Using Random Walks in Exchange Rate Modelling

Luis Kenji Hilasaca Sanchez

Supervised by Xiling Zhang

Submitted in accordance with the requirements for the
module MATH5872M: Dissertation in Data Science and Analitics
as part of the degree of

Master of Science in Data Science and Analytics

The University of Leeds, School of Mathematics

August 2025

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

Abstract

This dissertation confronts the long-standing challenge of outperforming a random walk in exchange rate forecasting. The Random Walk Hypothesis (RWH) is tested on daily data for three distinct currency pairs EUR/USD, ZAR/USD, and PEN/USD, using a rigorous, simulation-based out-of-sample evaluation framework. A suite of models, including standard time-series specifications (ARIMA, GARCH) and a machine learning approach (Efficient Kitchen Sink), are benchmarked against both a discrete and a continuous random walk. The primary contribution lies in the evaluation of a hybrid structural model that combines a traditional macroeconomic anchor, based on the framework of Meese & Rogoff (1983), with a sophisticated model for the dynamic, persistent, and heteroskedastic error process. The findings reveal a clear, horizon-dependent pattern of predictability. At short horizons (one month), no model consistently outperforms the random walk, supporting the weak-form efficiency of the market at high frequencies. At longer horizons (6 to 24 months), however, the hybrid structural models deliver substantial and robust forecast accuracy gains for EUR/USD and PEN/USD. In contrast, no model provides a useful signal for ZAR/USD, which remains consistent with a random walk across all horizons. This heterogeneity suggests that the "exchange rate disconnect" puzzle is not a universal phenomenon but rather a modelling challenge, where the key to unlocking long-term predictability lies in synergistic models that pair economic theory with an empirically-grounded treatment of error dynamics.

Summary

Title: Using Random Walks in Exchange Rate Modelling
Author: Luis Kenji Hilasaca Sanchez **Programme:** MSc Data Science and Analytics

Scope and Objective. This dissertation tests whether modern forecasting approaches can beat a random-walk benchmark in daily FX data. Three USD pairs are studied—EUR/USD, ZAR/USD, and PEN/USD—over 2015–2024. Competing models include: a discrete trinomial Random Walk (DRW), ARIMA(0,1,0) and ARIMA(1,1,0), GARCH(1,1), structural models with macro fundamentals and dynamic errors (AR/ARMA with GARCH), and an Efficient Kitchen Sink (E-KS) machine-learning model based on four standard predictors (UIP, PPP gap, Monetary Fundamentals gap, Taylor-rule composite).

Design and Data. Daily spot rates are evaluated out of sample (2023–2024) across horizons of 30, 182, 366 and 732 days. Monthly fundamentals (CPI, M1, IP, short rates) are mapped to daily frequency via log-linear interpolation (stocks) and forward-fill (rates). Forecast evaluation uses Monte Carlo simulation of $M = 1000$ paths per model/horizon, reporting expected RMSE and MAE in levels.

Main Findings.

- *Short horizon (1M):* The Random Walk remains exceptionally hard to beat. DRW is best (or statistically tied) for all pairs, confirming weak-form efficiency at high frequency.
- *Medium/Long horizons (6–24M):* Fundamentals re-emerge. For **EUR/USD**, a structural model with AR(1)+GARCH errors is consistently best from 6M onward ($\approx 30\%$ RMSE reduction vs. DRW at 24M). For **PEN/USD**, a simpler structural+AR(1) model dominates at 6–24M (over 50% RMSE reduction at 24M).
- **ZAR/USD** remains Random-Walk-like at all horizons; no specification outperforms DRW, pointing to drivers (commodity/risk factors) beyond a bilateral monetary model.
- **Machine Learning:** The E-KS (elastic-net) underperforms the RW and structural hybrids at all horizons, reflecting the low signal-to-noise ratio of daily fundamentals.

Contributions.

- A unified simulation-based evaluation (expected RMSE/MAE from 1000 paths) that fairly compares discrete RW, time-series, structural, and ML models on the same daily data.
- Evidence that *synergistic* models—macro anchors plus persistent/heteroskedastic error dynamics—unlock long-horizon predictability for some pairs.
- A clear demonstration of the heterogeneity of forecastability across different currency markets.

