# Short-term exchange rate forecasting with machine learning

**Edgar Cardoso**

Dissertation written under the supervision of Prof. Dan Tran.

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Católica Portuguesa, June 2025.

**Abstract**

**Title:** Short-term exchange rate forecasting with machine learning
**Author:** Edgar Cardoso

This master thesis aimed at forecasting exchange rates returns at a weekly and monthly frequency, focusing on the United States Dollar, Euro, Japanese Yen and British Pound. Exchange rate movements are notoriously hard to predict, which leads the literature to consider the random walk without drift as the de-facto benchmark which is consistently hard to outperform. The main goal of the prediction task was motivated by the hypothesis that sovereign risk priced by the market in various instruments would provide information into the future course of exchange rates. Machine learning models (Elastic Net, Random Forest, XGBoost) were built and trained using this data with the goal of producing predictions of the currency movements. Faithful to its role of benchmark, the random walk outperforms our models during the test phase. This test period included events like the COVID pandemic, the inflationary spikes that followed the reopening of the world economy as well as regional and global political tensions, a period which led to volatility spikes across asset classes and during which our models strongly underperformed. This unsuccessful attempt shines a light once again on the strenuous task of consistently predicting forecast exchange rates, especially at higher frequencies, and raised questions about the limitations of machine learning models when not purposely built with the limitations of the forecasted variable in mind.

**Keywords:** Machine Learning, Elastic Net, Random Forest, XGBoost, Exchange Rates, Interest Rates, Credit Default Swaps, Swap Spreads, Implied Volatility, Return Prediction

**Resumo**

**Título::** Previsão cambial de curto prazo com machine learning
**Autor:** Edgar Cardoso

Esta tese de mestrado teve como objetivo a previsão dos retornos das taxas de câmbio com uma frequência semanal e mensal, focando-se no Dólar dos Estados Unidos, Euro, Iene Japonês e Libra Esterlina. Os movimentos das taxas de câmbio são particularmente difíceis de prever, o que leva a literatura a considerar o random walk sem drift como o benchmark de facto que é consistentemente difícil de superar. O principal objetivo da tentativa de previsão baseou-se na hipótese de que a evolução do risco soberano, avaliado no mercado através de vários instrumentos, forneceria informações sobre a evolução futura das taxas de câmbio. Os modelos de aprendizagem automática (Elastic Net, Random Forest, XGBoost) foram construídos e treinados com base nestes dados, com o objetivo de produzir previsões das taxas de câmbio. Fiel ao seu papel de referência, o random walk superou os nossos modelos durante a fase de teste, que incluiu eventos como a pandemia de COVID, os picos inflacionistas que se seguiram à reabertura da economia mundial, bem como as tensões políticas regionais e globais, um período que conduziu a picos de volatilidade em todas as classes de activos e durante o qual os nossos modelos tiveram um desempenho fortemente inferior. Esta tentativa sem sucesso voltou a chamar a atenção para a árdua tarefa de prever taxas de câmbio de forma consistente, especialmente em frequências mais elevadas, e levantou questões sobre as limitações dos modelos de aprendizagem automática quando não são construídos propositadamente tendo em conta as limitações da variável prevista.

**Palavras-chave:** Aprendizagem automática, Elastic Net, Random Forest, XGBoost, Taxas de câmbio, Taxas de juro, Credit Default Swaps, Swap Spreads, Volatilidade implícita, Previsão do retorno

**Contents**

## 1 Introduction

In their 1983 study on currency markets, Meese and Rogoff [1983] concluded that accurately forecasting future foreign exchange spot rates posed significant challenges, arguing that the generally accepted principles guiding currency valuation were not always reliable in the short term. Reaching the same conclusion in Fama [1984], Eugene Fama introduced the ”forward premium puzzle,” describing the paradox of high yield currencies sometimes appreciating against low yielding currencies, contrary to established macroeconomic theories, as well proposing the existence of a time-varying risk premia. To understand the significance of these early contributions, we will firstly briefly explain the theory behind foreign exchanges rates, what is the carry trade and the source for its returns, as well as cover the ensuing literature that has put forth several possible causes to explain these excess returns, as well as predictors for this phenomenon, which has allowed willing investors to generate strong returns when currency deviate from what is considered their fair value. Offering superior risk-adjusted returns when compared to fixed income, equities and other publicly traded asset classes, it is easy to understand why it remains one of the most popular investment tactics for institutional investors, and why the literature on the topic is extensive. Having covered the fundamentals of foreign exchange and the developments exposed in the literature in explaining and forecasting returns, we will focus our attention on machine learning methods, their definition and advantages when compared to econometric models. This is important as econometric models remain the primary focus of the field, and although they offer a variety of advantages explaining their popularity, they also possess several limitations which reduce the predictive power found with the different variables. This will motivate our choice for this more complex approach. We will describe the data collected and considered for our study, before employing feature selection and engineering to extract more information from the data. We will run several model types and hyper-parameter combinations using training data, to reach a good compromise in model complexity and out-of-sample prediction capabilities, using the validation dataset to conduct preliminary tests. In short, this thesis will be composed of the following sections: literature review, description and analysis of data, feature selection and engineering, description and estimation of models, the methodology for analysis of results and a concluding section. The overarching research question can be summarized as follows:

*Can changes in spot exchange rates be partially explained by evolving perceptions of sovereign risk by financial institutions and capital allocators?*

## 2 Literature

### 2.1 Currencies, deviations from equilibrium and the carry trade

One fact about the carry trade that has provoked such extensive research and allocation of capital over the decades is the fact that it consistently provides superior risk-adjusted returns when compared to other publicly available asset classes Burnside et al. [2008]. Barro and Ursúa [2012] propose that the excess returns are compensation for an expected although unprecedented macroeconomic event, dubbed a "Peso problem" and Burnside et al. [2011] claim that peso events features high values of the stochastic discount factor rather than very large negative payoffs, giving a possible explanation for these excess returns. The argument of a specific risk could help to explain why its risk premium is uncorrelated with other known risk factors. Bacchetta and Van Wincoop [2007] argue that the violation of uncovered interest parity (UIP) can be explained by the fact that less than 1 percent of global FX positions are actively managed, since the benefit-to-cost ratio of actively managing those positions for those agents (as opposed to active money managers) is marginal after accounting for fees, expertise and transaction costs, unless the deviations and thus gains become substantial enough. Several empirical studies have found that carry trades tend to have excess kurtosis and negative skewnessFarhi and Gabaix [2008] and the findings of Harvey and Siddique [2000] and Blau [2017], suggest the existence of a risk premium for assets exhibiting systematic negative skewness, which goes against investors natural preferences for positive skewness.Farhi and Gabaix [2008] also argue that the forward premium puzzle is a risk premium for the possibility of rare economic disasters, which would impact some economies more than others, leading to varying risk premiums from country to country for currencies and financial assets. This seems consistent with the empirical observations of fixed income markets during times of great market uncertainty, which typically leads to a flight to quality, propping up the prices the assets of a few select countries, as seen in Ranaldo and Söderlind [2010]. Finally, we can also consider that price inefficiency could be reinforced by behavioural biases Froot and Thaler [1990], such as expectational errors, anchoring biases, as well as initial under-reactions and later overreactions leading to bubbles in different asset classes. Given all these possible explanations which offer some explanatory variables for past events, the reality is that the source of the returns, although more modest now due to large amounts of speculative capital Barroso and Santa-Clara [2015], is still partly a mystery.

### 2.2 Rise and fall of carry

The high risk-adjusted returns of currency carry trades lead to institutional players seeking to gain exposure to this strategy, resulting in strong market concentration among several large players. As seen in Jylhä and Suominen [2011], a portion of hedge fund returns (16% of returns on average for 2011) consistently came from carry trade activity. This is further exacerbated by the high leverage required to reasonably profit from the trade, resulting in few like-minded institutional investors building and unwinding simultaneously large positions causing drastic price movements. Coincidentally, Abreu and Brunnermeier [2003] note that investors might be aware of an asset being in a bubble, but will remain invested to generate additional profits from it, while attempting to time the exit before a possible crash, in a prisoner dilemma dynamic. This can lead to prolonged asset bubbles and stronger corrections and crashes. This could help to understand the crashes as they're happening, but we still require information as to possible triggers and predictors. Melvin and Taylor [2009] found that a sudden surge in financial uncertainty could lead to a lack of liquidity in several markets and cause a flight to safety, impacting exchange rates and currency markets. Brunnermeier et al. [2008] found that crashes can be triggered by the simultaneous liquidation and deleveraging by investors due to constraints in funding liquidity (measured by the TED spread). This unwinding of positions can also be motivated by an increase in expected market volatility (VIX) which may push investors over their risk allocation thresholds, reduce their available capital to fund the trade or increase its cost, reducing the profitability of the trade. The impact of funding liquidity is also discussed in Karnaukh et al. [2015]. The rapid sell-off in the market of higher yield currencies and the repurchase of the lower-yield investing currencies will result in funding currencies being momentarily propped up and investment currencies crashing, as funds all rush through the door to reduce their exposure. During times of market volatility and periods of flight to safety, this will mean selling riskier currencies at discounted prices and overpaying for safe-haven currencies highly desired by all market participants. Campbell et al. [2002] found that correlation between assets tend to increase significantly in strong corrections and bear markets, with the exception of assets perceived as safe, and Adrian et al. [2015] highlighted the role of repo markets and US commercial paper transactions in predicting the USD rate fluctuations against other currencies, as risk appetites evolve and investors increase their borrowing or deleverage using US-denominated debt as a funding currency during good times and as a safe-haven during uncertain periods. The predictive value of market volatility and the underperformance of carry trades during these market conditions is also discussed in Menkhoff et al. [2012a]. Brunnermeier et al. [2008] also found strong co-movement of different currencies possessing similar yields, which will also increase as the gap in yields decreases, hypothesised as an attempt at diversification of carry trades across different currencies by institutional players. This strong co-movement indicates a limited ability to diversify risk, as the market conditions impacting the carry trade of a pair will typically impact several others, leading to the unwinding of large portfolios and revealing non-diversifiable systematic risk components to the trade. Other factors mentioned as limiting the implementation of the strategy were low trading volumes and wide bid-ask spreads (trading liquidity), increasing the cost and risk profile of the trade. There is strong evidence to argue for the strategies excess returns as a compensation for systematic risk and as additional compensation for the negative skewness of the trade Blau [2017], and **?** sought to answer this postulate by constructing carry trade portfolios of G10 currencies with out-of-the-money puts to protect from crashes. Although the returns were modest due to the continuous cost of insurance, it still provided positive returns, raising doubts on the theory that excess returns were entirely a compensation for systematic risk.

### 2.3 Factor models and currency strategies

Lustig et al. [2011] proposed the $HML_{fx}$ factor, where one buys a portfolio of high yielding currencies and sells low-yielding currencies that generated satisfactory excess returns, but noting that the high yield currencies possessed more volatility and were typically associated with less stable economies, arguing for the presence of a systematic risk. Much like the asset pricing model seen in Fama and French [1992] and subsequent publications, this $HML_{fx}$ factor helps to better understand past returns, although it provides limited information in forecasting future performance. The $HML_{fx}$ factor has remained the basis for most currency strategies, but other factors have shown to improve risk-adjusted returns by offering diversification benefits and returns in different market conditions.Asness et al. [2013] found evidence of value and momentum strategies across different asset classes, whose negative correlation across time improved the performance of portfolios exposed to both factors simultaneously. Menkhoff et al. [2013] further explored currency momentum, showing weak correlation to carry and strong performance in currencies not explained by traditional risk factors, while also noting that momentum crashes were equally as important in currencies as in other asset classes. Berge et al. [2011] explored directional trading strategies based on carry, momentum, and value by employing receiver operating characteristic curve, a binary classifier model and Berge et al. [2011] developed upon this methodology while also including forward-curve data given its predictive nature for carry trads, as described in Ang and Chen [2010]. Barroso and Santa-Clara [2015] improved upon earlier findings by combining 3-month momentum and the magnitude of forward discounts with parametric portfolio policies, taking advantage of carry and momentum while dynamically weighing the different currencies based on their perceived potential to over-perform, surpassing the performance of previous published carry strategies published in the literature.Zhang [2022] finds that currency momentum is driven by highly correlated global factors such as the dollar factor and the momentum factor, concluding that strong past returns in these factors along with an expected low volatility for their future returns explains and helps to outperforms individual currency momentum. Adrian et al. [2015] measured the impact of changes in dollar denominated banking sector liabilities on the appreciation of the dollar against foreign currencies and found it depended on balance sheet capacity, funding liquidity constraints and an evolving risk appetite of financial institutions.

### 2.4 Business cycle exposure and macroeconomic risk

Menkhoff et al. [2013] employed a cross-sectional approach to currency valuation and found that macroeconomic variables such as interest rate differentials, real GDP and money growth and real exchange rates provided signals as to the performance of different currencies. This was significant because the literature had struggled to find relevant information using these variables when analysing time series of individual currency pairs. Ready et al. [2013] found that countries dependent on the exports of commodities and basic goods tend to have higher interest rates than their counterparts, which isn't fully accounted for in FX rates and leads to potential carry trade opportunities. They established explanatory power for FX forward discounts by analysing the composition of the balance of trade of different countries, justifying the higher and more volatile interest rates as risk premiums for a greater exposure to the global business cycle. Furthermore, they found that pro-cyclical variables such as freight shipping and commodity prices helped to predict higher returns in carry trade, consistent with Bakshi and Panayotov [2013]. As we've mentioned previously,Ang and Chen [2010] found that the evolution of interest rates and term spreads possessed predictive power for currency returns, which were weakly correlated with traditional carry, showing once more the impact of macroeconomic conditions in FX rates. Jordà and Taylor [2012] found that the equilibrium exchange rate defined by PPP helped in forecasting future FX rates, employing ECM and TECM models which allowed for short-term deviations from the fundamental exchange rate while keeping track of long-term trends, particularly useful in periods of reversal to equilibrium exchange rates. Della Corte et al. [2012] and Corte et al. [2016] argue that the impact of consistent trade imbalances have significant impact on the evolution of exchange rates, increasing in times of economic uncertainty where currencies perceived as riskier command higher risk premiums. Barroso et al. [2018] builds upon these findings by considering the dual impact of external trade and interest rate impact currencies' risk premiums, finding that interest rates differentials, when controlled for external imbalances, offer great explanatory power into carry returns. Proposing an orthogonal carry trade which improves the Sharpe ratio of traditional carry by 38%, due to the lower correlation with global market movements and lower portfolio concentration into individual currencies. Della Corte et al. [2012] test carry trades baed on the priced volatility of currencies, by building a long-short portfolio of Forward Volatility Agreements (FVA) based on the slope of implied volatility curves, buying downward slopping implied volatility and shorting upward slopping volatility, giving new insights into the risk premia of volatility in currency markets where returns don't appear to be connected to other factors such as dollar, carry or global imbalances. Zeng [2019](Ming, 2023) finds strong explanatory power of interest rate volatility (IRV) in predicting carry and momentum returns across 48 currencies and market conditions, theorizing that constraints felt by financial institutions during periods of increasing IRV lead to an unwinding of positions generating negative returns for both of these currency factors.Calice and Zeng [2021] provide evidence of the explanatory power of credit default swap levels and term premia into currency returns, providing once more proof about the relation between perceived sovereign/ systematic risk of an economic and the additional risk premia demanded by investors to invest on assets of those economies.

## 3 Data

Our choice of data was motivated by the hypothesis that an array of financial assets carry information regarding different risks of countries and economies and that these risk pricings could possess complementary information helping to assess the risk component of currencies. The variables chosen all possess high frequency data which reveals how market participants take into account new information as it is released to establish new fair values for the assets. Directly or indirectly, we believe these variables possess information regarding monetary and fiscal policy, price levels and inflation expectations, government credibility and commitment to its debt-holders, market liquidity and stability measures, uncertainty expectations through volatility measures, trends in markets and economies, business cycles phases and many more which we will try to present shortly.

### 3.1 Exchange Rates

Spot foreign exchange rates were collected using Datastream as the basis for our analysis, focusing on the Euro(EUR/USD), the British Pound(GBP/USD) and the Japanese Yen(USD/JPY) for reasons of data availability. The dollar index [ICE], a trade-weighted basket of currency pairs against the dollar, was also considered for our analysis. The dollar index has been offered by the Intercontinental Exchange since 1973 and represents a good proxy for the strength of the dollar against other currencies. As this ICE index is also one of the most popular futures contract [ICE], it serves as a good proxy for the views of market participants regarding exchange rates, international trade and the strength of the dollar, on a daily basis. Additionally, forward rates with one month of maturity were also retrieved to compute carry as a variable for the models.

