



Cross-Asset and Macro-Driven G10 FX Relative Value Strategy

Jay Lim
Aryan Sunilkumar
Diaz Ong
Shen Siyin

NUS Investment Society
Quantitative Finance Department
2026

Abstract

This paper studies the effects of macroeconomic and cross-asset features on relative movements in G10 foreign exchange markets and how it translates into a profitable trading strategy. Rather than forecasting absolute currency direction, we adopt a cross-sectional framework that ranks currencies by their expected relative strength against the U.S. dollar.

We construct a feature set grounded in financial and macroeconomic theory, including relative equity performance, momentum, relative volatility, oil-market variables, spillover effects, and crisis-safe-haven dynamics. These features are used to train a Random Forest classifier that outputs daily probabilities of currency appreciation against the U.S. dollar. The probabilities are then transformed into conviction scores and embedded within a systematic long-short portfolio construction framework.

Our findings suggest that while the classifier's standalone directional accuracy is modest, the model contains useful ranking information. When combined with portfolio construction techniques such as smoothing, hysteresis, inverse-volatility weighting, volatility targeting, turnover control, dispersion gating, and a panic overlay, the resulting strategy outperforms a buy-and-hold benchmark on a risk-adjusted basis. The results highlight that in foreign exchange, weak predictive signals can still be economically valuable when used as a relative-value ranking engine.

Contents

1	Introduction	3
2	Technical Concepts	3
2.1	Foreign Exchange as a Relative Value Market	3
2.2	Feature Construction	3
3	Implementation and Testing	6
3.1	Data Preparation and Pipeline Structure	6
3.2	Feature Engineering Implementation	7
3.3	Model Training Procedure	7
3.4	Model Evaluation	7
3.5	Signal Construction	8
3.6	Trading Strategy Implementation	8
3.7	Portfolio Construction	8
3.8	Transaction Costs and Turnover Constraints	9
3.9	Backtesting Framework	9
3.10	Performance Metrics	10
4	Critical Assessment	10
4.1	Model Performance	10
4.2	Interpretation of the Top Features	12
4.3	From Model to Strategy	13
4.4	Strategy Performance	13
4.5	Strengths of the Framework	15
4.6	Limitations	15
4.7	Potential Improvements	16
5	Conclusion	16
6	References	17

1 Introduction

Foreign exchange markets are widely regarded as noisy and difficult to forecast in absolute terms. Exchange rates react to monetary policy, trade conditions, global risk sentiment, geopolitical shocks, and portfolio flows, making outright directional prediction highly challenging. For this reason, many practical FX strategies focus not on absolute direction but on relative value. This paper studies whether macro and cross-asset information can be used to rank G10 currencies well enough to construct a profitable long-short FX basket.

2 Technical Concepts

2.1 Foreign Exchange as a Relative Value Market

Foreign exchange (FX) markets are inherently relative. A currency pair such as EUR/USD expresses the value of one currency in terms of another, rather than in absolute terms. For that reason, FX prediction is often more naturally framed as a *relative value* problem than as an absolute one. This is also consistent with empirical literature on currency strategies, where cross-sectional approaches such as momentum have been shown to be more robust than simple absolute forecasts (2).

In this project, the prediction problem is defined at the currency level. Let the log return of currency i at time t be

$$r_{i,t} = \log \left(\frac{P_{i,t}}{P_{i,t-1}} \right),$$

where $P_{i,t}$ denotes the relevant FX rate. The target variable is then defined as

$$Y_{i,t} = \begin{cases} 1, & \text{if } r_{i,t+1} > 0, \\ 0, & \text{otherwise.} \end{cases}$$

This converts the problem into a binary classification task. At the same time, the downstream trading layer remains cross-sectional: the model is ultimately used to rank currencies against one another, instead of making isolated directional calls.

2.2 Feature Construction

The feature set is designed around several economic channels through which capital flows, relative growth prospects, energy shocks, and market stress can influence exchange rates. The intention is to include variables that are economically interpretable and supported by prior research.

Relative Equity Returns We proxy relative growth and portfolio-attractiveness effects using the performance of local equity markets relative to the U.S. benchmark:

$$\text{RelRet}_{i,t} = \log \left(\frac{S_{i,t}^{\text{local}}}{S_{i,t-1}^{\text{local}}} \right) - \log \left(\frac{S_t^{\text{US}}}{S_{t-1}^{\text{US}}} \right),$$

where $S_{i,t}^{\text{local}}$ is the local equity index for country i , and S_t^{US} is the S&P 500.

The intuition is such that if foreign equities outperform U.S. equities, portfolio capital may be reallocated toward the foreign market, increasing demand for the domestic currency. This is closely related to the mechanism discussed by prior research (1), which shows that exchange rates and stock prices co-move through global stock-factor exposures and portfolio rebalancing. In that sense, relative equity performance also provides a flow-based signal.