Limitations and Future Work. Results are pair-specific and frequency-sensitive: daily fundamentals are slow-moving and may mask signal. Fixed-window estimates ignore parameter drift. Future work should prioritize: (i) evaluating models at a monthly frequency where the signal from fundamentals is stronger, (ii) employing rolling-window estimations or time-varying parameter models to account for instability, and (iii) developing factor-augmented models to capture multilateral currency movements.

Key Sources. Meese & Rogoff (1983); Bollerslev (1986); Engel, Mark & West (2012); Li, Tsiakas & Wang (2015); Greenaway-McGrevy, Mark, Sul & Wu (2018).

Keywords: exchange rate forecasting; random walk hypothesis; ARIMA; GARCH; structural models; machine learning; out-of-sample forecasting; EUR/USD; ZAR/USD; PEN/USD.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

The foreign exchange (Forex) market stands as the world's most substantial financial arena, with trillions of US dollars traded daily, making its behaviour a subject of intense academic and practical interest (Cramon, 2024). The movements of exchange rates hold profound significance for a wide range of stakeholders. For multinational corporations and investors, unexpected currency fluctuations introduce significant risk, affecting the value of international assets, the cost of imported goods, and the profitability of exports. For central banks and policymakers, the exchange rate is a key transmission mechanism for monetary policy and a critical indicator of economic health. Consequently, the capacity to forecast these rates, even with a small degree of accuracy, offers considerable economic advantages (Engel et al., 2012). These advantages are not merely theoretical; they translate into more effective hedging strategies for firms, superior asset allocation decisions for portfolio managers, and better-informed policy interventions by authorities.

This dissertation is motivated by one of the most persistent and well-documented puzzles in international finance: the remarkable difficulty of outperforming a simple random walk (RW) model in "out-of-sample" exchange rate forecasting. This term refers to the process of using a model, trained on a historical set of data (the "training period"), to predict future exchange rate values that the model has not yet seen (the "out-of-sample" or "test period"). It is considered the gold standard for evaluating a model's true predictive power, as it prevents "in-sample" overfitting, where a model looks good on historical data but fails in real-world application.

The seminal work of Meese & Rogoff (1983) established this challenge, demonstrating that structural models based on macroeconomic fundamentals failed to beat a naïve "no-change" forecast. This naïve forecast, the simplest form of a random walk, posits that the best prediction for tomorrow's exchange rate is simply today's rate. The structural models tested by Meese and Rogoff were not purely deterministic; they were stochastic models that attempted to link the exchange rate to economic variables like money supply, inflation, and interest rates, but they consistently failed to produce more accurate out-of-sample forecasts than this elementary benchmark. This finding gave rise to the *exchange rate disconnect* puzzle, where currencies appear detached from their theoretical economic anchors.

However, subsequent research has nuanced this interpretation. The apparent random walk behaviour does not necessarily mean that exchange rates are devoid of economic structure. A key insight comes from the *present-value perspective*. If the fundamental drivers of the exchange rate (like interest rate differentials or inflation expectations) are themselves highly persistent, then the exchange rate, which incorporates market expectations of the entire future path of these fundamentals, will also exhibit highly persistent, random-walk-like behaviour. This is related to the concept of a *unit root process*. A time series with a unit root is non-stationary, meaning its statistical properties (like its mean and variance) change over time, and it has an infinitely long memory of past shocks. For example, the price of a stock that follows a random walk is a unit root process; a shock (e.g., a surprisingly good earnings report) will permanently alter its future trajectory. Adams & Chadha (1991) showed that, in finite samples, it can be statistically impossible to distinguish a highly persistent but ultimately stationary process from a true unit root process, cautioning against the conclusion that exchange rates lack any structural foundation.