### 3.2 Sovereign yield curves

Sovereign bond yields represent the compensation required by investors to lend money to governments at various maturities, compensating investors for the opportunity cost of holding the securities when considering alternatives, including compensation for time value of money, inflation and inflation expectations, perceived riskiness of government accounts, growth trends and asset valuations, fiscal and monetary policies and their impact. Despite the ever more frequent interventions by central banks and governments on sovereign debt markets (QE/ QT), we hope the levels across maturities will provide some insight into the macroeconomic expectations of market participants that may influence currency movements over time. The choice was also made to include the central bank's monetary policy rate, which greatly impacts the borrowing costs of institutions, firms and market participants on the short-term.

### 3.3 Credit Default Swaps

Credit default swaps are instruments that serves as insurance on a given debt instrument, loan, portfolio of loans or index of debt securities, protecting its buyer from losses during credit events, agains the payment of a premium or spread. We note that the CDS described below have sovereign bonds as their underlying, the premiums reflecting market participants' opinions about the riskiness of government debt securities and their evolution through time, by comparing premiums across maturities. We must mention that for the case of the European Union, because of the historically low issuance of EU bonds by its institutions, the CDS premium of German government bonds were used instead as a proxy, as Germany is generally considered the benchmark for European and euro related financial markets transactions. Once again, we mention Calice and Zeng [2021] as a good reference for CDS premiums used as a determinant for exchange rate relative valuation.

### 3.4 Swap spread curve

Swap spreads are computed as the difference between the fixed leg of an Interest Rate Swap and the yield on a government bond with the same maturity. Swap spreads are a widely used indicator for tensions in the financial system, since during these periods swap rates tend to increase to compensate for higher perceived credit and liquidity risks, while government yields typically decrease during periods of flight to safety. Swap spreads are also indicators of supply and demand for government securities, with high issuances of debt pushing down swap spreads, sometimes into negative territory.

### 3.5 Implied volatility data for currencies and interest rates

Periods of high volatility are bad for carry trades, as they increase the VaR of portfolios, forcing portfolio managers to trim their highly leveraged positions to meet internal metrics and also causing higher collateral requirements by counterparties. In the worst cases, high volatility can lead to large losses, a feedback loop of stop losses and lack of liquidity in the markets. As discussed in Szakmary et al. [2003], implied volatility tends to be a strong predictor of future realized volatility in a variety of public markets, and for this reason the choice was made to include in our analysis implied volatility data related to foreign exchange and interest rates computed from call and put options on the relevant underlyings.

### 3.6 Financial Stability and business cycle

The choice was made to include weekly financial commercial paper outstanding, as Adrian et al. [2015] of the New York Federal Reserve showed that the growth of U.S.-dollar denominated banking sector liabilities revealed levels of risk appetite by financial institutions, possessing forecasting power regarding the appreciation of the US dollar, the second leg of of our exchange rate pairs. The choice to include the TED spread and VIX in our analysis as measures of uncertainty and market liquidity is based on extensive a priori research on these variables, as seen in Brunnermeier et al. [2008] and many other publications. Finally, gold and two commodity indices were considered as a proxy for global price levels, uncertainty (in the case of gold) and global demand (commodity indices). Commodities were shown to impact currency valuations Ready et al. [2013], as were freight prices Bakshi and Panayotov [2013], although the countries considered are service-based economics and not very dependent on commodity prices.

| **Variable** | **Definition** | **Source** | **Frequency** |
|---|---|---|---|
| DXY | Intercontinental Exchange US Dollar Index. | Bloomberg Terminal | Daily |
| Fx spot | Mid foreign exchange spot rates of EUR/USD, GBP/USD, USD/JPY. | LSEG Datastream | Daily |
| Forward Rates | Mid foreign exchange 1-month forward rates of EUR/USD, GBP/USD, USD/JPY. | LSEG Datastream | Daily |
| Central Bank | The short-term rate employed by central banks to steer monetary policy. | Bank of International Settlements | Daily |
| Govt 3M | Yield on 3-month government bond. | Bloomberg Terminal | Daily |
| Govt 1Y | Yield on 1-year government bond. | Bloomberg Terminal | Daily |
| Govt 2Y | Yield on 2-year government bond. | Bloomberg Terminal | Daily |
| Govt 5Y | Yield on 5-year government bond. | Bloomberg Terminal | Daily |
| Govt 10Y | Yield on 10-year government bond. | Bloomberg Terminal | Daily |
| Govt 20Y | Yield on 20-year government bond. | Bloomberg Terminal | Daily |
| 1Y cds | Premium on 1-year Credit Default Swap on Govt Bond. | LSEG Datastream | Daily |
| 5Y cds | Premium on 5-year Credit Default Swap on Govt Bond. | LSEG Datastream | Daily |
| 10Y cds | Premium on 10-year Credit Default Swap on Govt Bond. | LSEG Datastream | Daily |
| 1Y Swap Spread | 1-year swap rate minus the 1-year Govt bond yield. | Bloomberg Terminal | Daily |

*(continued on next page)*

(continued from previous page)

| Variable | Definition | Source | Frequency |
|---|---|---|---|
| 2Y Swap Spread | 2-year swap rate minus the 2-year Govt bond yield. | Bloomberg Terminal | Daily |
| 5Y Swap Spread | 5-year swap rate minus the 5-year Govt bond yield. | Bloomberg Terminal | Daily |
| 10Y Swap Spread | 10-year swap rate minus the 10-year Govt bond yield. | Bloomberg Terminal | Daily |
| Option Implied Vol 1M | Implied Volatility of 30-day options on FX rates for each FX pair mentioned earlier. | Bloomberg Terminal | Daily |
| Option Implied Vol 3M | Implied Volatility of 90-day options on FX rates for each FX pair mentioned earlier. | Bloomberg Terminal | Daily |
| DXY IV 1M | Implied Volatility of 30-day options on ICE US Dollar Index. | Bloomberg Terminal | Daily |
| DXY IV 3M | Implied Volatility of 90-day options on ICE US Dollar Index. | Bloomberg Terminal | Daily |
| TYX Index 10-day IV | Implied Volatility of 10-day options on 10-year US Treasury Futures. | Bloomberg Terminal | Daily |
| TYX Index 30-day IV | Implied Volatility of 30-day options on 10-year US Treasury Futures. | Bloomberg Terminal | Daily |
| TYX Index 60-day IV | Implied Volatility of 60-day options on 10-year US Treasury Futures. | Bloomberg Terminal | Daily |
| VIX | CBOE VIX index measuring implied volatility of options on the S&P500. | FRED database | Daily |
| TED spread | The TED spread is the difference between the 3-month Euro-Dollars (US LIBOR) and a 3-month Treasury bill. | FRED database | Daily |
| Fin CP | Log of domestic financial commercial paper outstanding (in billions of USD). | FRED database | Weekly |
| Gold | Gold spot price based on market transactions. | FRED database | Daily |

(continued on next page)

*(continued from previous page)*

| Variable | Definition | Source | Frequency |
|---|---|---|---|
| CRB Com index | CRB Commodity Index (raw industrials). | Bloomberg Terminal | Daily |
| BBG Com index | Bloomberg Commodity Index. | Bloomberg Terminal | Daily |
| BDI | Baltic Dry Index for shipping freight-cost. | Bloomberg Terminal | Daily |

Table 1: List of financial and macroeconomic variables used in the project containing information regarding price levels, interest rates, financial stability, asset price volatility and sovereign risk measures

## 4 Exploratory Data Analysis

This section is dedicated to deepening our knowledge of the data, its evolution over time and the relationships between variables, which will help us conceptualize a model able to deal with the specificities of the dataset. This is especially important because financial time series have several common issues. Zhang and Hua [2025] lists non-stationarity, high multicollinearity, low signal-to-noise ratios, imbalanced datasets and asynchronous data as issues impacting high -frequency financial data, for example, which affect linear models such as OLS regressions severely.

### 4.1 Summary statistics and overview

Summary statistics for all the variables considered initially can be seen below, with additional tables in the appendix. It is worth noting that the credit default swaps starts much later, as regulators forced this OTC derivatives to become more standardized and transparent. We also see that the data for swap spreads ends earlier than other time series, mainly due to the transition from LIBOR rates in most advanced economies. Both of these restrictions will limit the available period where we can train and test our model, as we seek to include all these variables in our forecasting model.

Not surprisingly, the summary statistics (skewness and kurtosis) quickly reveal that most time series concerned are non-normal, non-symmetric and possess fat tails, something often observed in financial time series. This has implications for the validity of results from models which depend on these assumptions.

### 4.2 Correlation of variables

Overall, we notice that within given term structures ( yields, cds, swap spread, volatility) variables have high correlations as expected, while correlations outside of the term structure

Table 2: Summary statistics for EUR related variables

| Variable | Count | Mean | Std.Dev. | Min. | 25% | 50% | 75% | Max. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Fx spot | 5718.00 | 1.226 | 0.128 | 0.959 | 1.117 | 1.209 | 1.322 | 1.599 | 0.497 | -0.481 |
| Govt 3M | 5718.00 | 0.895 | 1.596 | -1.021 | -0.558 | 0.167 | 2.053 | 4.310 | 0.636 | -1.061 |
| Govt 1Y | 5718.00 | 0.996 | 1.622 | -0.950 | -0.593 | 0.394 | 2.269 | 4.593 | 0.553 | -1.106 |
| Govt 2Y | 5718.00 | 1.065 | 1.606 | -1.016 | -0.555 | 0.562 | 2.503 | 4.666 | 0.411 | -1.232 |
| Govt 5Y | 5718.00 | 1.420 | 1.616 | -1.001 | -0.182 | 1.333 | 2.791 | 4.682 | 0.160 | -1.420 |
| Govt 10Y | 5718.00 | 1.978 | 1.611 | -0.853 | 0.422 | 1.988 | 3.496 | 4.687 | -0.030 | -1.396 |
| Govt 20Y | 5718.00 | 2.441 | 1.594 | -0.626 | 0.976 | 2.523 | 3.943 | 4.895 | -0.148 | -1.297 |
| 1Y swap spread | 5718.00 | 41.058 | 27.630 | 0.776 | 22.419 | 34.538 | 53.956 | 210.589 | 1.471 | 3.186 |
| 2Y swap spread | 5718.00 | 43.490 | 23.220 | 8.753 | 24.689 | 38.597 | 56.998 | 126.081 | 0.892 | 0.316 |
| 5Y swap spread | 5718.00 | 41.678 | 19.812 | 3.116 | 28.944 | 38.803 | 52.995 | 108.637 | 0.598 | 0.054 |
| 10Y swap spread | 5718.00 | 35.111 | 17.310 | -5.901 | 23.310 | 33.689 | 45.330 | 101.468 | 0.616 | 0.406 |
| 1Y cds | 4421.00 | 5.637 | 8.229 | 0.280 | 1.629 | 2.440 | 4.950 | 50.489 | 2.935 | 8.862 |
| 5Y cds | 4421.00 | 22.875 | 21.537 | 5.000 | 8.480 | 14.960 | 27.727 | 118.380 | 2.009 | 3.760 |
| 10Y cds | 4421.00 | 23.326 | 16.041 | 6.190 | 10.770 | 18.130 | 29.440 | 91.980 | 1.459 | 2.101 |
| Option Implied Vol 1M | 5718.00 | 8.896 | 3.112 | 3.772 | 6.686 | 8.448 | 10.450 | 28.885 | 1.603 | 4.766 |
| Option Implied Vol 3M | 5718.00 | 9.021 | 2.954 | 4.138 | 6.780 | 8.625 | 10.650 | 24.652 | 1.328 | 3.076 |
| Central Bank Rate | 5718.00 | 1.379 | 1.472 | 0.000 | 0.000 | 1.000 | 2.250 | 4.500 | 0.778 | -0.718 |

Table 3: Summary statistics for USD related variables

| Variable | Count | Mean | StdDev | Min | 25% | 50% | 75% | Max | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Govt 3M_USD | 5718.00 | 1.568 | 1.833 | -0.137 | 0.079 | 0.885 | 2.408 | 5.553 | 0.979 | -0.503 |
| Govt 1Y_USD | 5718.00 | 1.781 | 1.764 | 0.055 | 0.247 | 1.151 | 2.846 | 5.517 | 0.821 | -0.783 |
| Govt 2Y_USD | 5718.00 | 1.909 | 1.597 | 0.109 | 0.541 | 1.366 | 3.007 | 5.275 | 0.688 | -0.939 |
| Govt 5Y_USD | 5718.00 | 2.448 | 1.295 | 0.203 | 1.448 | 2.211 | 3.537 | 5.228 | 0.307 | -1.032 |
| Govt 10Y_USD | 5718.00 | 3.065 | 1.153 | 0.504 | 2.114 | 2.970 | 4.121 | 5.358 | -0.084 | -1.035 |
| Govt 20Y_USD | 5718.00 | 3.572 | 1.130 | 0.962 | 2.677 | 3.541 | 4.592 | 5.660 | -0.163 | -1.110 |
| 1Y cds_USD | 4410.00 | 16.297 | 17.289 | 1.000 | 7.530 | 12.020 | 18.034 | 169.110 | 4.368 | 27.323 |
| 5Y cds _USD | 4214.00 | 21.426 | 12.935 | 5.460 | 12.342 | 18.155 | 29.390 | 90.000 | 1.567 | 3.985 |
| 10Y cds_USD | 4410.00 | 36.572 | 15.499 | 10.250 | 25.770 | 31.840 | 47.385 | 97.000 | 0.654 | 0.004 |
| DXY IV 1M _USD | 5718.00 | 9.058 | 2.897 | 4.321 | 7.123 | 8.638 | 10.400 | 29.713 | 1.939 | 6.929 |
| DXY IV 3M _USD | 5718.00 | 9.158 | 2.701 | 4.628 | 7.262 | 8.848 | 10.411 | 23.807 | 1.576 | 4.384 |
| DXY spot_USD | 5718.00 | 89.847 | 9.208 | 71.329 | 81.490 | 90.076 | 97.074 | 114.106 | 0.067 | -0.997 |
| 1Y Swap Spread_USD | 5718.00 | 26.950 | 20.787 | -4.190 | 14.112 | 20.590 | 34.580 | 169.380 | 2.348 | 7.710 |
| 2Y Swap Spread_USD | 5718.00 | 26.871 | 22.234 | -9.850 | 12.900 | 21.520 | 34.570 | 168.620 | 2.253 | 7.047 |
| 5Y Swap Spread_USD | 5718.00 | 20.575 | 25.342 | -18.080 | 2.370 | 10.685 | 36.828 | 129.270 | 1.249 | 1.469 |
| 10Y Swap Spread_USD | 5718.00 | 8.775 | 21.281 | -52.880 | -5.265 | 3.485 | 24.502 | 68.360 | 0.654 | -0.343 |
| Central Bank Rate_USD | 5718.00 | 1.635 | 1.855 | 0.040 | 0.130 | 0.940 | 2.410 | 5.410 | 0.969 | -0.552 |
| TYX Index 10-day IV_USD | 5718.00 | 24.401 | 16.761 | 2.770 | 15.380 | 20.630 | 29.718 | 314.160 | 7.107 | 91.720 |
| TYX Index 30-day IV_USD | 5718.00 | 25.029 | 15.346 | 7.730 | 16.380 | 21.570 | 30.368 | 187.120 | 5.421 | 47.544 |
| TYX Index 60-day IV_USD | 5718.00 | 25.387 | 14.763 | 8.850 | 16.610 | 21.785 | 30.428 | 139.210 | 4.188 | 26.241 |

with other variables are generally moderate, ranging mainly between -0.6 to 0.6. We will engineer the highly correlated features later on to avoid issues of multicollinearity while

