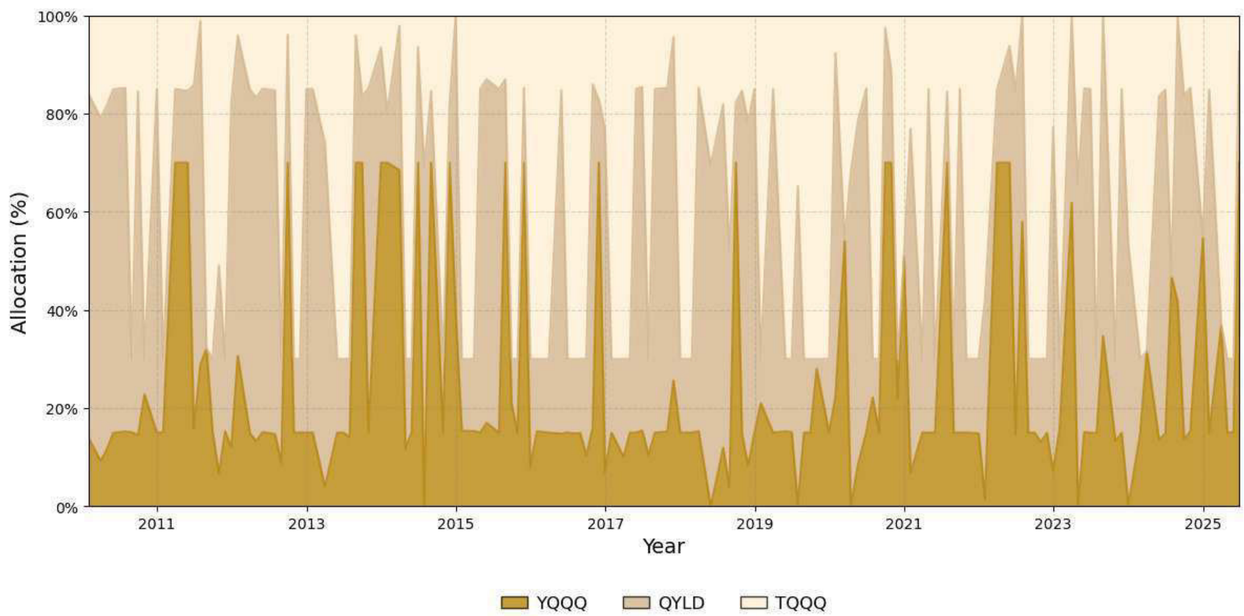


Figure 6: Portfolio allocation over time

### 3.3 Performance

Figure 7 reveals the portfolio's dividend-adjusted performance from 2010 to 2025 compared to QQQ. The portfolio incorporates dividends from all ETF holdings to adjust its returns. The benchmark QQQ provides annual dividends of $2.8628 per share through quarterly payments of $0.7157, while the portfolio receives annual dividends of $3.18 per share from YQQQ through monthly payments of $0.265, $1.98 per share from QYLD through monthly payments of $0.165, and $0.7909 per share from TQQQ through quarterly payments of $0.19773. The risk-free rate, referenced from SOFR, used for calculating the Sharpe and Sortino ratios is arbitrarily set at 4.29% annually. All performance metrics are calculated on a pre-tax and pre-transaction-cost basis to ensure unbiased and standardized evaluation.

The portfolio demonstrates strong performance during the training period, achieving an exceptional ER of 30.96% with a controlled TE of 12.62%. Risk-adjusted metrics are particularly impressive, with a Sharpe ratio of 2.34 significantly outperforming QQQ's 0.88. The Sortino ratio of 2.40 versus QQQ's 1.12 further demonstrates superior risk-adjusted performance when considering only downside deviations. The 95% daily CVaR of 3.12% remains only marginally higher than QQQ's 2.69%, indicating well-managed tail risk during optimization.

The validation period unveils notable performance deterioration, with ER declining to 7.60% and TE rising substantially to 19.65%. Risk-adjusted performance metrics decline substantially, with the Sharpe ratio dropping to 0.80 compared to QQQ's 0.67, while the Sortino ratio falls to 0.93 versus QQQ's 0.85. The 95% daily CVaR increases to 4.99% compared to QQQ's 3.85%, reflecting heightened tail risk exposure as market conditions diverge from training patterns.

The testing period shows continued challenges, with ER diminishing to 3.25% and TE reaching 20.56%. Most notably, risk-adjusted performance deteriorates below benchmark levels for the first time, with the Sharpe ratio declining to 0.55 versus QQQ's 0.70 and the Sortino ratio falling to 0.61 compared to QQQ's 0.92. This marks a fundamental shift where the portfolio underperforms on both total volatility-adjusted and downside deviation-adjusted metrics. The 95% daily CVaR deteriorates further to 5.05% versus QQQ's 3.34%, indicating that the portfolio's tail risk exposure has increased significantly relative to the benchmark.