Momentum Momentum is constructed over multiple horizons to capture persistence in relative performance:

$$\text{Mom}_{i,t}^{(k)} = \frac{S_{i,t}^{\text{local}}}{S_{i,t-k}^{\text{local}}} - \frac{S_t^{\text{US}}}{S_{t-k}^{\text{US}}}, \quad k \in \{21,60,252\}.$$

These horizon choices allow the model to capture short-, medium-, and longer-run trend effects. The inclusion of momentum is directly supported by prior studies (2), which documents that currencies with high past returns tend to continue outperforming currencies with low past returns, and that such returns are not fully explained by conventional risk factors. Thus, this makes momentum especially relevant in a cross-sectional FX setting.

Relative Volatility Relative riskiness is measured using rolling volatility ratios. For a given horizon k ,

$$\sigma_{i,t}^{(k)} = \sqrt{\text{Var}(r_{i,t-k:t})},$$

and the relative volatility feature is defined as

$$\text{RelVol}_{i,t}^{(k)} = \frac{\sigma_{i,t}^{\text{local},(k)}}{\sigma_t^{\text{US},(k)}}.$$

The rationale is that a local market that is substantially more volatile than the U.S. market may be perceived as riskier, weakening the attractiveness of its assets and, by extension, its currency. This suggests exchange rate volatility is economically meaningful and negatively related to stock returns, especially in internationally exposed sectors (3). In our context, the local-to-U.S. ratio makes this viable in a relative framework.

Multi-Market Spillover To capture broader global transmission effects, we construct a multi-market spillover feature using beta-type exposures. In simplified form, for market m ,

$$\beta_{i,m,t} = \frac{\text{Cov}(r_{i,t}^{\text{FX}}, r_{m,t}^{\text{mkt}})}{\text{Var}(r_{m,t}^{\text{mkt}})},$$

and the weighted spillover measure is based on beta-scaled market returns, further weighted by explanatory fit:

$$\text{Spillover}_{i,t} \propto \sum_m R_{i,m,t}^2 \beta_{i,m,t} r_{m,t}^{\text{mkt}}.$$

This feature is motivated by two related ideas. First, research (1) emphasises that exchange rate dynamics are closely linked to cross-market co-movement and portfolio reallocation across countries. Second, studies done (4) provide a broader multi-beta framework in which returns are driven by time-varying exposures to multiple global risk factors. Thus, we ensure currencies are not driven only by domestic conditions, but also by the extent to which they are exposed to common market shocks.

Brent Returns and Oil-Linked Features Oil is relevant because its macroeconomic effects differ sharply between exporters and importers. Daily Brent returns are defined as

$$r_t^{\text{oil}} = \log\left(\frac{O_t}{O_{t-1}}\right),$$

where O_t denotes the Brent crude oil price. Oil price movements affect exchange rates through several channels, including terms-of-trade adjustments, wealth transfers, and portfolio flows (5). These channels are especially important for commodity-linked currencies such as CAD and AUD.

We therefore include oil-linked features that allow the model to reflect this heterogeneity. One such feature is a direct oil exposure interaction:

$$\text{OilExposure}_{i,t} = r_t^{\text{oil}} \cdot \beta_i^{\text{oil}},$$

where β_i^{oil} is a currency-specific sensitivity parameter. Another feature is an export sensitivity index that uses trade data to scale oil shocks by a country's net external oil position:

$$\text{TradeRatio}_t = \frac{\text{Exports}_t - \text{Imports}_t}{\text{Total Trade}_t},$$

$$\text{ESI_Oil_Beta}_t = r_t^{\text{oil}} \cdot \text{TradeRatio}_t.$$

This construction is grounded in the trade-based logic that shows currencies of commodity exporters and importers respond differently to commodity-related shocks and that trade exposure is tied to currency risk premia (9).

Oil Spread Z-Score We also include the spread between WTI and Brent, standardised over a rolling window:

$$\text{Spread}_t = \text{WTI}_t - \text{Brent}_t,$$

$$\text{OilSpreadZ}_t = \frac{\text{Spread}_t - \mu_{21}(\text{Spread}_t)}{\sigma_{21}(\text{Spread}_t)}.$$

This feature is intended to capture dislocations in the oil market that may reflect transportation constraints, infrastructure bottlenecks, or shifts in relative U.S. energy conditions. Moreover, oil price differentials have undergone structural breaks over time, implying that spread levels should be interpreted relative to their prevailing regime rather than in raw historical terms (6). The rolling z-score formulation follows that logic and

allows the model to detect spread extremes within the current environment rather than against outdated historical benchmarks.