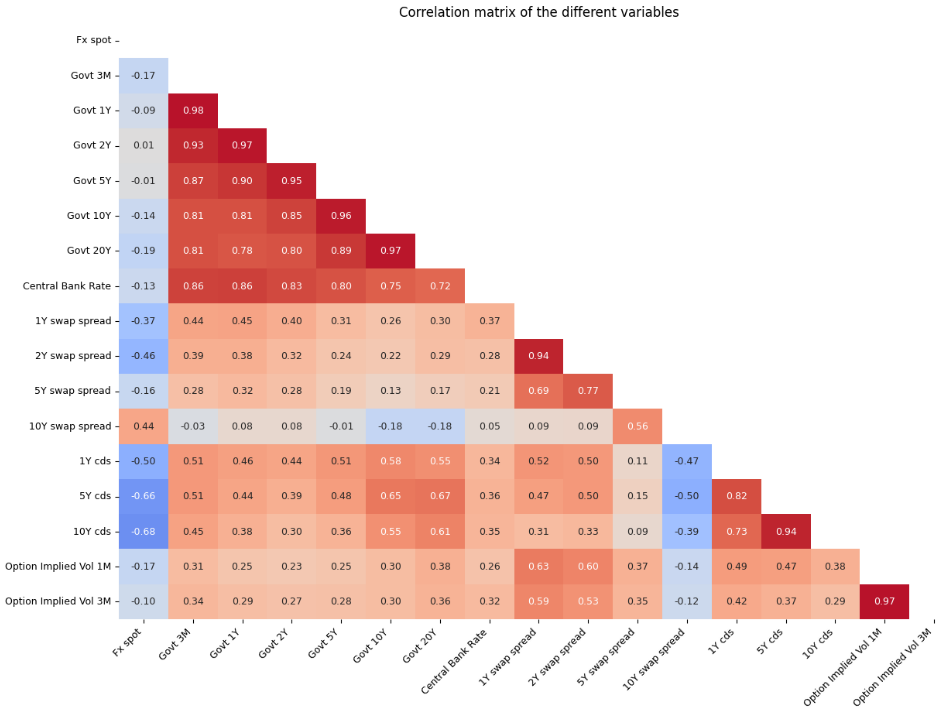



Figure 1: Correlation Matrix for JPY related variables

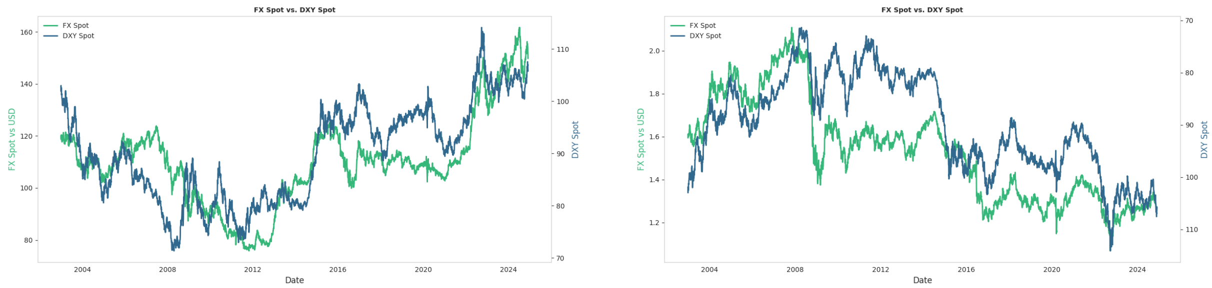
attempting to retain predictive power during our model training phase. Swap spreads appear to be weakly correlated with government rates for the Euro and the GBP, while the JPY and the USD show correlations of 0.3 to 0.4. Although swap spreads and cds premium should both express measures of sovereign risk, we note a correlation of only around 0.5 between these variables for all datasets, likely due to additional factors that affect swap spreads such as debt issuance, market liquidity and broker dealers balance sheet constraints Boyarchenko et al. [2018].In the case of the USD/JPY, we note that Japanese interest rates have much lower impact on the exchange rate compared with the other currencies, especially at the shorter end of the curve. This is interesting given the way monetary policy has been conducted in Japan when compared with other developed economies. We find an interesting correlation in several currencies between the Exchange rate implied volatility and the short end of the CDS and Swap spread curves, possibly indicating periods of structural macro uncertainty (FX, interest rates and credit risk).

### 4.3 Exchange Rates

Given that a currency pair is composed of two currencies, it is always hard to understand which currency/ economy is mainly driving the move of an FX pair. The choice was made to compare the evolution of all pairs against the dollar index, since the dollar is the second leg of each pair considered and because these same pairs are all considered in the dollar index with different weights. This helps to understand when movements are being led by the strength of the

USD (parallel move of inverted Dxy and Fx rate) or when it is being led by the target currency (widening/ tightening of spread in graph).

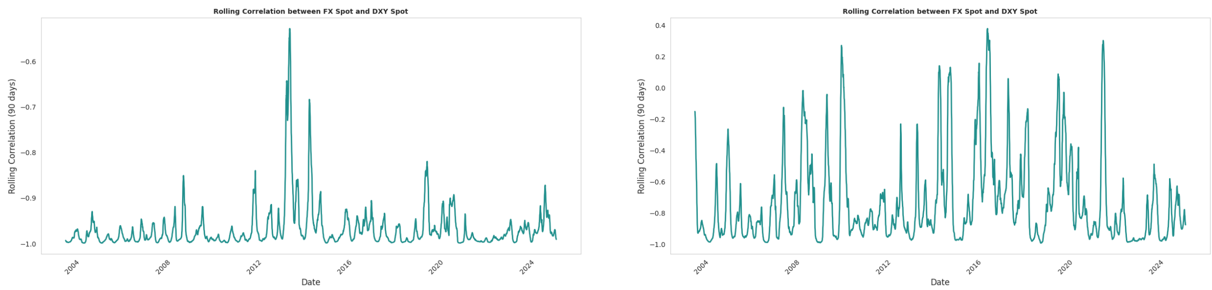
$$\text{Dollar Index (DXY)}=100\times\left(\frac{E_{\text{USD/EUR}}^{0.576}\times E_{\text{USD/JPY}}^{0.136}\times E_{\text{USD/GBP}}^{0.119}\times E_{\text{USD/CAD}}^{0.091}\times E_{\text{USD/SEK}}^{0.042}\times E_{\text{USD/CHF}}^{0.036}}{E_{\text{USD/EUR0}}^{0.576}\times E_{\text{USD/JPY0}}^{0.136}\times E_{\text{USD/GBP0}}^{0.119}\times E_{\text{USD/CAD0}}^{0.091}\times E_{\text{USD/SEK0}}^{0.042}\times E_{\text{USD/CHF0}}^{0.036}}\right)$$




(a) Impact of dollar strength on USD/JPY pair

(b) Impact of dollar strength on GBP/USD pair

Figure 2: Impact of dollar strength on currencies




(a) Rolling correlation of Dollar Index with EUR/USD pair

(b) Rolling correlation of DXY with GBP/USD pair

Figure 3: Rolling correlations of Dollar Index against FX pairs

The broad dollar movement tends to have a very strong impact on the pairs movements, but that influence will vary over time. It is therefore possible to say that the pair is mostly USD dominated, at least for these periods examined, which justifies our choice to consider variables from both currencies for our forecasting models. While EUR's correlation remains fairly stable throughout most of the period, we reach different conclusions for the JPY and the GBP, whose correlations vary widely, sometimes even changing signs from positive to negative. This could mean a more important weight of USD features for the prediction of Euro rates, and also be more country specific features for GBP and JPY, as well as added volatility in rates. For all currencies, we see big spikes during periods of financial stress such as the global financial crisis, the eurozone crisis and the COVID pandemic, to name a few.

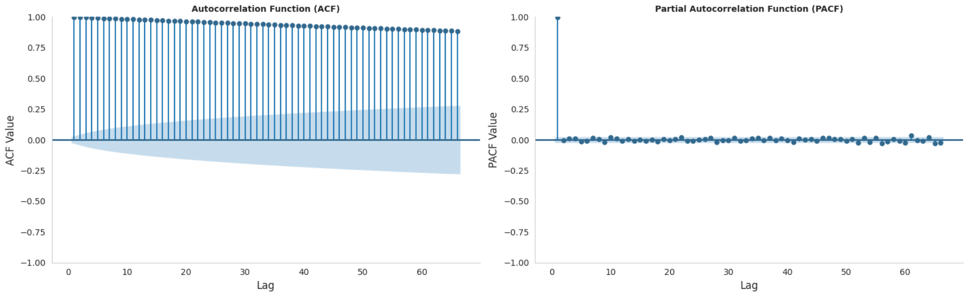



Figure 4: ACF and PACF plots for USD/JPY spot fx rate

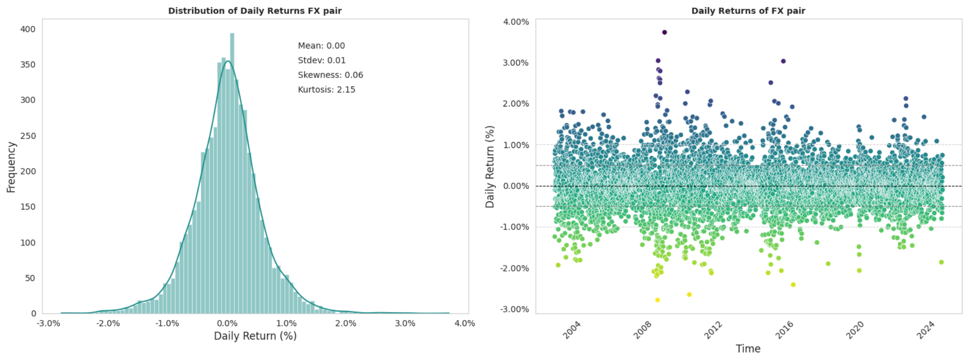
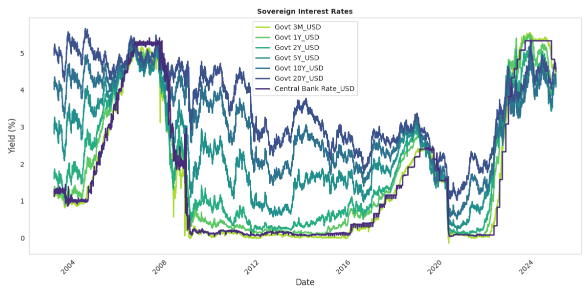



Figure 5: Log returns for EUR/USD spot fx rate

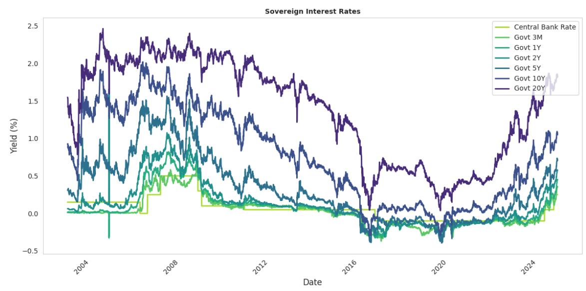
### 4.4 Evolution of Exchange Rates

An Autocorrelation Function (ACF) and a Partial Autocorrelation Function (PACF) were plotted to search for possible autocorrelations within the time series, with the 5% significance level shaded in light blue. The ACF plots indicates very high correlation with past values, but when isolating the lags using a PACF, we notice that the correlation is completely motivated by the first lag, with other values appearing not to be statistically significant. In other words, the current spot rate is highly correlated with the past spot rate, and does not appear to show mean reversion, trends or seasonality within the past quarter, when computing autocorrelation up to 66 daily lags. This indicates that the time series is unpredictable in the long term, non-stationary and that the best prediction value for the next value is simply the previous one observed. This time series shows characteristics of a random walk without drift, as discussed by researchers in the literature, motivating our choice to benchmark our forecasting models against it. To get a glimpse at our forecasted variable, we computed the log returns for daily exchange rates, the daily return distribution, and summary statistics of returns seen below. Although we will use weekly sampling, this decision provides valuable insights into the data. These plots seem to indicate the same conclusions as the autocorrelation plots earlier, as returns seems extremely uncorrelated from one period to the next, at least on a daily scale, the mean and standard deviation are extremely low and the return series shows different levels of kurtosis and skewness, underlying once again fundamental differences between the exchange rates during these periods.

## 4.5 Sovereign yield curves



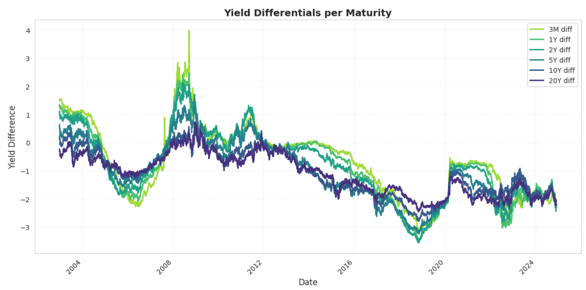

(a) Interest rate term structures (EUR)



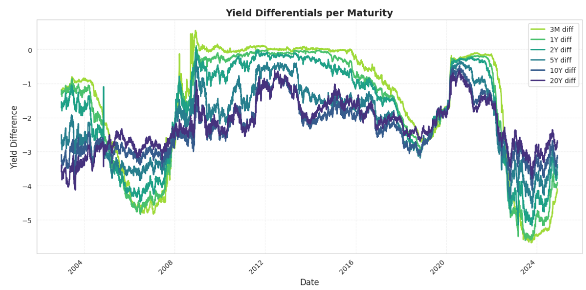

(b) Interest rate term structures (JPY)

Figure 6: Interest rate term structures

We see that interest rates were extremely low during our sample period when compared to past decades, up until the covid pandemic and the inflationary period that followed. It is also worth noting that Japanese and German Bunds (EU proxy) yields reached well below the central bank rates over time, hypothesizing investors preferences for safe haven assets following the GFC and the eurozone crisis. Throughout the past years, we also remark a rather flat yield curve for these economies, which is can the be the result of central banks injecting liquidity into the market through low rates, employing quantitative easing on the long end of the curve and other policies. Nonetheless, the term structures and the interest rate differentials between countries varied during these decades, motivating our decision to include these variables instead of selecting merely one or two measures for interest rates.




(a) Interest rate differentials (EUR minus USD)



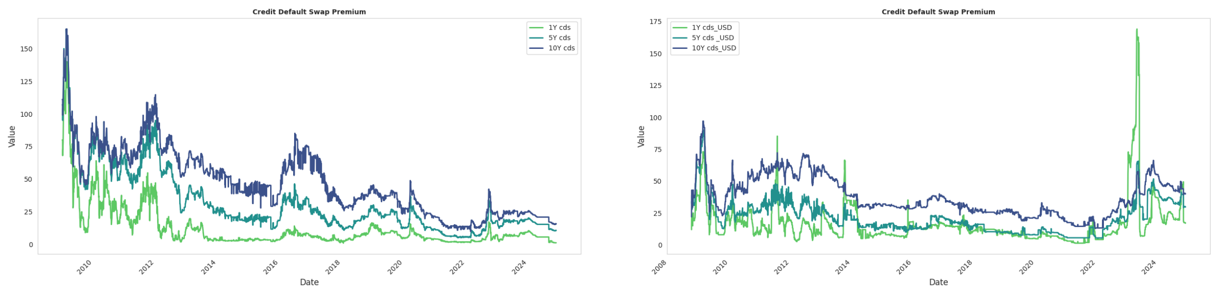

(b) Interest rate differentials (JPY minus USD)

Figure 7: Interest rate differentials

By plotting interest rate differentials, we notice the biggest yield differentials between Japanese and US yields, due to the other central banks leading different monetary policy cycles, while Japanese rates stayed low through the past decades following the lost decades.

## 4.6 Credit Default Swaps

European/German CDS swaps remained higher throughout the period as opposed to the US CDS premiums, due to the impact of the eurozone crisis on all European countries including




(a) CDS term structure for UK Govt bonds

(b) CDS term structure for US Treasuries

Figure 8: CDS term structure for sovereign bonds

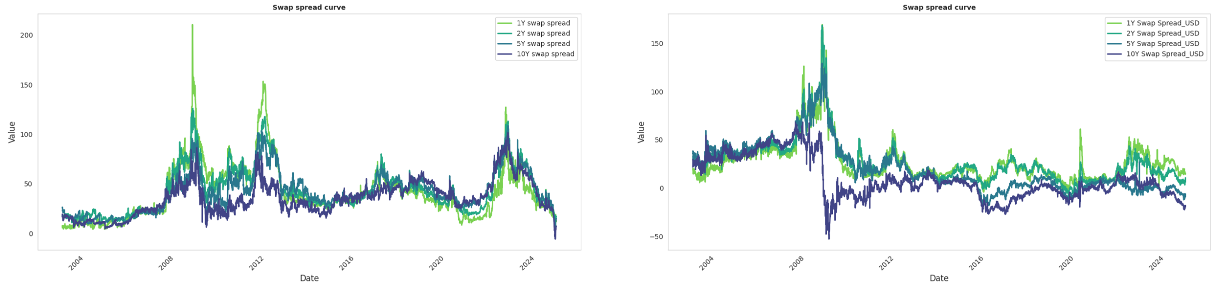
Germany, although premiums for more vulnerable European nations spiked multiple times the values seen here. The UK also seems to have been affected by this period,likely due to the close connection of its economy with the euro area. Overall CDS swaps followed a normal term structure during the period, since risk is logically cumulative over time, but sometimes exceptions were seen clearly indicating risks of credit event uncertainty, such as during 2012 at the peak of the EU crisis for EU German CDS and during early 2023 in the US due to the debt ceiling crisis. The other spikes seen simultaneously throughout these time series indicate other key macroeconomic events with risk priced in such as recessions, banking vulnerability, decisive elections, the COVID pandemic, inflationary spikes and global supply shocks, to name a few. During these periods the widening/ tightening of term structures was seen across asset classes, identical assets across a maturity spectrum have th epossibility to provide information for our forecasting.