Despite the formidable status of the RW benchmark, the academic pursuit of predictability has evolved into a sophisticated, multi-pronged effort. The research frontier has moved beyond simple linear models to explore several promising avenues:

1. *A more sophisticated use of fundamentals:* Rather than abandoning fundamentals, researchers now apply advanced econometric techniques to extract their predictive signal. This includes using models with time-varying parameters within a Bayesian framework to allow relationships to evolve, as demonstrated by Byrne et al. (2016) on monthly G10 currency data. While this approach can capture structural changes, its complexity can also make it prone to overfitting and its parameters difficult to interpret. Another successful strategy involves employing robust estimation methods like the *Generalized Method of Moments (GMM)*, which is particularly useful when standard OLS assumptions are violated. Chang et al. (2022) use GMM on quarterly data for G7 currencies to show that Taylor rule fundamentals can significantly outperform the random walk, though the predictive gains are modest and concentrated at specific horizons. Furthermore, specialized models like the copula-MIDAS approach are used to properly integrate the mixed-frequency nature of macroeconomic (monthly/quarterly) and financial (daily) data. Gong & Lin (2022) apply this to the USD/CNY exchange rate, finding that fundamentals do have a significant effect, but the complexity of the model makes it less transparent than simpler alternatives.

2. *The introduction of non-linearity and regime changes:* A growing body of work argues that the relationship between exchange rates and their determinants is not linear. Models that explicitly allow for different behaviours in different market states (e.g., high vs. low volatility), such as Threshold Autoregressive (TAR) or Markov-switching models, have shown an ability to outperform the random walk. For instance, Çatık & Karaca (2020) find

success applying these models to monthly data for a panel of emerging market currencies. The main advantage is their ability to capture sudden structural breaks. However, these models carry the significant disadvantage of requiring the pre-specification of the number of regimes and the threshold variable, which can be subjective and may not be stable out-of-sample.

3. *The application of Machine Learning and AI:* More recently, a new wave of research has applied powerful, data-driven machine learning algorithms to the problem. These methods are agnostic to economic theory and excel at identifying complex, non-linear patterns. Studies have shown success with a wide range of techniques, from tree-based methods like Random Forests to advanced deep learning models. Colombo & Manera (2020) test a battery of ML models on monthly G10 currencies, finding that they often outperform the random walk, but also noting that their performance is sensitive to hyperparameter tuning and they suffer from a lack of interpretability (the "black box" problem). In the deep learning domain, *Long Short-Term Memory (LSTM)* networks, a type of Recurrent Neural Network (RNN) specifically designed for sequential data, have been applied. Hong et al. (2023) use LSTMs on daily data for several major currency pairs and find superior performance. The key advantage of LSTMs is their ability to capture long-range dependencies, but their major drawbacks include high computational cost, a large number of parameters that increases the risk of overfitting, and even greater difficulty in interpretation compared to simpler ML models.

4. *A shift to a multilateral perspective:* Another key insight is that bilateral exchange rates are not determined in a vacuum. Research on factor models has shown that a large portion of currency movements can be explained by a small number of common factors, such as a general "dollar factor" (Engel et al., 2012). Identifying these systemic drivers, like a "dollar" and "euro" factor, suggests that a multilateral view is crucial for forecasting (Greenaway-McGrevy et al., 2018).

This study connects directly to this ongoing debate. Its primary objective is to conduct a rigorous, comparative forecast evaluation on high-frequency daily data, testing several of these modelling philosophies against each other. A sophisticated bilateral model, which combines a structural macroeconomic specification with a flexible AR-GARCH error structure, is tested against not only the classic random walk but also an adaptation of the modern E-KS approach from Li et al. (2015). By doing so, light is shed on which modelling strategies hold the most promise in the challenging environment of daily exchange rate forecasting.

1.2 Problem Statement and Research Questions

The central issue this thesis confronts is the empirical strength of the Random Walk Hypothesis (RWH) when applied to modern, high-frequency exchange rate data. While the literature has explored numerous avenues for predictability, a consensus remains elusive, particularly for daily data where the noise-to-signal ratio is notoriously high. This study aims to provide a clear, comparative assessment of several key modelling paradigms within a unified forecasting framework.