Precautionary Demand Shock To capture risk-driven oil shocks, we define a precautionary demand shock as

$$PDS_t = r_t^{\text{oil}} \cdot \Delta OVX_t,$$

where ΔOVX_t is the daily change in implied oil volatility. This interaction term is designed to distinguish ordinary price moves from those accompanied by sharp increases in uncertainty. Risk-related oil shocks also have stronger and more immediate effects on exchange rates than standard supply shocks, particularly at short horizons (7).

Refinery Utilization Proxy The refinery utilisation proxy is measured using year-on-year changes in a downstream industrial production series:

$$\text{RefineryYoY}_t = \text{Refinery}_t - \text{Refinery}_{t-252}.$$

This aims to capture downstream energy demand and potential bottlenecks in crude processing. Refinery utilisation is also closely related to oil price dynamics and can affect spot-market conditions through nonlinear supply-demand adjustments (8). In the FX context, changes in U.S. refinery capacity and margins can alter relative energy prices, trade balances, and ultimately dollar strength.

Safe-Haven Dynamics The strategy also accounts for the fact that currencies do not all respond symmetrically to crises. In particular, JPY and CHF are treated as safe-haven currencies. This is supported by recent literature, which finds that JPY exhibits the strongest safe-haven characteristics among major currencies, while CHF also retains defensive properties, though more modestly (10). More broadly, traditional safe-haven assets, remain effective stabilisers during recessionary and stress episodes (11).

3 Implementation and Testing

3.1 Data Preparation and Pipeline Structure

The implementation follows a structured pipeline beginning from raw data collection to final strategy evaluation. Daily price data for foreign exchange rates, equity indices, and commodity prices were obtained using the `yfinance` API, while macroeconomic variables such as trade flows were sourced from the Federal Reserve Economic Data (FRED) database.

All data were aligned to a common daily frequency. Lower-frequency macroeconomic variables were forward-filled to ensure consistency across time. Missing values were removed to prevent distortions in model training.

The final dataset was structured as a panel indexed by currency and date, allowing both cross-sectional and time-series relationships to be captured.

3.2 Feature Engineering Implementation

Feature engineering was implemented directly in code through vectorised operations over the dataset. For each currency, features were constructed based on relative comparisons against the U.S. benchmark.

Lagging was applied systematically to all features:

$$X_{i,t}^{\text{used}} = X_{i,t-1}$$

This ensures that only information available at time t is used to predict returns at $t + 1$, preventing look-ahead bias.

Each currency was processed independently and then concatenated into a unified panel dataset. Observations with insufficient lookback history were dropped to maintain consistency across features such as rolling volatility and momentum.

3.3 Model Training Procedure

The predictive model was implemented using a Random Forest classifier. The objective of the model is to estimate the probability:

$$P_{i,t} = \mathbb{P}(Y_{i,t+1} = 1 \mid X_{i,t})$$

where $X_{i,t}$ denotes the full feature vector. A Random Forest aggregates predictions across many trees:

$$\hat{P}_{i,t} = \frac{1}{B} \sum_{b=1}^B T_b(X_{i,t}),$$

where each T_b is a decision tree trained on a bootstrap sample.

To ensure realistic evaluation, a strict time-based split was applied. The dataset was divided into a training set (first 80% of dates) and a test set (remaining 20%). This preserves temporal ordering and avoids leakage of future information.

Within the training set, a walk-forward validation scheme was used. The training period was divided into sequential folds, where each fold expands the training window and evaluates performance on the next time segment.

Hyperparameters such as tree depth, number of estimators, and subsampling ratios were selected based on average validation log-loss across folds.

3.4 Model Evaluation

Model performance was evaluated on the out-of-sample test set using several classification metrics:

- Accuracy
- Log-loss
- Area Under the ROC Curve (AUC)

While accuracy was modest (approximately 53–55%), this is consistent with expectations in financial prediction tasks. Small predictive edges are sufficient when combined with appropriate portfolio construction.

The model outputs were transformed into probabilities, which serve as the primary input into the trading strategy.

3.5 Signal Construction

Predicted probabilities were converted into continuous signals:

$$\text{Conv}_{i,t} = 2 \cdot (P_{i,t} - 0.5)$$

This transformation produces a symmetric signal in the range $[-1,1]$, where magnitude reflects confidence and sign reflects direction.

To reduce noise, an exponential moving average (EMA) smoothing was applied:

$$\tilde{P}_{i,t} = \text{EMA}(P_{i,t})$$

This helps stabilise signals and reduces excessive turnover in the strategy.

3.6 Trading Strategy Implementation

The trading strategy was implemented as a cross-sectional long-short portfolio. At each rebalancing date, currencies were ranked based on their conviction scores.