### 4.7 Swap rates curve

All swap spreads followed similar trajectories before and during the great financial crisis of 2008, however Germany and the UK lingered higher for longer, as the eurozone crisis took hold of European economies at large with drastic impacts on fiscal and monetary policy, on investment, debt and growth. It is worth noting that US swap spreads spent several months below zero on multiple occasions. This is theorized to be the impact of large US debt issuance and balance sheet constraints of broker dealers for US-Treasuries Boyarchenko et al. [2018].

### 4.8 Implied volatility data for currencies and interest rates

Implied volatility of the different FX pairs and the dollar index moved in tandem. The choice was then made to look at correlation between realized and implied volatility, considering different maturities and rolling window sizes for interest rates of both countries in the FX pair, in addition to implied volatility measures collected earlier. Each dataset proved different trends but overall we can conclude that correlations were mostly low for a lot of variable combinations, indicating that different volatility measures can provide different insights.




(a) Swap Spread structure for Europe against Bunds

(b) Swap Spread term structure for USA

Figure 9: Spreads of swap rate over sovereign bonds of equal maturity

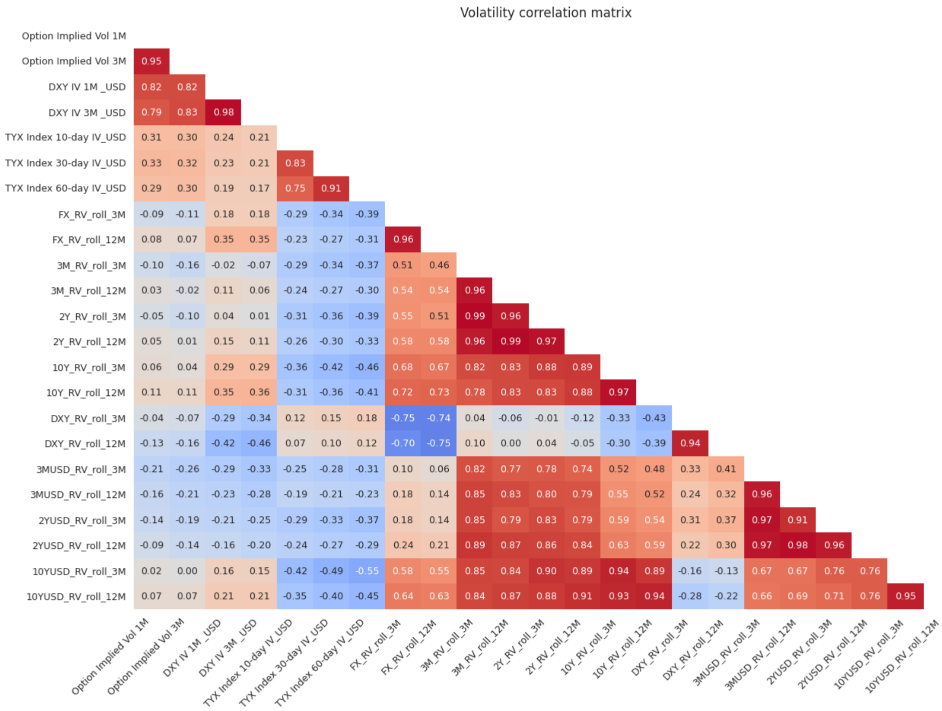



Figure 10: Correlation matrix for GBP and USD volatility variables

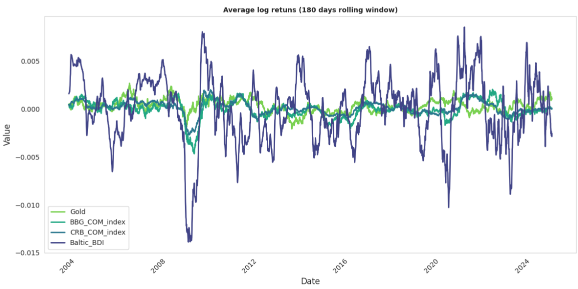
### 4.9 Financial stress and business cycle

We decided to group the global variables into two groups, one with commodities and another one with indicators of market uncertainty, with gold overlapping both groups.We hope to find consistent evidence of risk-on/ risk-off climates by searching consistent correlations between volatility indicators and business cycle metrics such as global demand for commodities and their respective price levels.

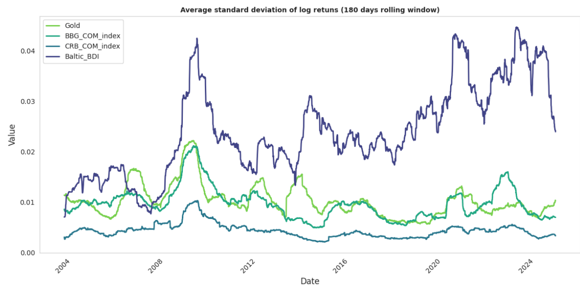
### 4.10 Commodities

Commodity indices are a proxy for global demand, business cycles and price levels. Because both the CRB are the BBG are commodities indices, we expect them to have some correlation, but not necessarily extremely high, as weighting is unequal and less energy focusedd for the

BBG index. The Baltic Dry Index is the international reference for dry bulk shipping across the world, and was also included as a proxy for global demand in commodities, signalling increased production and trade. Both the CRB and the Baltic dry indices have been used previously, as discussed in the literature review section.




(a) Average log returns (180 day rolling window)




(b) Standard deviation of log returns (180 rolling day window)

Figure 11: analysis of log returns of commodity related variables

The Baltic dry index seems like a much more sensitive indicator for commodities and business cycle activity, as seen by rolling logarithmic changes and its standard deviations, and hypothesize that it will provide to our weekly forecasts more informational value than the competing variables seen in the graph below.

### 4.10.1 Financial Stress

Overall, the correlations seem to continuously evolve, and can sometimes completely reverse during exceptional periods such as crisis, fast deleveraging, flight to quality events and other exceptional events. It was hard to find consistent patterns among these global variables so we hope each of them brings additional information into the model regarding different elements of macroeconomic risk. Finally, despite the potential information of these variables regarding global macro, we still need to find if they impacted our specific currency pairs, as they might show more resilience and dependence on commodity and global business cycles.

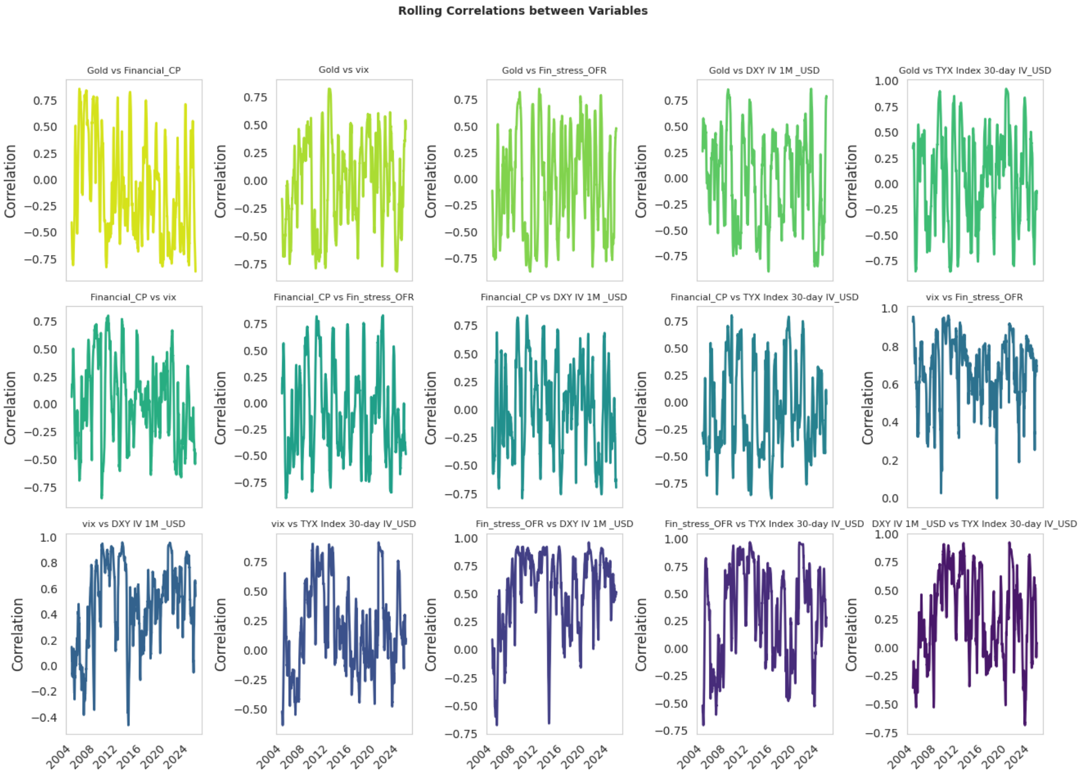



Figure 12: Rolling correlations for macroeconomic variables

## 5 Data preparation

### 5.1 Stationarity

We employed the Augmented Dickey-Fuller (ADF) test to test for the stationarity of the time series, to check for the possible evolution of its properties over time, such as changes in mean, variance and autocorrelations.

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \epsilon_t$$

The Null Hypothesis ($H_0$) is the existence of a unit root (the non-stationarity) of the data, which we will reject given the appropriate p-values. Significance levels are indicated using * for 10%, ** for 5% and *** for 1%.

Table 4: Augmented Dickey-Fuller test for stationarity - USD data

| Variable | p-value | p-value_diff |
|---|---|---|
| Govt 3M_USD | 0.810 | 0.000*** |
| Govt 1Y_USD | 0.748 | 0.000*** |
| Govt 2Y_USD | 0.698 | 0.000*** |
| Govt 5Y_USD | 0.530 | 0.000*** |
| Govt 10Y_USD | 0.295 | 0.000*** |
| Govt 20Y_USD | 0.318 | 0.000*** |
| 1Y cds_USD | 0.000*** | 0.000*** |
| 5Y cds _USD | 0.004*** | 0.000*** |
| 10Y cds_USD | 0.034** | 0.000*** |
| DXY IV 1M _USD | 0.009*** | 0.000*** |
| DXY IV 3M _USD | 0.028** | 0.000*** |
| DXY spot_USD | 0.4979 | 0.000*** |
| 1Y Swap Spread_USD | 0.047** | 0.000*** |
| 2Y Swap Spread_USD | 0.139 | 0.000*** |
| 5Y Swap Spread_USD | 0.523 | 0.000*** |
| 10Y Swap Spread_USD | 0.334 | 0.000*** |
| Central Bank Rate_USD | 0.814 | 0.000*** |
| TYX Index 10-day IV_USD | 0.000*** | 0.000*** |
| TYX Index 30-day IV_USD | 0.000*** | 0.000*** |
| TYX Index 60-day IV_USD | 0.000*** | 0.000*** |

Our tests conclude that most time series were not stationarity initially, but they found stationarity when computing the ADF test on the first differences. Although non-linear models like Random Forests and XGBoost are not impacted by these properties of the data, linear models like OLS and elastic nets will provide less than optimal results if these properties are not considered while using these variables.

### 5.2 Multicollinearity

Multicollinearity refers to the phenomenon that two or more independent variables used in a regression model are highly correlated with each other, having implications on the estimation of coefficients and making it difficult to understand the true contribution of each concerned variable on the predicted variable. This leads to an increased variance of these coefficient estimates and might make them statistically insignificant or at the very least, unreliable. Financial time series by nature always present some degree of correlation and this can sometimes be overcome by models if it is moderate enough. The correlation matrix presented earlier reveals that correlation among variables is not only present but sometimes - due to the use of term structures and yield curves - extremely high among similar variables. To correct for this while keeping the informational value of those time series, we will engineer new features which will have new moments and weaker correlations. Furthermore, several variables were removed from our dataset during validation iterations, to help with this issue impacting model estimation.

## 6 Feature Engineering

Feature engineering can be described as the process of transforming raw data into variables that better represent the prediction problem and overcome model and data shortcomings, using both statistical reasoning and domain expertise. This can include simple transformations such as first differences and log transformations, the creation of lagged variables in time-series to account for lead/lagged indicators or the combination of different series to create new variables and retrieve implicit information in hopes of improving model accuracy, for example. Different variables were created and considered but the choice was made to include here only the ones used during the out-of-sample test to remain concise.

Table 5: Derived Variables and Definitions

| **Variable** | **Definition** |
|---|---|
| **Yield differentials and term structures (UIP)** | |
| diff_3m | Difference between 3M government bond yield and 3M USD yield |
| diff_2y | Difference between 2Y government bond yield and 2Y USD yield |

*(continued on next page)*

(continued from previous page)

| Variable | Definition |
|---|---|
| diff_10y | Difference between 10Y government bond yield and 10Y USD yield |
| term_policy | 2Y yield minus domestic central bank policy rate |
| term_policy_usd | 2Y USD yield minus USD policy rate |
| term_premia | 10Y yield minus 2Y yield (domestic) |
| term_premia_usd | 10Y USD yield minus 2Y USD yield |
| **Exchange rate returns** | |
| carry_past_1M | Log-difference between spot and lagged 1M forward FX rate |
| fx_mom3 | 3-month FX momentum, computed as rolling sum of 1M FX returns |
| **Sovereign risk and financial system stability** | |
| swap_diff_1y | 1Y swap spread minus 1Y USD swap spread |
| swap_diff_5y | 5Y swap spread minus 5Y USD swap spread |
| swap_diff_10y | 10Y swap spread minus 10Y USD swap spread |
| **Isolated pricing of sovereign risk** | |
| cds_level | Log-level of 5Y CDS spread |
| cds_term | Difference in log-levels of 10Y and 1Y CDS spreads |
| cds_level_usd | Log-level of 5Y USD CDS spread |
| cds_term_usd | Difference in log-levels of 10Y and 1Y USD CDS spreads |
| **Market, interest and currency risk** | |
| vix | Log-level of VIX index |
| fx 30dIV level | Log-level of 1M implied FX volatility |
| fx 90dIV level | Log-level of 3M implied FX volatility |
| DXY 30dIV level | Log-level of 1M implied volatility on the USD DXY index |

(continued on next page)

*(continued from previous page)*

| Variable | Definition |
|---|---|
| RV_fx_1M | Realized 1M FX volatility (20-day rolling standard deviation, annualized) |
| TYX 30dIV level | Log-level of 30-day implied volatility on bond future |
| **Global variable transformations** | |
| Financial_CP | Log-level of financial commercial paper rate |
| Gold | Log-level of gold price |
| Baltic_BDI | Log-level of the Baltic Dry Index |

Table 5: Engineered features using available data described in earlier section, solving time series issues and reducing dimensionality of panel data

## 7 Data processing

### 7.0.1 Data splitting

Our final dataset was split into three distinct groups: training, cross-validation and test datasets, as is best practice in this field. This additional data split allows for the test of multiple model prototypes, variables and parameters choices in a "fake" out-of-sample dataset, before finalizing our model training and testing it once in the true out-of-sample portion of the data. This brings flexibility by offering multiple attempts to consider variables and parameters, while avoiding data snooping and look ahead bias. Welch and Goyal [2008] and Goyal et al. [2024] have criticized models models and predictor variables used in financial research which performed apparently good in sample but struggled out-of-sample. By creating a hold-out test dataset from the very beginning and testing model variations in the intermediate validation set, we increase our ability to test and verify our results accurately. All datasets start in January 2003 due to data availability of some variables and end with the last business day of December 2021, marking the period when LIBOR rates began to be replaced by alternatives across most developed markets. In order to consider cds data which began much later (January 2009), we shortened the weekly dataset to skip missing observations, while cds data was not used during the monthly frequency tests to maintain enough observations to train, validate and test those models.

| | EUR | GBP | JPY |
|---|---|---|---|
| Weekly initial train obs. | 275 | 263 | 275 |
| Weekly val obs. | 240 | 240 | 240 |
| Weekly test obs. | 176 | 176 | 176 |
| Weekly total obs. | 691 | 679 | 691 |
| | | | |
| Monthly initial train obs. | 69 | 69 | 69 |
| Monthly val obs. | 96 | 96 | 96 |
| Monthly test obs. | 60 | 60 | 60 |
| Monthly total obs. | 225 | 225 | 225 |

Table 6: Summary of datasets by currency pair ending

### 7.0.2 Standardization

Standardizing time series using computed means and standard deviations is a common practice in machine learning as it helps more data-sensitive algorithms to search for gains in parameter optimization across the panel data instead of focusing on optimizing values in absolute terms, which would disproportionately benefit some variables over others. This transformation to zero mean and unit variance allows the models to learn more effectively from all available information and improve generalizations to unseen data.The standardization was done for each variable as follows:

$$\tilde{\mathbf{x}}_{i,t} = \frac{\mathbf{x}_{i,t} - \bar{\mathbf{x}}_{i\cdot}}{\hat{\sigma}_{i\cdot}}$$

where:

$$\bar{\mathbf{x}}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_{i,t} \quad (1)$$

$$\hat{\sigma}_{i\cdot} = \left(\frac{1}{T-1}\sum_{t=1}^{T}(\mathbf{x}_{i,t} - \bar{\mathbf{x}}_{i\cdot})^2\right)^{0.5} \quad (2)$$

To avoid look ahead bias, the data was split before being standardized. For the validation phase, the mean and standard deviation were computed on the training data and used to standardize the training data and the validation data. A similar approach was used during the testing phase, computing moments on all past data and using it to transform the future variables.