The primary research question is therefore articulated as: *Do key currency exchange rates, representing diverse market structures, behave as a random walk, or can their future movements be predicted with greater accuracy by more sophisticated forecasting models when evaluated on daily out-of-sample data?*

This overarching question gives rise to a set of specific, operational sub-questions that will guide the empirical analysis:

1. In an out-of-sample context, do standard time series models (ARIMA, GARCH) offer a statistically significant improvement in forecast accuracy over a simple Random Walk benchmark for the EUR/USD, ZAR/USD, and PEN/USD pairs across horizons of 30, 182, 366, and 732 days?

2. Can a structural model, which incorporates macroeconomic fundamentals, outperform the Random Walk? More specifically, does augmenting this structural model with a dynamic AR(1)-GARCH error structure lead to a meaningful reduction in forecast error?

3. Does a modern, machine-learning-based approach like the Efficient Kitchen Sink, adapted to a daily frequency, offer predictive gains, or does its performance suffer from the high noise-to-signal ratio of daily data, failing to beat even the simplest benchmarks?

The performance of each model will be rigorously evaluated based on its ability to minimize forecast error outside of the training period. The primary evaluation metrics will be the *expected Root Mean Squared Error (RMSE)* and the *expected Mean Absolute Error (MAE)*, which will be computed through a comprehensive Monte Carlo simulation framework. This simulation-based approach allows for the assessment of not just a single forecast path, but the entire distribution of potential outcomes, providing a more robust measure of a model's predictive power.

1.3 Objectives and Scope

The principal objective of this dissertation is to conduct a rigorous, systematic, and comparative test of the Random Walk Hypothesis. A key secondary objective is to explore the horizon-dependency of any predictability found by assessing each model's performance across multiple forecast horizons, spanning from one month to two years. Furthermore, the same set of models is applied to three distinct currency pairs in order to evaluate the robustness of the findings across varying market structures and economic environments.

The scope of this study is carefully defined to ensure a focused and manageable analysis. We concentrate exclusively on forecasting three bilateral exchange rates against the U.S. dollar: the EUR/USD, the ZAR/USD, and the PEN/USD. This selection was deliberate to capture a diverse range of market dynamics. The EUR/USD is the world's most traded currency pair; the ZAR/USD (South African Rand) represents a highly volatile emerging market currency often influenced by global risk sentiment and commodity prices; and the PEN/USD (Peruvian Sol) provides a case from a different emerging market with potentially different characteristics. The analysis utilizes a 10-year sample of daily closing prices from 2015 to 2024.

The study's scope is also defined by what it excludes. We do not venture into ultra-high-frequency intraday data, which would require an analysis of market microstructure. Similarly, the modelling toolkit, while comprehensive, does not include the most computationally intensive deep learning models, such as LSTMs or transformers, as the focus is on the intersection between traditional econometrics and established machine learning techniques.

1.4 Thesis Structure

This dissertation is structured across seven chapters, each designed to build logically upon the last, from the establishment of the research problem to the final conclusions.

Chapter 2, the *Literature Review*, delves into the key academic debates that form the foundation of this study. It begins by dissecting the seminal work on the Random Walk Hypothesis and the "exchange rate disconnect" puzzle. It then explores the evolution of forecasting approaches, reviewing the literature on models based on macroeconomic fundamentals, the introduction of non-linearity and time-varying parameters, the application of modern machine learning techniques, and the important shift towards multilateral factor models.

Chapter 3, the *Data Description*, provides a thorough description of the data used in the empirical analysis. It details the sources, transformations, and construction of all exchange rate and macroeconomic variables.

Chapter 4, the *Methodology*, provides a rigorous technical description of each forecasting model employed, from the simple Random Walk benchmarks to the more complex ARIMA, GARCH, E-KS, and hybrid structural specifications. This chapter also details the robust model evaluation framework based on Monte Carlo simulations.

Chapter 5 presents the *Empirical Results*. It is structured to first present the in-sample estimation results for each model, followed by the main out-of-sample forecasting performance, allowing for a direct comparison of their predictive power.

Chapter 6, the *General Discussion and Conclusions*, moves beyond a simple reporting of the findings to interpret their broader implications. It analyzes why certain models performed better than others, connects the results to the literature, discusses limitations, and provides a final, evidence-based answer to the research questions.