Positions were constructed as follows:

- Long positions in the top 2-ranked currencies
- Short positions in the bottom 2-ranked currencies

To prevent excessive trading and noise sensitivity, hysteresis rules were applied:

- Positions are only entered if conviction exceeds a threshold
- Existing positions are maintained unless conviction weakens significantly

3.7 Portfolio Construction

Positions are taken by going long the top 2-ranked currencies and short the bottom 2-ranked ones. Let $w_{i,t}$ denote the weight assigned to currency i at time t . Weights are scaled using inverse volatility:

$$w_{i,t} \propto \frac{\text{Conv}_{i,t}}{\sigma_{i,t}},$$

and then normalised so that the long and short sides remain balanced.

Portfolio variance is given by

$$\sigma_p^2 = w_t^\top \Sigma_t w_t,$$

where Σ_t is the covariance matrix of returns. To maintain a stable risk budget, the weights are then rescaled to meet a fixed target volatility:

$$w_t \leftarrow w_t \cdot \frac{\sigma^*}{\sigma_p},$$

where σ^* is the chosen target volatility.

This combination of ranking, inverse-volatility weighting, and volatility targeting were done to reflect standard practice in systematic macro and relative-value trading.

3.8 Transaction Costs and Turnover Constraints

Transaction costs were incorporated to ensure realistic performance estimates. Costs were modelled as proportional to portfolio turnover:

$$\text{Cost}_t = c \cdot \sum_i |w_{i,t} - w_{i,t-1}|$$

where c is a constant representing trading costs.

Additionally, daily turnover was capped to avoid excessive rebalancing. When desired changes in weights exceeded this cap, positions were adjusted gradually.

3.9 Backtesting Framework

The strategy was evaluated using a fully out-of-sample backtest. At each time step:

1. Signals are generated using the trained model
2. Portfolio weights are constructed
3. Returns are realised over the next period
4. Transaction costs are deducted

Portfolio returns were computed as:

$$R_t = \sum_i w_{i,t} \cdot r_{i,t+1} - \text{Cost}_t$$

Cumulative performance was obtained by compounding daily returns.

Buy-and-Hold Benchmark To provide a baseline for comparison, a buy-and-hold benchmark was constructed using the same set of tradable currency pairs. At the beginning of the out-of-sample period, a fixed portfolio of currencies was selected based on the initial model signals. Specifically, the top-2 ranked currencies were assigned long positions, while the lowest 2-ranked currencies were assigned short positions.

The corresponding portfolio weights were determined using the same risk-based weighting scheme as the main strategy, ensuring comparability in terms of volatility exposure. However, unlike the dynamic strategy, these weights were held constant throughout the entire test period, with no subsequent rebalancing.

The return of the buy-and-hold portfolio is therefore given by:

$$R_t^{BH} = \sum_i w_i^{BH} \cdot r_{i,t+1}$$

where w_i^{BH} denotes the fixed initial weights.

This benchmark captures the performance of a static allocation based on initial information, and serves to isolate the value added by dynamic rebalancing, signal updating, and turnover control in the main strategy.

3.10 Performance Metrics

Strategy performance was evaluated using standard financial metrics:

- Annualised return:

$$\left(\prod_{t=1}^T (1 + R_t) \right)^{\frac{252}{T}} - 1$$

- Annualised volatility:

$$\sigma = \sqrt{252} \cdot \text{Std}(R_t)$$

- Sharpe ratio:

$$\frac{\mathbb{E}[R_t]}{\sigma}$$

- Maximum drawdown:

$$\min_t \left(\frac{C_t}{\max_{s \leq t} C_s} - 1 \right)$$

4 Critical Assessment

4.1 Model Performance

Before discussing the strategy results, it is useful to evaluate the predictive model on its own. The final Random Forest classifier was trained using a walk-forward validation procedure and then tested on a fully out-of-sample period. Figure 1 shows the feature importance ranking of the Random Forest model, while Table 1 summarises the out-of-sample classification performance.

On the surface, the model performance appears only modest. However, in the context of foreign exchange forecasting, this level of performance is not trivial. FX returns are noisy, heavily influenced by macro news, and difficult to predict even for highly structured models. What matters is whether it contains enough information to rank currencies in a way that can be monetised.

The class-wise metrics also suggest that the model is only mildly informative rather than decisively predictive. Precision and recall are relatively balanced across the two classes, which is useful in the sense that the model is not collapsing into one-sided predictions. At the same time, the improvement over a coin-flip is still narrow. This means that the raw classification model, taken alone, is not strong enough to justify aggressive directional trading.