## 8 Machine Learning models

### 8.1 Machine learning in finance

Machine learning can be defined as a branch of artificial intelligence, which has as its primary focus the development of algorithms and techniques to improve predictions through data-driven optimization. Some of its strengths over traditional econometrics are its ability to handle large volumes of data which can be structured (time series, panel data) or unstructured (images, sound, unlabelled data) and its the focus on prediction accuracy without a need for theoretical assumptions for the model or the data. Giglio et al. [2022] and Dixon and Halperin [2019] surveyed the field's literature and categorized most financial machine learning research into the topics of expected return and risk premia estimation, factor modelling, portfolio management, methods for risk management, derivative pricing and optimal hedging models, using mostly data-driven approaches to reach actionable outcomes. Rundo et al. [2019] and Kelly et al. [2023] also discussed the literature covering machine learning applied to financial markets and quantitative finance, showing numerous examples of these methods consistently outperforming traditional approaches. Kelly et al. [2024] concluded that increased model complexity and high volumes of data in certain contexts can strongly outperform parsimonious models out-of-sample, and Gu et al. [2020] discussed the benefits of neural networks and other deep learning models for financial applications, such as the flexibility of model definition and estimation and the superior ability to deal with complex relationships of data often found in financial datasets.

## 8.2 Panel predictive regressions

### 8.2.1 Forecasting model specification

Our forecasting model for weekly exchange rate returns can be written in matrix notation as follows:

$$\hat{y} = \tilde{X}b + \epsilon$$

where our dependent variable is defined as $y_t = \log(\mathrm{Fx}_{t+1}) - \log(\mathrm{Fx}_t)$, which represents the logarithmic returns of the exchange rate for country $i$ at time $_{t+1}$ given the value at time $t$ and $\tilde{X}$ corresponds to the standardized list of independent variables at time $_t$ as discussed in earlier sections, using only information available today to forecast the future exchange rate of the coming week.

### 8.2.2 A comparison with Filippou et al. [2023]

We recall that Filippou et al. [2023] aimed at forecasting an identical $y_t$ variable at a monthly frequency and is therefore the most suitable benchmark for our work. Their model can similarly be written as:

$$\hat{y} = \tilde{X}b + \epsilon$$

where $X$ is a (Z*G) vector representing the Kronecker product of the country characteristics vector $\mathbf{z}_{i,t}$ and the global variables vector $\mathbf{g}_t$ (G x 1), defining their dependent variable as

$y_t = \log(\mathrm{Fx}_t) - \log(\mathrm{Fx}_{t-1})$ for country $i$ at time $t$ using the independent variables listed here:

- *Country Variables* ($\mathbf{z}_{i,t}$)*:* inflation differential, unemployment rate gap differential Christiano and Fitzgerald [2003], bill yield differential note yield differential, bond yield differential, dividend yield differential, price-earnings differential, stock market time-series momentum differential, idiosyncratic volatility and idiosyncratic skewnessLustig et al. [2011];

- *Global Variables* ($\mathbf{g}_{i,t}$)*:* economic policy uncertainty and monetary policy uncertainty Baker et al. [2016], geopolitical risk Caldara and Iacoviello [2022] , global FX volatility and global FX illiquidity Menkhoff et al. [2012b], global FX correlation Mueller et al. [2017].

### 8.3 Benchmark: Random walk without drift

The random walk without drift model assumes that exchange rates follow a path where each new value is expected to equal the previous one plus a random term. The model can be specified as:

$$\mathrm{Fx}_{t+1} = \mathrm{Fx}_t + \varepsilon_t, \qquad \varepsilon_t \sim \text{i.i.d. } (0, \sigma^2) \tag{3}$$

$$\mathbb{E}_t[\mathrm{Fx}_{t+h}] = \mathrm{Fx}_t \qquad \text{for all } h \geq 1. \tag{4}$$

where $\mathrm{Fx}_t$ denotes the exchange rate at time $t$, and $\varepsilon_t$ is a zero-mean, serially uncorrelated error term with constant variance. Under these assumptions, the optimal forecast for $\mathrm{Fx}_{t+h}$ at time $t$ is simply $\mathrm{Fx}_t$, implying an expected logarithmic return of 0 between every two periods.

$$E(y_t) \approx \ln(E(\mathrm{Fx}_{t+1})) - \ln(\mathrm{Fx}_t) \approx \ln(\mathrm{Fx}_t) - \ln(\mathrm{Fx}_t) = 0 \tag{5}$$

This model will serve as our benchmark for exchange rate forecasting due to its simplicity and strong empirical performance. The no-change model performs remarkably well despite its simplicity, beating much more complex economic models in out-of-sample tests, especially in short forecasting horizons. Rossi [2013] reviewed decades of research literature on exchange rates and concluded that no economic model could consistently outperform a random walk without drift in forecasting exchange rates out-of-sample. This stylized fact known as the Meese and Rogoff puzzle Meese and Rogoff [1983] makes this an appropriate benchmark for all exchange rate models.

### 8.4 Model 1: Ordinary Least Squares(OLS) Regression

Using an Ordinary Least Squares (OLS) regression as an additional benchmark for our machine learning models, the out-of-sample forecast for the log return of exchange rates denoted as $\Delta s_{i,t+1}$, using the regression equation including data up to to time $t$ for forecasting:

$$\hat{y} = \tilde{\mathbf{x}}'_{i,t}\hat{\mathbf{b}}_{1:t} \tag{6}$$

where $\hat{\mathbf{b}}_{1:t}$ is the estimate of $\mathbf{b}$ using each time data available up to time $t$. The many assumptions of linear models combined with the large amounts of variables included in the panel regression are likely to lead to over-fitting and poor OOS performance for such a hard forecasting task. We use it mainly to understand the model estimation process and its downfalls, while also having it as an additional comparison for our main models.

### 8.5 Model 2: Elastic Net Regression

The Elastic Net (L1 + L2) algorithm, put forth by Zou and Hastie [2005] is a regression method that provides feature selection and coefficient shrinkage by combining penalty terms from L1 (LASSO) Tibshirani [1996] and L2 (Ridge) Hoerl and Kennard [1970]. This model is capable of exploring relationships between variables outside of linear spaces to improve forecasting capabilities, while also penalizing feature weights and excessive complexity, assisting in model generalization to unseen data. The objective function that Elastic Net algorithm minimizes is as follows:

$$\min_{\beta_0,\beta}\left\{\frac{1}{2n}\sum_{i=1}^{n}(y_i-\beta_0-x_i^T\beta)^2+\lambda\left((1-\alpha)\frac{1}{2}||\beta||_2^2+\alpha||\beta||_1\right)\right\}$$

where $n$ is the number of observations, $y_i$ is the target value, $x_i$ is the vector of features, $\beta_0$ is the intercept term and $\beta$ is the vector of regression coefficients. To this initial regression function, a penalty term is added which contains two distinct parts: $||\beta||_2^2 = \sum_{j=1}^{p}\beta_j^2$, Ridge penalty for large coefficients and $||\beta||_1 = \sum_{j=1}^{p}|\beta_j|$,the Lasso penalty encouraging sparsity. The term $\alpha \in [0,1]$ defines the balance between the L1 and L2 penalty terms, while $\lambda \geq 0$ is the parameter that controls the strength of the the total penalty term of the function.

### 8.6 Model 3: Random Forest

Random Forests are an ensemble machine learning method that can be used both for classification and regression. This model, as explained by Breiman [2001] is trained by building multiple decision trees, sampling different features and observations within the dataset throughout the process, and then averaging their predictions. Random Forests are both more useful and more efficient to train than simple decision tree models because of this process of random sampling, helping to avoid overfitting and performing feature selection, while keeping the other benefits of decision tress such as the ability to deal with non-linear relationships, missing data and providing good predictive performance out-of-sample by generalizing relationships. They are also a very easy to understand model, regarding parameter and hyperparameter selection and feature importance. The Random Forest prediction $\hat{f}_{RF}(x')$ is the average of the predictions from all of its decision trees:

$$\hat{f}_{RF}(x) = \frac{1}{B}\sum_{i=1}^{B} h(x; \Theta_i)$$

where $x$ represents the input features and $y$ a continuous target variable, $B$ is the total number of trees, $\Theta_i$ the random parameters for $i$-th tree (data and feature subsets), $h(x; \Theta_i)$ the prediction of $i$-th tree for input $x$, with results in $\hat{f}_{RF}(x')$ as the final RF prediction for a new input $x'$.

### 8.7 Model 4: Extreme Gradient Boosting (XGB)

Extreme Gradient Boosting*(XGB or XGBoost)* is a machine learning model put forth by Chen and Guestrin [2016] that combines the qualities of both regression trees and gradient boosting models, using techniques such as variable subsampling, sparsity-aware learning, feature selection and various regularization mechanisms, which assist in defining optimal model complexity, while employing gradient descent between each iteration to reduce residual errors and speed up training. XGBoost's ability to quickly optimize the bias and variance of models, defined as the search for the optimal balance between complexity and ability to perform on different datasets, is another one of its strongest assets, which regularly puts it ahead of other algorithms and methods in the field of forecasting. The XGB ensemble prediction can be defined as:

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(\mathbf{x}_i)$$

where $x_i$ represents the input features, $y_i$ a continuous target variable for the $i$-th observation and $\hat{y}_i^{(t)}$ is the predicted value after $t$ boosting rounds, built by sequentially adding the predictions $f_t(x_i)$ of each new tree. This process aims to minimize the objective loss function $\mathcal{L}^{(t)}$:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \ell(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t)$$

This function balances the loss $\ell$ between the true $y_i$ and the predicted value $\hat{y}_i$, while also including a regularization term $\Omega(f_t)$ that penalizes the complexity of the $t$-th tree:

$$\Omega(f_t) = \gamma T_t + \frac{1}{2}\lambda \sum_{j=1}^{T_t} w_{j,t}^2$$

where $T_t$ is the number of leaves, $w_{j,t}$ the weight of the $j$-th leaf, $\gamma$ is the the penalty per leaf, and $\lambda$ is the penalty for squared leaf weights. This regularization aims at encouraging complexity to improve predictions while also restraining it to avoid overfitting.

## 9 Model selection and training

After data exploration, cleaning, transformation and preprocessing, different variations of the model were built and tested using the validation dataset. This step of prototyping models had three different goals: 1) understand the value of the variables chosen and test different ways to extract more information from them, 2) establish the appropriate parameters for our model architectures and 3) define the appropriate lags and time dependent features to use with our time series, to account for time dependency.

### 9.1 Optuna framework for choice of parameters

Model parameters are the specifications of a Machine Learning algorithm defining its constraints, and given identical model architecture and datasets, a poor definition of parameters can lead to widely different results: therefore, one can't overstate the importance of this step of model selection. Two of the most frequent options for parameter optimization are GridSearch, which goes through an exhaustive combination of parameters (very resource intensive) and RandomSearch, which tests a variety of hyperparameter combinations using random sampling from distributions defined by the user. RandomSearch is generally preferred for baseline models as it is more resource efficient, although the optimized values might be subject to a small degree of randomness. Optuna, an open-source hyperparameter framework was employed for this task, which provides an alternative to both of these techniques by relying on Bayesian optimization (Tree-structured Parzen Estimator) and parallelization (running multiple trials at a given time). This method of optimization develops preferences for sampling values from distribution areas which provided better scores in previous trials and avoids value ranges with less promising past results, allowing for both a very resource efficient method than GridSearch and a more precise and less random process than RandomSearch.

### 9.2 Validation of the model

During each iteration of our optimization function, values were selected from ranges defined earlier for the parameters and hyperparameters, leading to a specific model structure which was tested using different iterations of training and predicting on an expanding dataset. Thanks to Optuna's probabilistic approach, parameters which led to successful results in past iterations gained more weight and those that led to bad results lost importance when it came to choose the next set of parameter values, leading to a efficient exploration of possibilities and a fast improvement in the test scores of the successive models during the optimization phase. The existence of an additional hold-out dataset (test data) also held us accountable to avoid extensively iterating through data and tinkering which may improve results in-sample but may lead to results out-of-sample.

| **Validation MSE** | $EUR_w$ | $GBP_w$ | $JPY_w$ | $EUR_m$ | $GBP_m$ | $JPY_m$ |
|---|---|---|---|---|---|---|
| Random Walk | 1.4004 | 1.4955 | 1.7413 | 7.5557 | 5.9667 | 7.4169 |
| OLS regression | **1.3903** | 1.5182 | 1.7831 | 7.5705 | **5.7949** | **7.1222** |
| Elastic Net | 1.4242 | 1.5328 | 1.7870 | **7.4727** | 5.9728 | 7.5937 |
| Random Forest | 1.4375 | 1.5412 | 1.8206 | **7.2952** | 6.1599 | **7.4119** |
| XGBoost | 1.4086 | 1.5344 | **1.7786** | **7.4243** | 5.8122 | 7.5005 |

Table 7: Mean Squared Error (MSE) for validation dataset predictions against the actual log returns of currency pairs. Bold indicates models that outperform the Random Walk baseline.

### 9.3 Optimization metric

The metric chosen for our optimization function was the Mean-Squared-Error (MSE), which penalizes large errors more heavily, appropriate in a context of return prediction in assets with sometimes extreme moves. The RMSE and the Mean Absolute Error (MAE) were considered but not chosen over the MSE because of this penalization. The Mean Absolute Percentage Error (MAPE), another very popular metric for optimization, was not chosen because it struggles when values are very small or close to zero, which is the case of exchange rate returns. The expanding window approach was used during the validation phase, using a nested-cross validation approach, where the train data is split into different sections (folds), leading to several simulated out-of-sample performances, the average mean-squared-error of which was used to assess the performance of the models over different time periods. Models using weekly data were trained and validated using 4 folds, each containing 60 data points, while models using monthly data were more limited due to the size of the data, and only 4 folds were created containing 24 observations each, all of which were sequential and represented the later half of the training dataset of each currency. This adds up to the data splitting described in an earlier table. The choice to skip a period between every train and validation dataset in every fold was made, to avoid data leakage issues.

## 10 Results

The optimal model structure was decided based on the best performing iterations obtained through the optimization routines of each model using the validation data. The training data and validation were then merged, to give a bigger sample for the initial training of the final model, using again the expanding window methodology.

### 10.1 Statistical Accuracy

The choice was made to compare the forecasts with the no-change benchmark discussed in Rossi [2013], measuring our performance using our out-of-sample accuracy as discussed in

Fama and French [1989] and Campbell and Thompson [2008], equally employed by Filippou et al. [2023] as a benchmark for their models. Out-of-sample $R^2$ statistics were computed on a dataset unknown to the algorithms during model estimation and validation, measuring the model's performance against our no-change benchmark, measuring the relative improvement in model Mean Squared Prediction Error(MSPE). Out-of-sample $R^2$ were computed:

$$R^2_{i,\mathrm{OS}} = 1 - \frac{\sum_{t=T_1+1}^{T_2} \left(\hat{e}^{\mathrm{Forecast}}_{i,t|t-1}\right)^2}{\sum_{t=T_1+1}^{T_2} \left(\hat{e}^{\mathrm{Benchmark}}_{i,t|t-1}\right)^2} \tag{7}$$

A Clark-West test was also computed following Clark and West [2007] to understand the statistical significant of our results, accounting for different levels of model complexity. The Clark-West test statistic is computed from a regression of the loss differential between the two models on a constant, where $\ell_{i,t}$ is the adjusted loss differential accounting for competing model complexity, $a_{0,i}$ is the intercept, and $\varepsilon_{i,t}$ is the error term.

$$\ell_{i,t} = \left(\hat{e}^{\mathrm{Benchmark}}_{i,t|t-1}\right)^2 - \left(\hat{e}^{\mathrm{Forecast}}_{i,t|t-1}\right)^2 + \left(\Delta\hat{s}^{\mathrm{Benchmark}}_{i,t|t-1} - \Delta\hat{s}^{\mathrm{Forecast}}_{i,t|t-1}\right)^2 = a_{0,i} + \varepsilon_{i,t} \tag{8}$$

| **OOS** $R^2$ | $EUR_w$ | $GBP_w$ | $JPY_w$ | $EUR_m$ | $GBP_m$ | $JPY_m$ |
|---|---|---|---|---|---|---|
| OLS | -0.012 | -0.007 | -0.019 | -0.092 | -0.027 | -0.002 |
| Elastic Net | -0.001 | -0.001 | -0.000 | -0.026 | -0.011 | -0.008 |
| Random Forest | -0.012 | -0.020 | -0.029 | -0.241 | -0.076 | -0.156 |
| XGBoost | -0.011 | -0.005 | -0.003 | -0.039 | -0.051 | -0.033 |

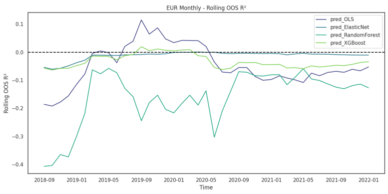
Table 8: Out-of-sample $R^2$ for each model by currency and frequency, with Clark-West test significance noted as ****,**,* for 1%, 5% and 10% respectively.