Table 1: Portfolio Performance Summary Across All Periods

| | **Training** (2010-2019) | | **Validation** (2019-2024) | | **Testing** (2024-2025) | |
|---|---|---|---|---|---|---|
| **Metric** | **Portfolio** | **QQQ** | **Portfolio** | **QQQ** | **Portfolio** | **QQQ** |
| Annual Return (%) | 50.50 | 19.54 | 29.01 | 21.41 | 23.47 | 20.22 |
| Excess Return (%) | 30.96 | – | 7.60 | – | 3.25 | – |
| Annual Volatility (%) | 19.79 | 17.42 | 30.78 | 25.49 | 34.62 | 22.61 |
| Tracking Error (%) | 12.62 | – | 19.65 | – | 20.56 | – |
| Sharpe Ratio | 2.34 | 0.88 | 0.80 | 0.67 | 0.55 | 0.70 |
| Sortino Ratio | 2.40 | 1.12 | 0.93 | 0.85 | 0.61 | 0.92 |
| Maximum Drawdown (%) | −20.68 | −22.74 | −51.99 | −37.78 | −34.91 | −23.42 |
| Beta | 0.88 | 1.00 | 0.93 | 1.00 | 1.26 | 1.00 |
| 95% Daily VaR (%) | 1.59 | 1.80 | 3.20 | 2.59 | 2.83 | 2.40 |
| 99% Daily VaR (%) | 3.94 | 3.20 | 6.18 | 4.38 | 5.57 | 3.75 |
| 95% Daily CVaR (%) | 3.12 | 2.69 | 4.99 | 3.85 | 5.05 | 3.34 |
| 99% Daily CVaR (%) | 5.65 | 4.00 | 7.67 | 6.11 | 10.61 | 5.05 |

Note: ER and TE are not applicable (–) for QQQ as it serves as the benchmark. All metrics calculated using dividend-adjusted returns.

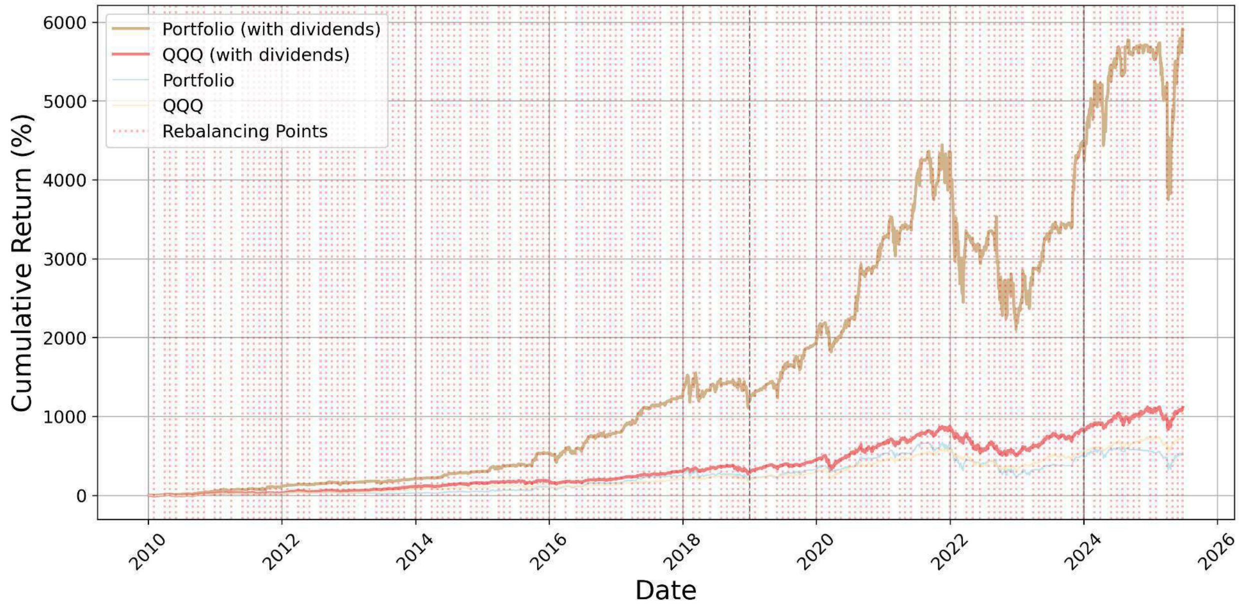


Figure 7: Cumulative Returns Comparison: Portfolio vs QQQ Benchmark (2010-2025)