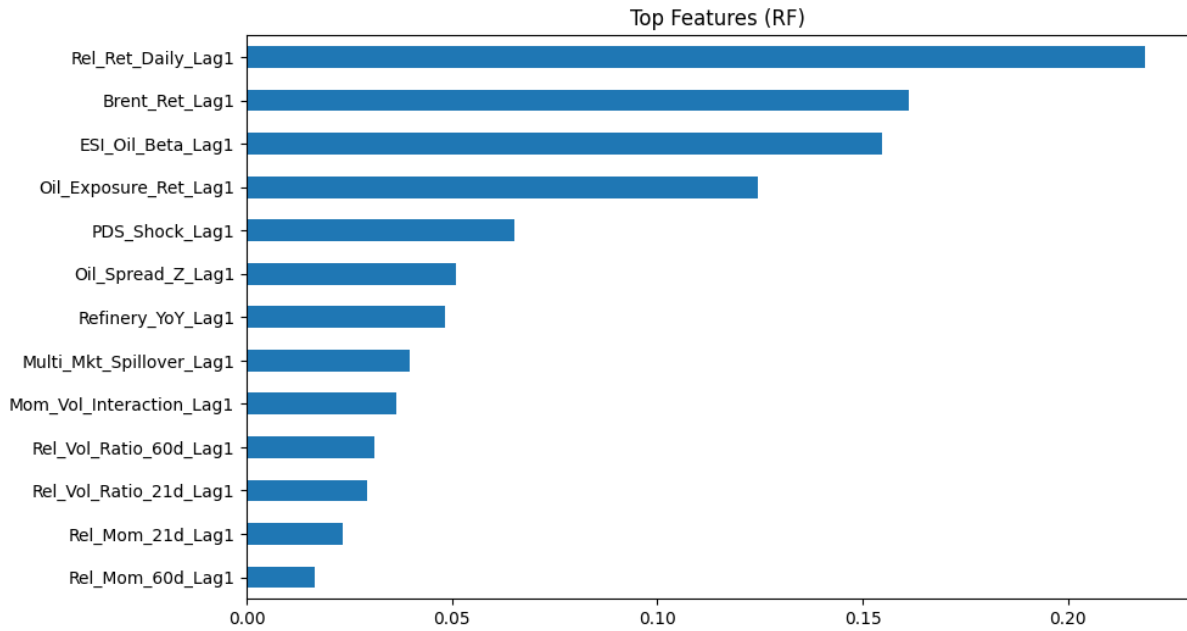


Figure 1: Random Forest feature importance ranking.

Table 1: Out-of-Sample Random Forest Classification Performance

Metric	Value
Accuracy	54.81%
Precision (Class 0)	0.55
Recall (Class 0)	0.60
F1-score (Class 0)	0.57
Precision (Class 1)	0.54
Recall (Class 1)	0.50
F1-score (Class 1)	0.52
Macro Average F1-score	0.55
Weighted Average F1-score	0.55

4.2 Interpretation of the Top Features

The feature importance chart provides a clearer picture of *what* the model is actually responding to. The top-ranked variables are economically intuitive, and that is one of the more encouraging findings of the project.

Relative Daily Return. The most important feature is `Rel_Ret_Daily_Lag1`. This suggests that short-horizon relative equity performance versus the U.S. remains the strongest predictor in the model. If local equities outperform the U.S. benchmark, it may reflect stronger relative growth expectations or stronger portfolio inflows, both of which can support the domestic currency. In other words, the model appears to be learning a short-run capital flow channel

Brent Return and ESI Oil Beta. The next most important features are `Brent_Ret_Lag1` and `ESI_Oil_Beta_Lag1`. Oil shocks affect net exporters and net importers asymmetrically, so it makes sense that a cross-sectional FX model picks up oil-linked variation. The fact that the trade-weighted oil sensitivity feature ranks highly is especially meaningful: it suggests that oil prices matter not just in isolation, but through a country's economic exposure to them. This gives support to the decision to engineer oil features beyond a simple headline price return.

Oil Exposure Return. `Oil_Exposure_Ret_Lag1` also appears among the leading features. While it is partly related to the previous two variables, its importance indicates that the model is using fixed directional commodity sensitivities in a useful way. This is consistent with the idea that currencies such as CAD and AUD behave differently from importers such as JPY or EUR when energy prices move.

PDS Shock. `PDS_Shock_Lag1` has non-negligible importance as well. This is encouraging because it was designed to capture stress-driven or volatility-amplified oil shocks rather than ordinary oil price changes. In practical terms, it suggests that the model is sensitive to both the direction of energy prices and the moves occurring under stressed market conditions.

Oil Spread Z-score and Refinery Proxy. `Oil_Spread_Z_Lag1` and `Refinery_YoY_Lag1` rank in the middle of the importance chart. These features are not dominant, but they still contribute. They are more indirect macro structure variables, so one would not expect them to drive day-to-day currency movements as strongly as relative equity returns or broad oil shocks. Still, their presence suggests that the model is extracting some signal from energy-market dislocations and downstream industrial conditions.