As the table reveals, the out-of-sample performance of the models was consistently lacklustre, something that was expected as the mean squared error and other metrics computed during the validation period pointed towards that conclusion for several models. Nonetheless, it is surprising that none managed to outperform during the period. This is likely a consequence of a poor choice and treatment of variables, as well as of a weak validation phase which didn't explore enough the importance of features and parameters, which given such a hard task, is something that should be avoided at all costs. The Clark-West statistics seem to indicate no significant advantage or disadvantage in most cases, as the models' predictions often laid closely to the random walk benchmark. As most results don't show significance using the CW-test, it is hard to reach definite conclusions, we can however notice that for these periods, the more complex models like Random Forests and XGBoost underperformed sometimes easier benchmarks, as their predictions strayed from the benchmark more often, especially during the initial COVID period, impacting its global performance due to the small size of the test data.

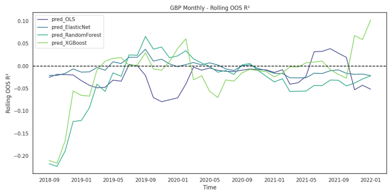
## 10.2 Performance over time

Given the many economic and political events that took place during the test period, the test data was in many ways different from both the training and the validation date, therefore the choice was made to compute rolling metrics for the performance of the forecasts to understand their evolution over these many events which induced a lot of exchange rate volatility across developed markets.

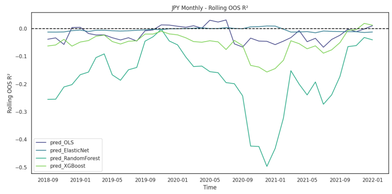
### 10.2.1 Rolling OOS $R^2$



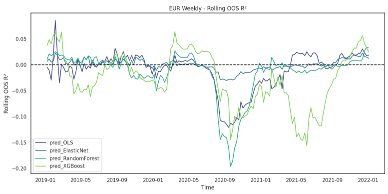

(a) EUR Monthly



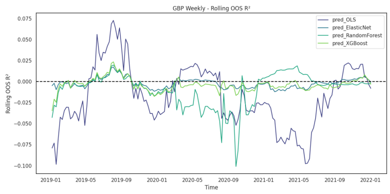

(b) GBP Monthly



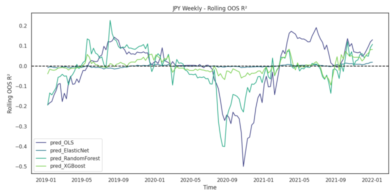

(c) JPY Monthly




(d) EUR Weekly



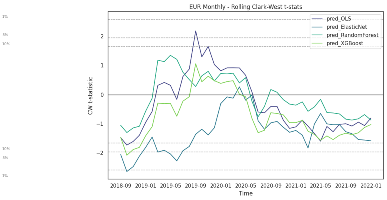

(e) GBP Weekly



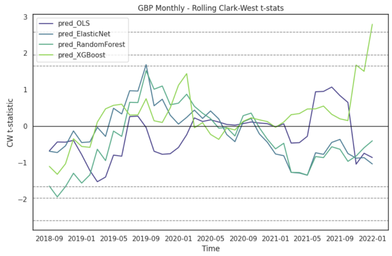

(f) JPY Weekly

Figure 13: Rolling OOS $R^2$ for all currencies and frequencies, comparing the predicted power of each model against the random walk benchmark, with values above zero indicating superior predictive performance from the model.

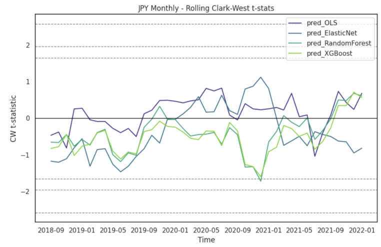
### 10.2.2 Rolling Clark West Statistic



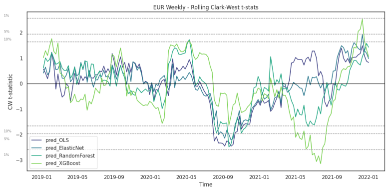

(a) EUR Monthly



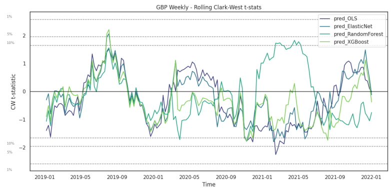

(b) GBP Monthly



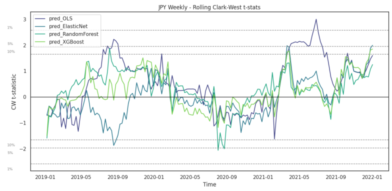

(c) JPY Monthly




(d) EUR Weekly




(e) GBP Weekly



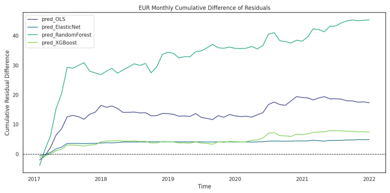

(f) JPY Weekly

Figure 14: Rolling Clark-West statistic for all currencies and frequencies with significance levels indicated with horizontal lines, indicating the significance of outperformance/ underperformance of each model against the benchmark

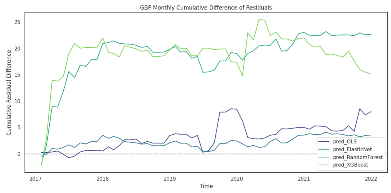
When assessing the rolling statistics and the cumulative difference in residuals, it is apparent that the covid pandemic situated in the middle of the test data led to strong underperformance of models for a few periods, which translated into a spike in cumulative residual differences, which exhibited stable trends before and after the event. This also impacted the rolling R2 and rolling CW statistic, with both indicating significant strong underperformance during that period.

### 10.2.3 Cumulative Residuals

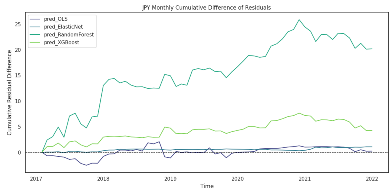
The stable trend seen in residuals in the beginning and end of the test period reveals that most models' predictions were quite close to 0 most times, leading to a small difference in residuals against the random walk during most periods, with the occasional larger prediction that brought the residual higher or lower depending on its success. During the phase of training the model, the predictions made by the models most likely generated large prediction errors during most periods, leading models to reduce or omit the influence of most variables, and nudge its predictions close to zero, in an attempt to minimize the loss function, another strong sign of the random walk nature of the predicted variable.



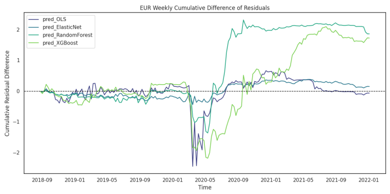

(a) EUR Monthly



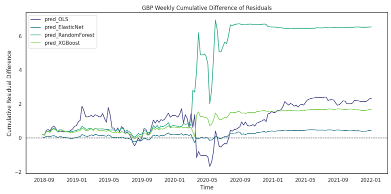

(b) GBP Monthly



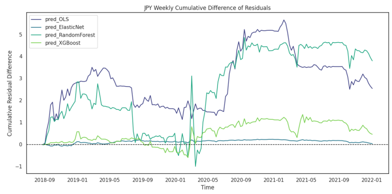

(c) JPY Monthly




(d) EUR Weekly




(e) GBP Weekly




(f) JPY Weekly

Figure 15: Cumulative differences in residuals for all currencies and frequencies, considering the predictions of each model against the prediction of the random walk benchmark

## 11 Conclusion

We sought to apply machine learning methods to the notoriously difficult task of predicting short exchange rates, choosing a wide array of variables related to both currencies of each Fx pair in our dataset, putting together ideas from past research, economic theory and other financial market indicators, all of which were resampled to weekly and monthly frequencies. Different models were chosen and optimized using training and validation datasets, employing common methodologies from the field of machine learning. This led to unsuccessful results, and our predictions joined the vast amount of models that under-performed a simple random walk without drift. These results do not necessarily lead to the conclusion that the data or the techniques used are unable to produce results in exchange rate prediction but rather indicates that the choices made during this project led to a weak performance during our validation and test periods. A more careful selection of variables, a more extensive model validation phase and deeper knowledge of the fields of finance, statistics and machine learning would have contributed to better model specifications and results. This unsuccessful attempt also highlights the necessity of developing strong domain knowledge in different fields to be able to employ these tools successfully, especially when dealing with complex prediction tasks, proving that ML and AI are not a magical solution for all tasks but rather a solution to be considered and adapted to each task, if it is deemed an appropriate avenue.

The major limitations of this research were the shortened time-frame of the regression analysis and the small pool of currency pairs used, both a consequence of data unavailability, despite numerous currencies being initially considered. With the emergence of cheap hardware and the development of data science as a field, data collection and availability is expected to increase, allowing for more complete analysis of these relationships in the future. This could mean extending the forecasting analysis of the three pairs by using longer time series, or even adapting the approach to analyse additional developed, emerging and frontier markets. Other contributing factors that limited the scope and results of this project were a non-extensive knowledge of the field of machine learning by the author which severely impacted the results of such a arduous prediction task. Researchers interested in the topic might develop upon this work by considering more advanced feature engineering and data transforming, more extensive validation phases and by learning from other fields of research which have attempted prediction tasks on data posessing similar characteristics.

## References


Dilip Abreu and Markus K Brunnermeier. Bubbles and crashes. *Econometrica*, 71(1):173–204, 2003.

Tobias Adrian, Erkko Etula, and Hyun-Song Shin. Risk appetite and exchange rates. Technical report, Staff Report, 2015.

Andrew Ang and Joseph Chen. Yield curve predictors of foreign exchange returns. In *AFA 2011 Denver Meetings Paper*, 2010.

Clifford S Asness, Tobias J Moskowitz, and Lasse Heje Pedersen. Value and momentum everywhere. *The journal of finance*, 68(3):929–985, 2013.

Philippe Bacchetta and Eric Van Wincoop. Random walk expectations and the forward discount puzzle. *American Economic Review*, 97(2):346–350, 2007.

Scott R Baker, Nicholas Bloom, and Steven J Davis. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636, 2016.

Gurdip Bakshi and George Panayotov. Predictability of currency carry trades and asset pricing implications. *Journal of financial economics*, 110(1):139–163, 2013.

Robert J Barro and José F Ursúa. Rare macroeconomic disasters. *Annu. Rev. Econ.*, 4(1):83–109, 2012.

Pedro Barroso and Pedro Santa-Clara. Beyond the carry trade: Optimal currency portfolios. *Journal of Financial and Quantitative Analysis*, 50(5):1037–1056, 2015.

Pedro Barroso, Frickson Kho, Florent Rouxelin, and Li Yang. Cutting the gordian knot of carry and imbalances. *Available at SSRN 3232396*, 2018.

Travis Berge, Òscar Jordà, and Alan M Taylor. Currency carry trades. In *NBER International Seminar on Macroeconomics*, volume 7, pages 357–388. University of Chicago Press Chicago, IL, 2011.

Benjamin M Blau. Skewness preferences, asset prices and investor sentiment. *Applied Economics*, 49(8):812–822, 2017.

Nina Boyarchenko, Pooja Gupta, Nick Steele, and Jacqueline Yen. Negative swap spreads. *Economic Policy Review*, 24(2):1–14, 2018.

Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

Markus K Brunnermeier, Stefan Nagel, and Lasse H Pedersen. Carry trades and currency crashes. *NBER macroeconomics annual*, 23(1):313–348, 2008.

Craig Burnside, Martin Eichenbaum, and Sergio Rebelo. Carry trade: The gains of diversification. *Journal of the European Economic Association*, 6(2-3):581–588, 2008.

Craig Burnside, Martin Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo. Do peso problems explain the returns to the carry trade? *The Review of Financial Studies*, 24(3):853–891, 2011.

Dario Caldara and Matteo Iacoviello. Measuring geopolitical risk. *American economic review*, 112(4):1194–1225, 2022.

Giovanni Calice and Ming Zeng. The term structure of sovereign credit default swap and the cross-section of exchange rate predictability. *International Journal of Finance & Economics*, 26(1):445–458, 2021.

John Y Campbell and Samuel B Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531, 2008.

Rachel Campbell, Kees Koedijk, and Paul Kofman. Increased correlation in bear markets. *Financial Analysts Journal*, 58(1):87–94, 2002.

Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

Lawrence J Christiano and Terry J Fitzgerald. The band pass filter. *International economic review*, 44(2):435–465, 2003.

Todd E Clark and Kenneth D West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, 138(1):291–311, 2007.

Pasquale Della Corte, Steven J Riddiough, and Lucio Sarno. Currency premia and global imbalances. *The Review of Financial Studies*, 29(8):2161–2193, 2016.

Pasquale Della Corte, Lucio Sarno, and Ilias Tsiakas. Volatility and correlation timing in active currency management. *Handbook of exchange rates*, pages 421–447, 2012.

Matthew Francis Dixon and Igor Halperin. The four horsemen of machine learning in finance. *Available at SSRN 3453564*, 2019.

Eugene F. Fama. Forward and spot exchange rates. *Journal of Monetary Economics*, 14(3): 319–338, 1984. ISSN 0304-3932. doi: https://doi.org/10.1016/0304-3932(84)90046-1. URL `https://www.sciencedirect.com/science/article/pii/0304393284900461`.

Eugene F Fama and Kenneth R French. Business conditions and expected returns on stocks and bonds. *Journal of financial economics*, 25(1):23–49, 1989.

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

Emmanuel Farhi and Xavier Gabaix. Rare disasters and exchange rates: A theory of the forward premium puzzle. Technical report, October 2007. Working Paper Harvard University, 2008.

Ilias Filippou, David Rapach, Mark P Taylor, and Guofu Zhou. Economic fundamentals and short-run exchange rate prediction: A machine learning perspective. *Available at SSRN 3455713*, 2023.

Kenneth A Froot and Richard H Thaler. Anomalies: foreign exchange. *Journal of economic perspectives*, 4(3):179–192, 1990.

Stefano Giglio, Bryan Kelly, and Dacheng Xiu. Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics*, 14(1):337–368, 2022.

Amit Goyal, Ivo Welch, and Athanasse Zafirov. A comprehensive 2022 look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 37(11):3490–3557, 2024.

Shihao Gu, Bryan Kelly, and Dacheng Xiu. Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5):2223–2273, 02 2020. ISSN 0893-9454. doi: 10.1093/rfs/hhaa009. URL `https://doi.org/10.1093/rfs/hhaa009`.

Campbell R Harvey and Akhtar Siddique. Conditional skewness in asset pricing tests. *The Journal of finance*, 55(3):1263–1295, 2000.

Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

Intercontinental Exchange (ICE). Ice dollar index futures (dx), 2025a. URL `https://www.ice.com/products/194/US-Dollar-Index-Futures`. Accessed: 2025-02-15.

Intercontinental Exchange (ICE). Ice dollar index (usdx), 2025b. URL `https://www.ice.com/forex/usdx`. Accessed: 2025-02-15.

Òscar Jordà and Alan M Taylor. The carry trade and fundamentals: Nothing to fear but feer itself. *Journal of International Economics*, 88(1):74–90, 2012.

Petri Jylhä and Matti Suominen. Speculative capital and currency carry trades. *Journal of Financial Economics*, 99(1):60–75, 2011. URL `https://EconPapers.repec.org/RePEc:eee:jfinec:v:99:y:2011:i:1:p:60-75`.

Nina Karnaukh, Angelo Ranaldo, and Paul Söderlind. Understanding fx liquidity. *The Review of Financial Studies*, 28(11):3073–3108, 2015.

Bryan Kelly, Dacheng Xiu, et al. Financial machine learning. *Foundations and Trends® in Finance*, 13(3-4):205–363, 2023.

Bryan Kelly, Semyon Malamud, and Kangying Zhou. The virtue of complexity in return prediction. *The Journal of Finance*, 79(1):459–503, 2024.