# 4. Conclusion

The reinforcement learning-based portfolio optimization framework successfully constructs a portfolio of NASDAQ-100 derivative ETFs including YQQQ, QYLD, and TQQQ that generates consistent positive excess returns relative to the QQQ benchmark across all evaluation periods. Nevertheless, the temporal analysis reveals a clear deterioration pattern: while the strategy maintains positive excess returns, both tracking error and tail risk metrics progressively worsen, and risk-adjusted performance deteriorates from substantial outperformance during training to underperformance during testing periods. This progression suggests that the portfolio's risk profile becomes increasingly unfavorable as it encounters market conditions different from the training environment.

The framework demonstrates the potential of DRL for multi-objective portfolio optimization, effectively balancing efficiency through TQQQ, income generation via QYLD, and hedging through YQQQ. However, the observed parameter instability across time periods indicates that future research should implement more robust validation methodologies, such as rolling window cross-validation, to improve generalization capability and reduce overfitting risks in evolving market conditions.

# References

Canakgoz, N. A. and Beasley, J. E. (2009) "Mixed-integer programming approaches for index tracking and enhanced indexation" European Journal of Operational Research 196, 384-399.

Chaouki, A., Hardiman, S., Schmidt, C., Série, E., and Lataillade, J. D. (2020) "Deep deterministic portfolio optimization" The Journal of Finance and Data Science 6, 16-30.

Crama, Y. and Schyns, M. (2003) "Simulated annealing for complex portfolio selection problems"

European Journal of Operational Research 150, 546-571.
Cui, T., Du, N., Yang, X., and Ding, S. (2024) "Multi-period portfolio optimization using a deep reinforcement learning hyper-heuristic approach" Technological Forecasting & Social Change 199, 1-11.
de Paulo, W. L., de Oliveira, E. M., and do Valle Costa, O. L. (2016) "Enhanced index tracking portfolio selection" Finance Research Letters 16, 93-102.
Guastaroba, G. and Speranza, M. G. (2012) "Kernel Search: An application to the index tracking problem" European Journal of Operational Research 217, 54-68.
Israelov, R. and Nielsen, L. N. (2018) "Covered Call Strategies: One Fact and Eight Myths" Financial Analysts Journal 70, 23-31.
J.P. Morgan/Reuters (1994) RiskMetrics—Technical Document, J.P. Morgan Technical Report: New York.
Liu, F. T., Ting, K. M., and Zhou, Z. H. (2008) "Isolation Forest" Eighth IEEE International Conference on Data Mining, 413-422.
Maringer, D. and Kellerer, H. (2003) "Optimization of cardinality constrained portfolios with a hybrid local search algorithm" OR Spectrum 25, 481-495.
Markowitz, H. M. (1952) "Portfolio Selection" Journal of Finance 7, 77-91.
Neuneier, R. (1995) "Optimal asset allocation using adaptive dynamic programming" Advances in Neural Information Processing Systems 8, 952-958.
Rockafellar, R. T. and Uryasev, S. P. (2000) "Optimization of Conditional Value-at-Risk" Journal of Risk 2, 21-42.
Rudolf, M., Wolter, H. J., and Zimmermann, H. (1999) "A linear model for tracking error minimization" Journal of Banking & Finance 23, 85-103.
Schwarz, G. (1978) "Estimating the dimension of a model" The Annals of Statistics 6, 461-464.
Sharpe, W. F. (1966) "Mutual Fund Performance" The Journal of Business 39, 119-138.
Sims, C. A. (1980) "Macroeconomics and Reality" Econometrica 48, 1-48.
Sortino, F. A. and Price, L. N. (1994) "Performance Measurement in a Downside Risk Framework" Journal of Investing 3, 59-64.
Su, R., Chi, C., Tu, S., and Xu, L. (2024) "A deep reinforcement learning approach for portfolio management in non-short-selling market" IET Signal Processing 2024, 1-9.
Wu, L. C., Chou, S. C., Yang, C. C., and Ong, C. S. (2007) "Enhanced index investing based on goal programming" Journal of Portfolio Management 33, 49-56.
Wu, X., Chen, H., Wang, J., Troiano, L., Lois, V., and Fujita, H. (2020) "Adaptive stock trading strategies with deep reinforcement learning methods" Information Sciences 538, 142-158.
Yu, P., Lee, J. S., Kulyatin, H., Shi, Z., and Dasgupta, S. (2019) "Model-based deep reinforcement learning for dynamic portfolio optimization" ArXiv preprint 1901.08740.