Multi-market spillover and interaction terms. `Multi_Mkt_Spillover_Lag1` and `Mom_Vol_Interaction_Lag1`'s lower ranking indicates that cross-market dependence matters, though less than the more direct relative-return and oil features. Spillover effects are likely to be diffuse rather than dominant in daily forecasting, and interaction terms often matter only in certain regimes.

Momentum and Relative Volatility. The lower ranking of momentum and relative volatility features does not necessarily imply they are unimportant. Rather, it may mean that at the daily forecasting horizon, these features are either weaker, more redundant with other variables, or more useful as conditioning variables within the trading layer.

Taken together, the feature importance profile is one of the stronger parts of the project. Even though the model's classification metrics are only modest, the variables it relies on are economically intuitive. This suggests that the framework is identifying plausible macro-financial channels.

4.3 From Model to Strategy

A key lesson from the results is that model quality and strategy quality are not the same thing. A model with modest predictive accuracy can still generate useful trading performance if it ranks currencies well and if the portfolio construction process extracts that information carefully.

This is particularly relevant here because the strategy is not implemented as a naive one-step directional trade. Instead, the Random Forest output is converted into a cross-sectional ranking, then filtered through smoothing, hysteresis, dispersion gating, inverse-volatility weighting, turnover control, and transaction costs. In that sense, it is trying to isolate and stabilise the small edge that the model provides.

4.4 Strategy Performance

The performance of the trading strategy is shown in Figure 2, with the key summary statistics presented in Table 2.

At a basic level, the strategy does improve on the benchmark. It delivers a positive annualised return of 0.78% versus a negative annualised return for buy-and-hold, and its Sharpe ratio is positive while the benchmark's is negative.

At the same time, the magnitude of the outperformance should not be overstated. A Sharpe ratio of 0.120 is still weak, and the maximum drawdown is meaningfully worse than the benchmark. This tells us that the strategy is not yet robust enough to be considered production-ready. It improves on the benchmark, but only marginally, and with a risk profile that remains difficult to justify.

The PnL curve itself gives a more nuanced picture. The strategy does not generate returns in a smooth or stable way. Instead, performance seems to come in distinct bursts. There are extended stretches where the strategy is flat or under pressure, followed by relatively sharp recovery periods. This pattern is consistent with a model that has a weak signal that becomes useful only in certain regimes. In other words, the edge is probably not constant through time.

A second observation is that the strategy performs better when cross-sectional separation across currencies is stronger. This fits the design of the framework, since the dispersion gate was introduced precisely to avoid trading when model scores are too similar. The fact that performance improves after these filters are added suggests that the raw model outputs contain useful information, but that much of the difficulty lies in deciding *when* those signals are worth trading.

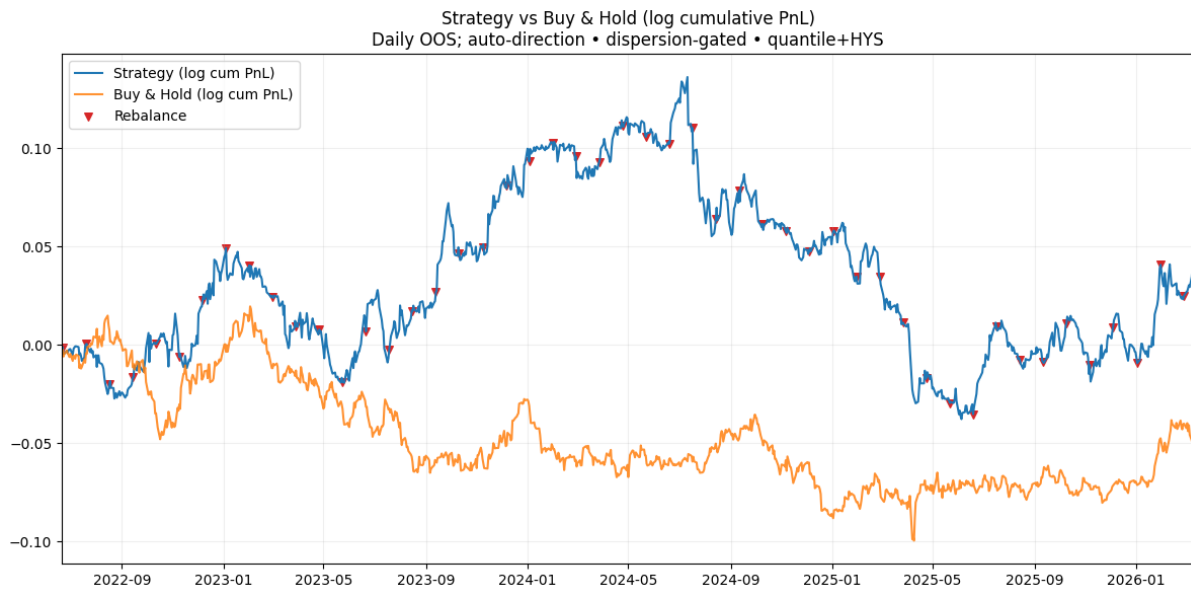


Figure 2: Cumulative log returns of the strategy versus buy-and-hold benchmark.