Hanno Lustig, Nikolai Roussanov, and Adrien Verdelhan. Common risk factors in currency markets. *The Review of Financial Studies*, 24(11):3731–3777, 2011.

Richard A Meese and Kenneth Rogoff. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, 14(1-2):3–24, 1983.

Michael Melvin and Mark P Taylor. The crisis in the foreign exchange market. *Journal of International Money and finance*, 28(8):1317–1330, 2009.

Lukas Menkhoff, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. Currency momentum strategies. *Journal of Financial Economics*, 106(3):660–684, 2012a. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2012.06.009. URL `https://www.sciencedirect.com/science/article/pii/S0304405X12001353`.

Lukas Menkhoff, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. Carry trades and global foreign exchange volatility. *The Journal of Finance*, 67(2):681–718, 2012b.

Lukas Menkhoff, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. Currency risk premia and macro fundamentals. *Available at SSRN*, 2013.

Philippe Mueller, Andreas Stathopoulos, and Andrea Vedolin. International correlation risk. *Journal of Financial Economics*, 126(2):270–299, 2017.

Angelo Ranaldo and Paul Söderlind. Safe haven currencies. *Review of finance*, 14(3):385–407, 2010.

Robert Ready, Nikolai Roussanov, and Colin Ward. Commodity trade and the carry trade: a tale of two countries. Working Paper 19371, National Bureau of Economic Research, August 2013. URL `http://www.nber.org/papers/w19371`.

Barbara Rossi. Exchange rate predictability. *Journal of economic literature*, 51(4):1063–1119, 2013.

Francesco Rundo, Francesca Trenta, Agatino Luigi di Stallo, and Sebastiano Battiato. Machine learning for quantitative finance applications: A survey. *Applied Sciences*, 9(24), 2019. ISSN 2076-3417. doi: 10.3390/app9245574. URL `https://www.mdpi.com/2076-3417/9/24/5574`.

Andrew Szakmary, Evren Ors, Jin Kyoung Kim, and Wallace N Davidson Iii. The predictive power of implied volatility: Evidence from 35 futures markets. *Journal of Banking & Finance*, 27(11):2151–2175, 2003.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288, 1996.

Ivo Welch and Amit Goyal. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508, 2008.

Ming Zeng. Currency carry, momentum, and global interest rate volatility. *Journal of Financial and Quantitative Analysis*, pages 1–35, 2019.

Lu Zhang and Lei Hua. Major issues in high-frequency financial data analysis: A survey of solutions. *Mathematics*, 13(3):347, 2025.

Shaojun Zhang. Dissecting currency momentum. *Journal of Financial Economics*, 144(1): 154–173, 2022.

Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301–320, 2005.

## 12 Appendix

Table 9: Summary statistics for GBP related variables

| Variable | Count | Mean | Std.Dev. | Min. | 25% | 50% | 75% | Max. | Mean | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Fx spot | 5718.000 | 1.533 | 0.236 | 1.069 | 1.309 | 1.543 | 1.669 | 2.108 | 1.533 | 0.396 | -0.848 |
| Govt 3M | 5718.000 | 1.948 | 1.979 | -0.164 | 0.408 | 0.601 | 4.233 | 5.886 | 1.948 | 0.618 | -1.407 |
| Govt 1Y | 5718.000 | 1.947 | 1.979 | -0.156 | 0.398 | 0.665 | 4.229 | 5.816 | 1.947 | 0.596 | -1.438 |
| Govt 2Y | 5718.000 | 2.031 | 1.900 | -0.134 | 0.450 | 0.966 | 4.203 | 5.823 | 2.031 | 0.533 | -1.435 |
| Govt 5Y | 5718.000 | 2.420 | 1.702 | -0.134 | 0.894 | 1.921 | 4.185 | 5.790 | 2.420 | 0.263 | -1.433 |
| Govt 10Y | 5718.000 | 2.884 | 1.509 | 0.104 | 1.501 | 2.908 | 4.312 | 5.537 | 2.884 | -0.114 | -1.394 |
| Govt 20Y | 5718.000 | 3.285 | 1.320 | 0.514 | 1.985 | 3.549 | 4.474 | 5.172 | 3.285 | -0.420 | -1.245 |
| Central Bank Rate | 5718.000 | 2.036 | 2.054 | 0.100 | 0.500 | 0.500 | 4.500 | 5.750 | 2.036 | 0.614 | -1.450 |
| 1Y swap spread | 4958.000 | 44.035 | 27.132 | 2.831 | 27.440 | 36.231 | 52.269 | 174.894 | 44.035 | 1.670 | 3.207 |
| 2Y swap spread | 4958.000 | 46.806 | 23.930 | 2.450 | 32.810 | 38.728 | 54.848 | 160.482 | 46.806 | 1.430 | 1.977 |
| 5Y swap spread | 4958.000 | 36.655 | 17.675 | 1.070 | 25.125 | 32.300 | 41.490 | 115.460 | 36.655 | 1.300 | 1.696 |
| 10Y swap spread | 4958.000 | 21.111 | 15.352 | -24.750 | 9.049 | 19.738 | 32.603 | 70.375 | 21.111 | 0.268 | -0.272 |
| 1Y cds | 4151.000 | 12.306 | 17.197 | 0.910 | 3.520 | 6.480 | 12.064 | 140.000 | 12.306 | 3.663 | 17.292 |
| 5Y cds | 4151.000 | 31.641 | 24.685 | 5.040 | 15.810 | 22.000 | 40.228 | 165.000 | 31.641 | 1.818 | 4.125 |
| 10Y cds | 4151.000 | 47.268 | 25.363 | 9.850 | 25.250 | 41.410 | 66.130 | 165.000 | 47.268 | 0.877 | 1.069 |
| Option Implied Vol 1M | 5718.000 | 8.961 | 3.085 | 4.335 | 7.145 | 8.200 | 10.024 | 29.622 | 8.961 | 2.470 | 9.004 |
| Option Implied Vol 3M | 5718.000 | 9.147 | 2.762 | 4.900 | 7.442 | 8.351 | 10.250 | 24.900 | 9.147 | 1.935 | 5.512 |

Table 10: Summary statistics for global variables

| Variable | Count | Mean | Std.Dev. | Min. | 25% | 50% | 75% | Max. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Gold | 5718.000 | 1246.890 | 527.468 | 322.750 | 874.250 | 1267.600 | 1661.668 | 2787.610 | 0.049 | -0.505 |
| Financial_CP | 5718.000 | 335.535 | 118.347 | 174.051 | 246.959 | 281.586 | 419.864 | 657.408 | 0.940 | -0.308 |
| vix | 5718.000 | 19.098 | 8.494 | 9.140 | 13.472 | 16.740 | 21.910 | 82.690 | 2.527 | 9.693 |
| TEDRATE | 4958.000 | 0.405 | 0.413 | 0.060 | 0.200 | 0.280 | 0.430 | 4.580 | 3.958 | 22.028 |
| EPU_JPY | 5718.000 | 106.547 | 31.666 | 48.395 | 84.621 | 103.535 | 122.839 | 239.279 | 1.127 | 2.179 |
| EPU_EUR | 5718.000 | 186.883 | 83.539 | 47.692 | 120.001 | 179.658 | 235.986 | 433.278 | 0.584 | -0.064 |
| BBG_COM_index | 5718.000 | 122.047 | 34.588 | 59.480 | 89.502 | 120.577 | 145.992 | 237.953 | 0.454 | -0.412 |
| CRB_COM_index | 5718.000 | 474.960 | 92.901 | 248.560 | 427.338 | 491.540 | 536.265 | 689.010 | -0.496 | -0.196 |
| Fin_stress_OFR | 5718.000 | -0.151 | 1.808 | -2.008 | -1.200 | -0.643 | 0.111 | 13.279 | 2.990 | 12.062 |
| Baltic_BDI | 5718.000 | 2349.173 | 1986.967 | 290.000 | 1061.000 | 1704.500 | 2868.000 | 11793.000 | 2.096 | 4.916 |
| EPU_GBP | 5718.000 | 291.665 | 197.235 | 0.000 | 155.852 | 248.300 | 382.205 | 2610.060 | 2.161 | 11.159 |
| EPU_USD | 5718.000 | 112.776 | 82.694 | 3.320 | 59.732 | 91.295 | 140.928 | 1026.380 | 2.575 | 11.415 |

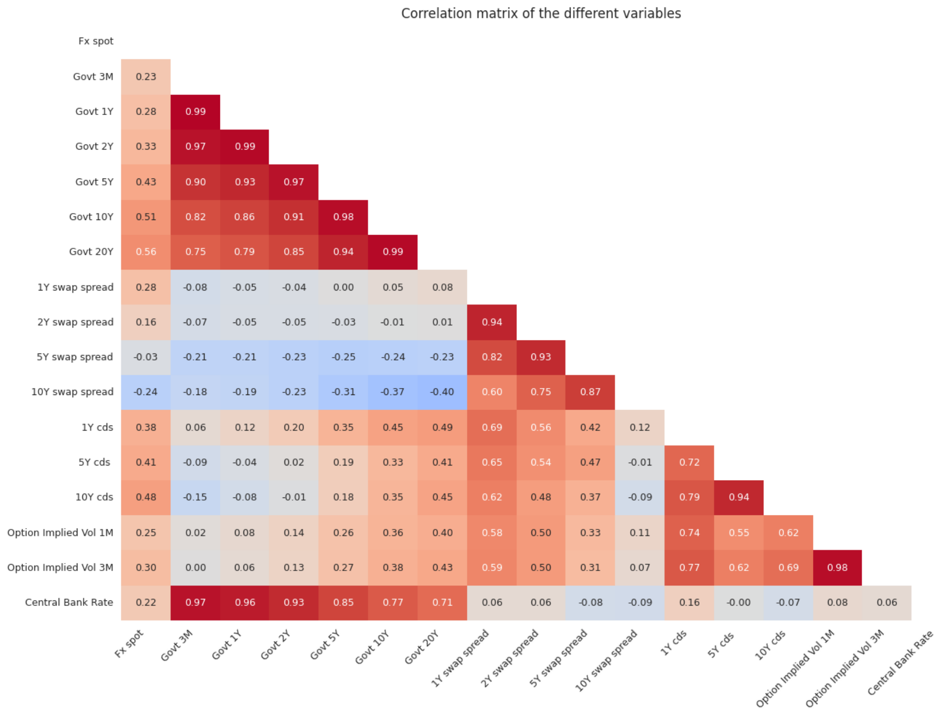



Figure 16: Correlation Matrix for EUR related variables

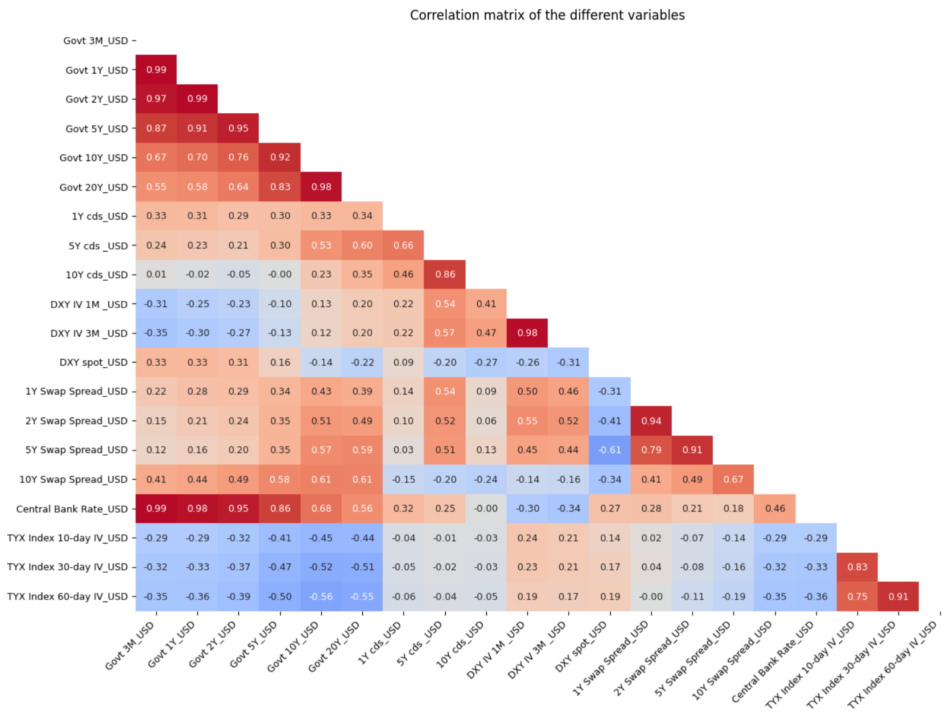



Figure 17: Correlation Matrix for USD related variables

Table 11: Summary statistics for JPY related variables

| Variable | Count | Mean | StdDev. | Min. | 25% | 50% | 75% | Max. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Fx spot | 5718.000 | 109.728 | 17.181 | 75.820 | 101.846 | 109.495 | 117.569 | 161.695 | 0.433 | 0.570 |
| Govt 3M | 5718.000 | 0.021 | 0.196 | -0.370 | -0.126 | 0.010 | 0.096 | 0.670 | 0.784 | 0.341 |
| Govt 1Y | 5718.000 | 0.058 | 0.240 | -0.363 | -0.126 | 0.018 | 0.120 | 0.822 | 1.192 | 1.070 |
| Govt 2Y | 5718.000 | 0.119 | 0.293 | -0.367 | -0.118 | 0.081 | 0.183 | 1.721 | 1.216 | 1.085 |
| Govt 5Y | 5718.000 | 0.330 | 0.438 | -0.396 | -0.071 | 0.229 | 0.592 | 1.564 | 0.751 | -0.229 |
| Govt 10Y | 5718.000 | 0.765 | 0.611 | -0.297 | 0.112 | 0.754 | 1.326 | 2.020 | 0.119 | -1.289 |
| Govt 20Y | 5718.000 | 1.380 | 0.670 | 0.036 | 0.637 | 1.536 | 2.014 | 2.464 | -0.342 | -1.349 |
| Central Bank Rate | 5718.000 | 0.060 | 0.164 | -0.100 | -0.100 | 0.050 | 0.150 | 0.500 | 1.273 | 1.444 |
| 1Y swap spread | 4958.000 | 18.872 | 10.861 | 0.100 | 11.350 | 16.010 | 23.695 | 55.000 | 0.946 | 0.222 |
| 2Y swap spread | 4958.000 | 16.781 | 8.585 | 1.800 | 10.638 | 14.518 | 21.525 | 44.950 | 0.820 | -0.106 |
| 5Y swap spread | 4958.000 | 15.067 | 5.855 | -1.700 | 11.100 | 15.000 | 19.350 | 39.850 | 0.032 | -0.144 |
| 10Y swap spread | 4958.000 | 11.001 | 8.484 | -17.050 | 3.700 | 12.750 | 18.450 | 29.250 | -0.225 | -1.107 |
| 1Y cds | 4370.000 | 8.444 | 9.361 | 0.560 | 3.010 | 5.095 | 9.000 | 85.000 | 3.097 | 13.543 |
| 5Y cds | 4370.000 | 28.097 | 19.954 | 8.960 | 13.340 | 18.020 | 38.344 | 120.000 | 1.323 | 1.147 |
| 10Y cds | 4370.000 | 41.934 | 23.579 | 13.070 | 21.520 | 31.725 | 59.775 | 130.000 | 0.737 | -0.446 |
| Option Implied Vol 1M | 5718.000 | 9.630 | 2.655 | 4.795 | 7.925 | 9.205 | 11.179 | 27.388 | 1.122 | 2.717 |
| Option Implied Vol 3M | 5718.000 | 9.625 | 3.173 | 3.922 | 7.531 | 9.100 | 11.255 | 38.420 | 1.673 | 6.667 |

Table 12: Augmented Dickey-Fuller test for stationarity - GBP data

| Variable | p-value | p-value_diff |
|---|---|---|
| Fx spot | 0.642 | 0.000*** |
| Govt 3M | 0.741 | 0.000*** |
| Govt 1Y | 0.742 | 0.000*** |
| Govt 2Y | 0.767 | 0.000*** |
| Govt 5Y | 0.649 | 0.000*** |
| Govt 10Y | 0.611 | 0.000*** |
| Govt 20Y | 0.676 | 0.000*** |
| Central Bank Rate | 0.885 | 0.000*** |
| 1Y swap spread | 0.027** | 0.000*** |
| 2Y swap spread | 0.038** | 0.000*** |
| 5Y swap spread | 0.032** | 0.000*** |
| 10Y swap spread | 0.009*** | 0.000*** |
| 1Y cds | 0.000*** | 0.000*** |
| 5Y cds | 0.000*** | 0.000*** |
| 10Y cds | 0.017** | 0.000*** |
| Option Implied Vol 1M | 0.000*** | 0.000*** |
| Option Implied Vol 3M | 0.000*** | 0.000*** |