Table 2: Performance Metrics

Metric	Strategy	Buy & Hold
Annualised Return	0.78%	-1.04%
Annualised Volatility	6.45%	4.71%
Sharpe Ratio	0.120	-0.222
Maximum Drawdown	-16.0%	-11.2%

A third point is that the strategy appears to benefit from being interpreted as a ranking engine rather than as a pure directional forecaster. The underlying classification accuracy is modest, but the long-short basket still manages to outperform the benchmark. It suggests that the model is better at ordering currencies from stronger to weaker than at making binary directional calls on each one.

4.5 Strengths of the Framework

Even though the final strategy is not especially strong in absolute terms, the overall framework has several clear merits.

First, the feature set is not arbitrary. The top features identified by the Random Forest line up with plausible macro-financial channels, especially relative equity performance and oil-linked exposures. Thus, the project is not just a technical strategy, but an economically backed one as well.

Second, the portfolio construction framework is rigorous. Signal smoothing, hysteresis, turnover control, volatility targeting, and dispersion filtering are all techniques that address known practical problems in systematic trading. Even if the final Sharpe ratio remains low, these components make the strategy far more realistic and thorough.

Third, the cross-sectional setup is conceptually appropriate for FX. Currencies are relative instruments, and the project's ranking-based approach reflects that structure better than a directional framework would.

4.6 Limitations

The results also reveal important limitations.

The first limitation is that the predictive edge of the Random Forest is still weak. The model is informative enough to help with ranking, but not strong enough to support aggressive trading. This leaves the strategy heavily dependent on filtering and portfolio construction.

The second limitation is that the framework still assumes a single pooled model across currencies. In reality, the drivers of JPY and CHF are unlikely to be the same as those of AUD and CAD. A global model may therefore blur heterogeneous relationships that matter in FX.

The third limitation is that some important macro variables are still missing. Interest rate differentials, inflation surprises, yield curve information, and more explicit monetary policy variables are all central to currency valuation and were not fully incorporated in the current specification.

The fourth limitation is that the final strategy remains somewhat regime-dependent. The PnL curve suggests that performance is not evenly distributed over time. This raises the possibility that the strategy is only effective in certain macro environments and that its signal is unstable outside those periods.

Finally, although transaction costs are modelled, real implementation frictions may still be understated. In practice, slippage, spreads, and liquidity conditions would likely further reduce realised performance.

4.7 Potential Improvements

Several natural extensions follow from these findings.

A strong next step would be to move from a single pooled model to a clustered framework, where currencies are grouped into economically similar categories such as safe-havens, commodity currencies, and growth-sensitive currencies. Separate models could then be trained for each cluster.

A second improvement would be to strengthen the macro feature set, particularly by including rate spreads, inflation-linked variables, and broader global financial conditions. This would make the framework more consistent with how institutional FX desks typically think about currency valuation.

A third improvement would be to refine the signal extraction step further. Since the model appears to be more useful as a ranking engine than as a directional classifier, it may be worth explicitly optimising for cross-sectional ranking performance rather than classification accuracy.

Finally, the strategy could be made more regime-aware. If performance is concentrated in particular macro conditions, then adding a regime filter or switching framework may improve stability.

5 Conclusion

This study set out to investigate whether machine learning models, combined with economically backed features, can be used to predict foreign exchange movements and construct a profitable trading strategy. While the predictive performance of our model remains modest, the findings provide several meaningful insights into both the modelling process and the challenges of translating predictions into actionable strategies.

Overall, the project does not produce a particularly strong standalone predictive model, and the final trading strategy is not yet compelling enough in risk-adjusted terms. However, this outcome is not unexpected in the context of foreign exchange markets, where price movements are notoriously noisy and difficult to forecast. Importantly, the results demonstrate that even weak predictive signals can contain useful information when interpreted appropriately.

Our Random Forest model learns relationships that are economically intuitive, particularly in relation to relative equity performance, commodity exposure, and global risk conditions. These findings suggest that the feature engineering process is well-founded and aligns with established macro-financial theory. More importantly, when the model outputs are embedded within a disciplined cross-sectional trading framework, they are sufficient to generate performance that modestly exceeds a simple benchmark.

We believe this highlights a key insight from the project: in quantitative finance, the primary challenge is not only to predict returns, but to effectively extract and utilise weak predictive signals. The transformation of model outputs into portfolio positions—through ranking, filtering, and risk management—plays a central role in determining overall performance. In this regard, the framework developed in this study represents a coherent and credible approach to systematic FX trading.

At the same time, several avenues for improvement remain. Future work could focus on enriching the feature set with additional macroeconomic variables, such as interest

rate differentials and monetary policy indicators, as well as exploring clustering-based approaches to account for heterogeneity across currencies. Refinements to the trading strategy, including stronger signal filtering and regime-aware allocation, may also improve robustness.

In conclusion, while the strategy presented here should be viewed as an initial prototype rather than a finished product, the project successfully demonstrates a complete quantitative research pipeline—from data collection and feature engineering to model training and portfolio construction. The results underscore both the difficulty of the problem and the importance of combining predictive modelling with thoughtful strategy design. As such, this work provides a solid foundation for further exploration in systematic FX trading.

6 References

References

- [1] Ding, L. (2021). *Conditional correlation between exchange rates and stock prices*. *The Quarterly Review of Economics and Finance*, 80, 452–463.
- [2] Menkhoff, L., Sarno, L., Schmeling, M., & Schrimpf, A. (2012). *Currency momentum strategies*. *Journal of Financial Economics*, 106, 660–684.
- [3] Bianconi, M., & Cai, Z. (2017). *Higher moment exchange rate exposure of S&P500 firms*. *North American Journal of Economics and Finance*, 42, 513–530.
- [4] Lim, G. C. (2005). *Currency risk in excess equity returns: A multi time-varying beta approach*. *Journal of International Financial Markets, Institutions and Money*, 15, 189–207.
- [5] Beckmann, J., Czudaj, R., & Arora, V. (2017). *The relationship between oil prices and exchange rates: Theory and evidence*. U.S. Energy Information Administration Working Paper.
- [6] Plante, M., & Strickler, G. (2019). *Closer to one great pool? Evidence from structural breaks in oil price differentials*. Federal Reserve Bank of Dallas Working Paper No. 1922.
- [7] Zhang, R., Zhang, H., Gao, W., Li, T., & Yang, S. (2022). *The dynamic effects of oil price shocks on exchange rates*. *Sustainability*, 14(14), 8452.
- [8] Kaufmann, R. K., Dees, S., Gasteuil, A., & Mann, M. (2008). *Oil prices: The role of refinery utilization, futures markets and non-linearities*. *Energy Economics*, 30(6), 2609–2622.
- [9] Ready, R., Roussanov, N., & Ward, C. (2017). *Commodity trade and the carry trade: A tale of two countries*. *Journal of Finance*, 72(6), 2629–2684.
- [10] Michelis, L., Ning, C., & Ponrajah, J. (2025). *Safe haven currencies: A dependence-switching copula approach*. *Finance Research Letters*, 86.

- [11] Panazan, O., Gheorghe, C., Syed, A. A., & Jeribi, A. (2026). *Dynamic interactions between safe-haven assets and macroeconomic indicators*. Cogent Economics & Finance, 14(1).
- [12] Su, C.-W., Ding, Y.-M., Wang, K.-H., & Wang, X.-Q. (2026). *Testing safe-haven properties of green bonds, gold, and bitcoin*. Australian Economic Papers, 65(1), 16–29.

Disclaimer

This research material has been prepared by NUS Invest. NUS Invest specifically prohibits the redistribution of this material in whole or in part without the written permission of NUS Invest. The research officer(s) primarily responsible for the content of this research material, in whole or in part, certifies that their views are accurately expressed, and they will not receive direct or indirect compensation in exchange for expressing specific recommendations or views in this research material. Whilst we have taken all reasonable care to ensure that the information contained in this publication is not untrue or misleading at the time of publication, we cannot guarantee its accuracy or completeness, and you should not act on it without first independently verifying its contents. Any opinion or estimate contained in this report is subject to change without notice. We have not given any consideration to and we have not made any investigation of the investment objectives, financial situation or particular needs of the recipient or any class of persons, and accordingly, no warranty whatsoever is given and no liability whatsoever is accepted for any loss arising whether directly or indirectly as a result of the recipient or any class of persons acting on such information or opinion or estimate. You may wish to seek advice from a financial adviser regarding the suitability of the securities mentioned herein, taking into consideration your investment objectives, financial situation or particular needs, before making a commitment to invest in the securities. This report is published solely for information purposes, it does not constitute an advertisement and is not to be construed as a solicitation or an offer to buy or sell any securities or related financial instruments. No representation or warranty, either expressed or implied, is provided in relation to the accuracy, completeness or reliability of the information contained herein. The research material should not be regarded by recipients as a substitute for the exercise of their own judgement. Any opinions expressed in this research material are subject to change without notice.