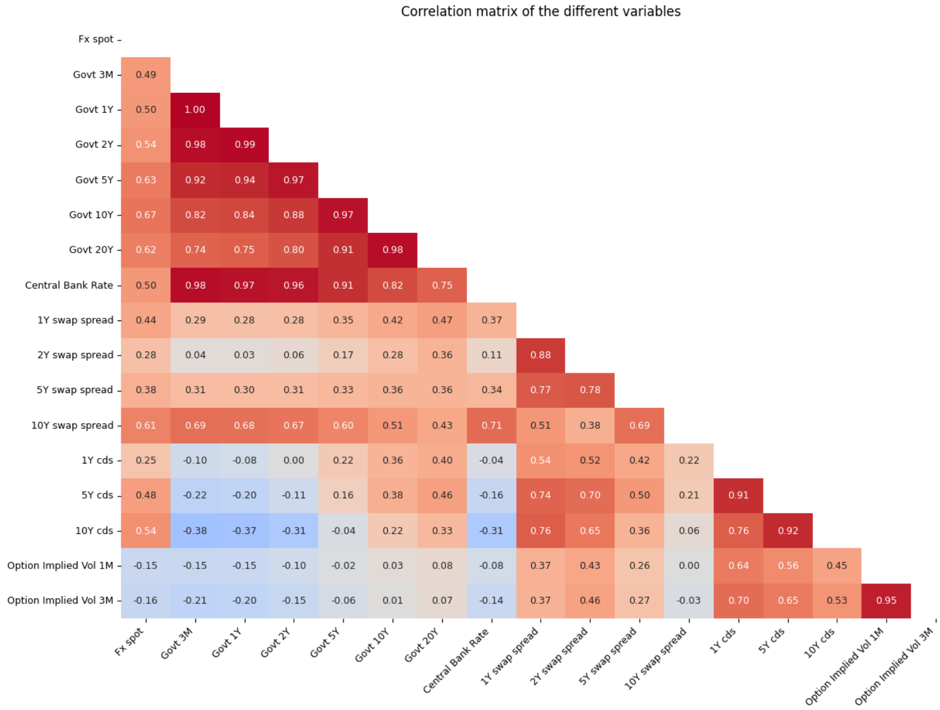
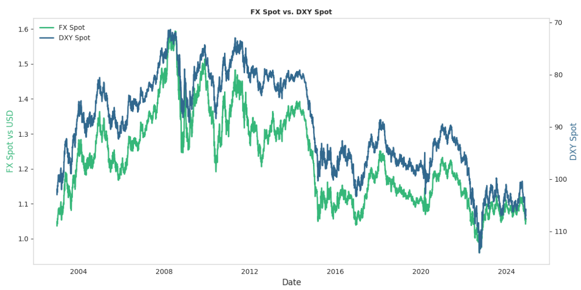
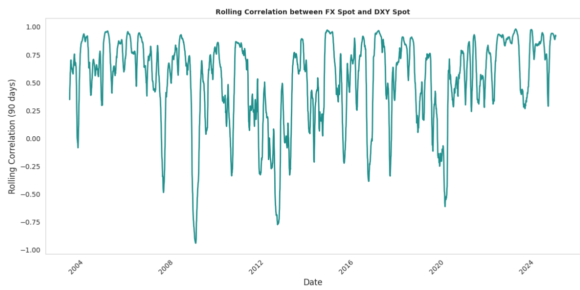



Figure 18: Correlation Matrix for GBP related variables

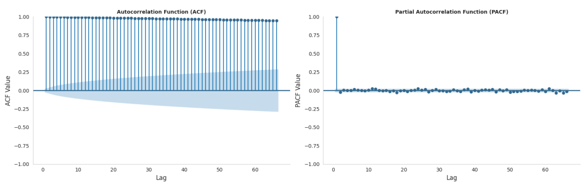



(a) Impact of dollar strength on EUR/USD pair

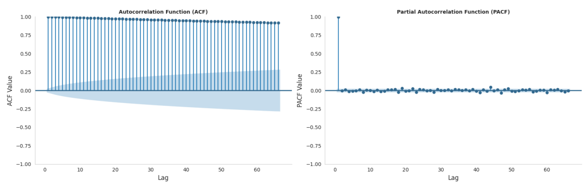



(b) Rolling correlation of DXY with USD/JPY pair

Figure 19

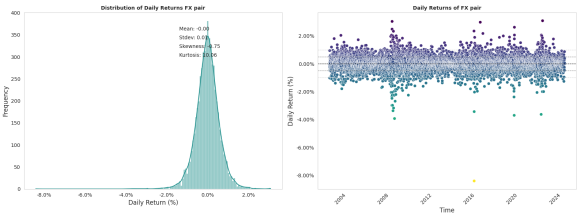



(a) ACF and PACF plots for GBP/USD spot fx rate

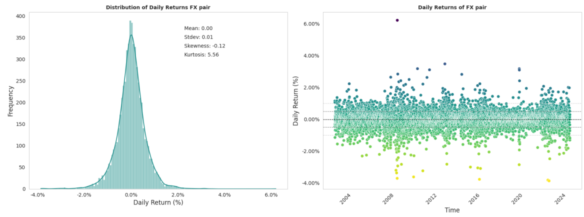



(b) ACF and PACF plots for EUR/USD spot fx rate

Figure 20: Autocorrelation tests

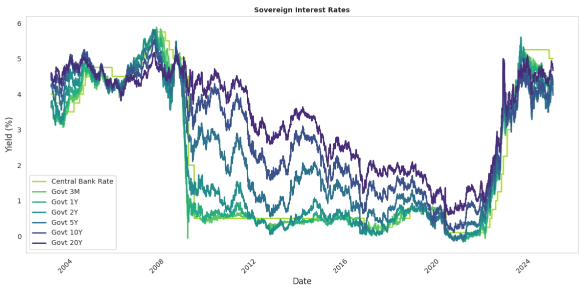



(a) Log returns for GBP/USD spot fx rate

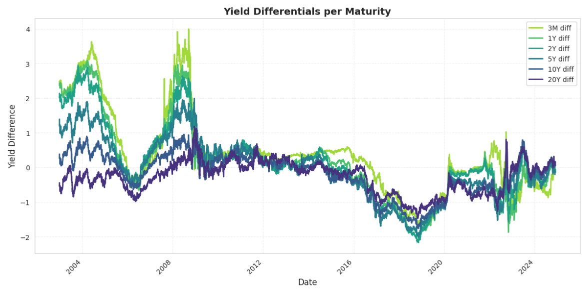



(b) Log returns for USD/JPY spot fx rate

Figure 21: Log returns for fx pairs

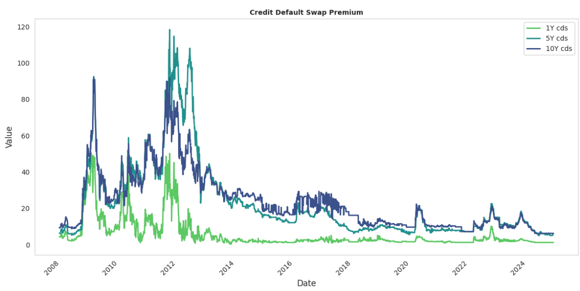



(a) Interest rates term structure (GBP)

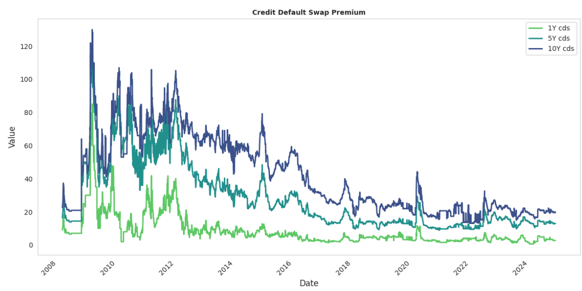



(b) Interest Rates Differentials (GBP minus USD)

Figure 22: Interest Rate term structures

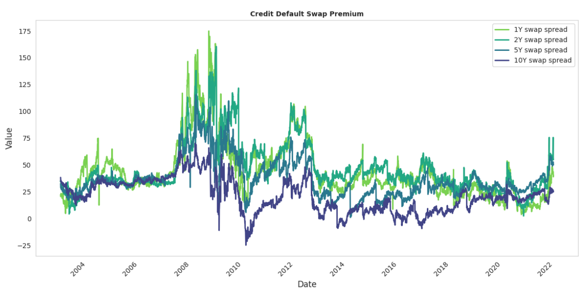



(a) CDS term structure for EUR Govt bonds

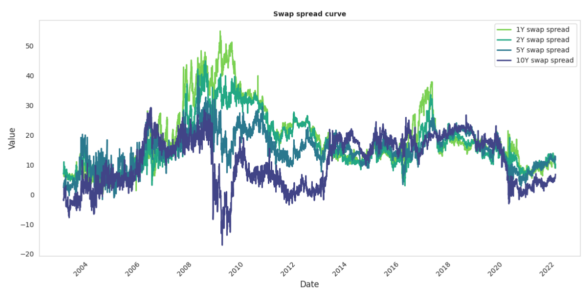



(b) CDS term structure for JPY Treasuries

Figure 23: CDS term structure for sovereign bonds




(a) Swap Spread structure for United Kingdom




(b) Swap Spread term structure for Japan

Figure 24: Spreads of swap rate over sovereign bonds of equal maturity

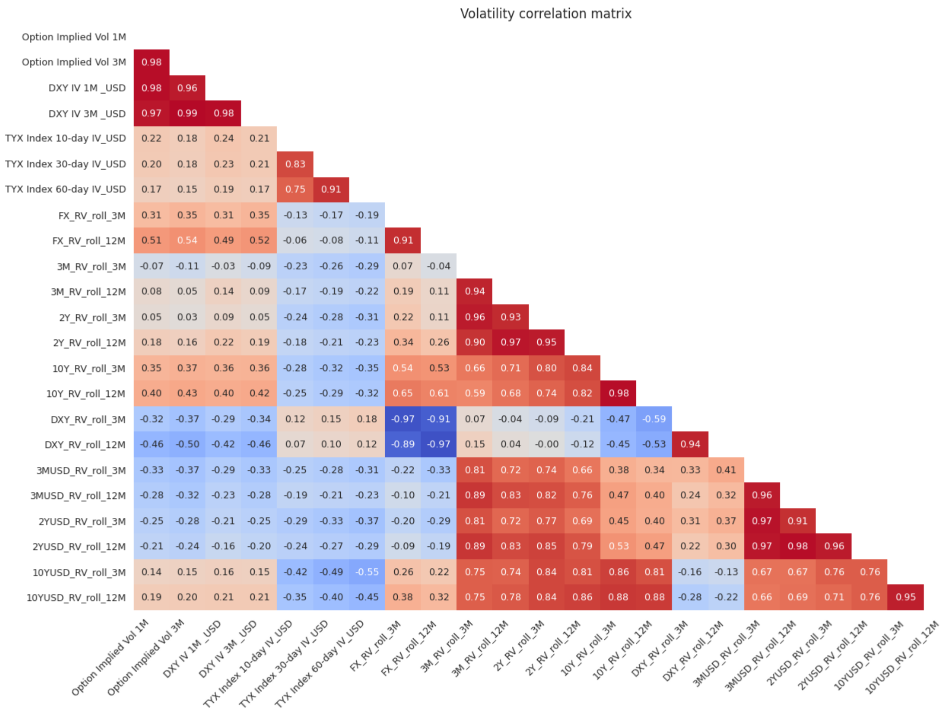



Figure 25: Correlation matrix for EUR and USD volatility variables

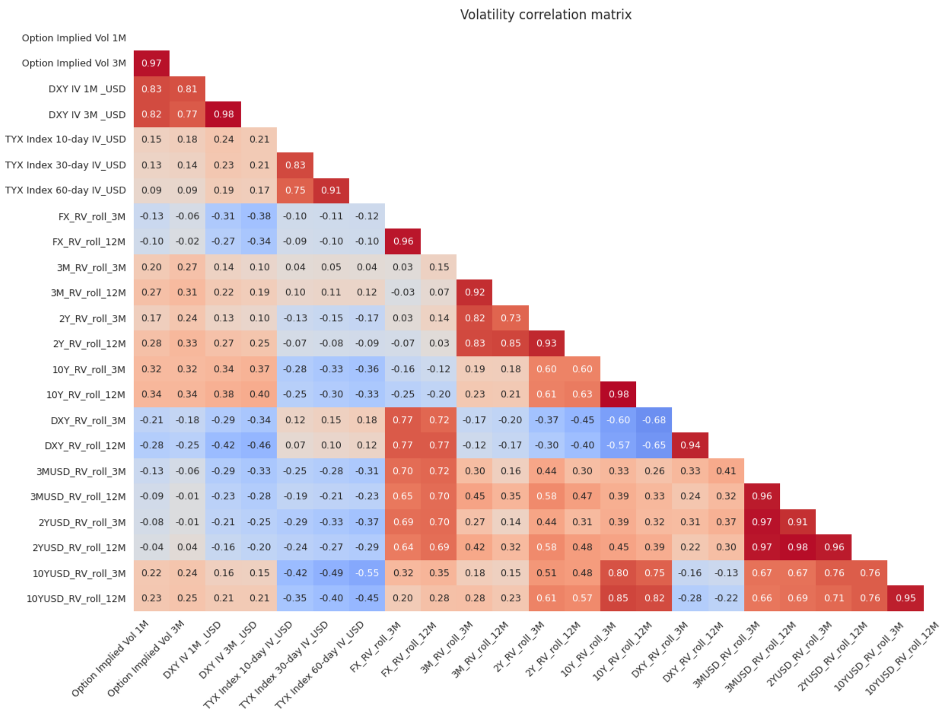



Figure 26: Correlation matrix for JPY and USD volatility variables

Table 13: Augmented Dickey-Fuller test for stationarity - JPY data

| Variable | p-value | p-value_diff |
|---|---|---|
| Fx spot | 0.837 | 0.000*** |
| Govt 3M | 0.682 | 0.000*** |
| Govt 1Y | 0.771 | 0.000*** |
| Govt 2Y | 0.666 | 0.000*** |
| Govt 5Y | 0.561 | 0.000*** |
| Govt 10Y | 0.648 | 0.000*** |
| Govt 20Y | 0.655 | 0.000*** |
| Central Bank Rate | 0.502 | 0.000*** |
| 1Y swap spread | 0.253 | 0.000*** |
| 2Y swap spread | 0.180 | 0.000*** |
| 5Y swap spread | 0.002*** | 0.000*** |
| 10Y swap spread | 0.052* | 0.000*** |
| 1Y cds | 0.001*** | 0.000*** |
| 5Y cds | 0.085* | 0.000*** |
| 10Y cds | 0.084* | 0.000*** |
| Option Implied Vol 1M | 0.005*** | 0.000*** |
| Option Implied Vol 3M | 0.000*** | 0.000*** |

Table 14: Augmented Dickey-Fuller test for stationarity - Global data

| Variable | p-value | p-value_diff |
|---|---|---|
| Gold | 0.963 | 0.000*** |
| Financial_CP | 0.498 | 0.000*** |
| vix | 0.000*** | 0.000*** |
| TEDRATE | 0.004*** | 0.000*** |
| EPU_JPY | 0.000*** | 0.000*** |
| EPU_EUR | 0.029** | 0.000*** |
| BBG_COM_index | 0.544 | 0.000*** |
| CRB_COM_index | 0.182 | 0.000*** |
| Fin_stress_OFR | 0.006*** | 0.000*** |
| Baltic_BDI | 0.031** | 0.000*** |
| EPU_GBP | 0.000*** | 0.000*** |
| EPU_USD | 0.000*** | 0.000*** |

Table 15: Augmented Dick Fuller test for stationarity - EUR data

| **Variable** | **p-value** | **p-value_diff** |
| --- | --- | --- |
| Fx spot | 0.236 | 0.000*** |
| Govt 3M | 0.644 | 0.000*** |
| Govt 1Y | 0.620 | 0.000*** |
| Govt 2Y | 0.612 | 0.000*** |
| Govt 5Y | 0.606 | 0.000*** |
| Govt 10Y | 0.536 | 0.000*** |
| Govt 20Y | 0.496 | 0.000*** |
| 1Y swap spread | 0.011** | 0.000*** |
| 5Y swap spread | 0.051* | 0.000*** |
| 10Y swap spread | 0.084* | 0.000*** |
| 1Y cds | 0.001*** | 0.000*** |
| 5Y cds | 0.081* | 0.000*** |
| 10Y cds | 0.083* | 0.000*** |
| Option Implied Vol 1M | 0.015** | 0.000*** |
| Option Implied Vol 3M | 0.0358** | 0.000*** |
| Central Bank Rate | 0.783 | 0.000*** |