Chapter 2

Literature Review

This chapter reviews the theoretical and empirical literature relevant to exchange rate modelling and forecasting, focusing on the evolution of approaches designed to challenge the random walk benchmark.

2.1 The Random Walk Benchmark and its Interpretation

The Random Walk Hypothesis (RWH) has been the central and most formidable benchmark in exchange rate forecasting for over four decades. A simple random walk model, which corresponds to an ARIMA(0,1,0) process without drift, posits that the change in the exchange rate from one period to the next is a purely random innovation, independent of all past information. Formally, for the log exchange rate X_t, this is expressed as $X_t = X_{t-1} + \varepsilon_t$, where ε_t is a white-noise error term. The implication is stark: the best forecast for tomorrow's exchange rate is simply today's exchange rate.

The dominance of this benchmark was cemented by the seminal work of Meese & Rogoff (1983), who found that prevailing structural macroeconomic models could not outperform it in out-of-sample tests. This result has proven remarkably robust over time, leading to what is often called the "Meese and Rogoff puzzle" (Grossmann & Chudik, 2015). The reason the RW is such a difficult benchmark to beat is rooted in the efficient market hypothesis, which suggests that all predictable information is already incorporated into the current price. Furthermore, from a present-value perspective, if the fundamental drivers of exchange rates are themselves highly persistent (near unit-root processes), and if the discount factor applied to future fundamentals is close to one, then the exchange rate will theoretically behave very much like a random walk, even if it is fundamentally determined (Engel et al., 2012).

Evaluating performance against this benchmark requires a rigorous out-of-sample framework. The most common metric is the Root Mean Squared Prediction Error (RMSPE), often compared using Theil's U statistic, where a value less than one indicates that the alternative model is more accurate than the random walk. However, simply comparing RMSPE values is insufficient to claim superiority. Formal statistical tests, such as the Diebold-Mariano test or, more appropriately for nested models like those tested against a random walk, the Clark-West test, are necessary to determine if the observed improvement in accuracy is statistically significant and not due to chance (Engel et al., 2012). The practical implication of the RW's dominance is that for short horizons (from intraday to several weeks), predictability is exceptionally low. However, the literature has found that as the forecast horizon extends to several months or years, the signal from fundamentals or other factors can begin to emerge, and some models have shown an ability to modestly, but significantly, outperform the random walk (Greenaway-McGrevy et al., 2018).

2.2 Predictability from Macroeconomic Fundamentals

Despite the historical failures, the notion that economic fundamentals should drive exchange rates is too theoretically compelling to abandon. The literature has focused on several key channels through which fundamentals are expected to influence currency values.

2.2.1 Key Fundamental Theories

- **Purchasing Power Parity (PPP):** This is one of the oldest theories, suggesting that the exchange rate between two currencies should equalize the prices of a common basket of goods. Empirically, this is often tested by modelling the real exchange rate (the nominal rate adjusted for price level differences). While PPP fails spectacularly in the short run, there is strong evidence that deviations from PPP are mean-reverting over long horizons (typically with a half-life of 3-5 years), providing a long-term anchor for the exchange rate.

- **Uncovered Interest Parity (UIP):** The UIP condition states that the expected change in the exchange rate should equal the interest rate differential between two countries. If UIP holds, the forward exchange rate should be an unbiased predictor of the future spot rate. However, empirically, the forward premium has been found to be a biased predictor, often pointing in the wrong direction. This ”forward premium puzzle” has led researchers to use the forward premium or interest rate differentials not as a direct forecast, but as a predictor variable in more complex models, where it may capture risk premia (Li et al., 2015).

- **Monetary Models:** These models, such as the flexible-price or sticky-price (Dornbusch) variants, link the exchange rate to differentials in money supply, output, and interest rates. For example, a relative increase in the domestic money supply is expected to lead to domestic currency depreciation. While these models formed the basis of the failed tests in Meese & Rogoff (1983), they remain a cornerstone of theoretical exchange rate determination.

- **The Taylor Rule:** A more modern approach is to link exchange rates to the systematic way central banks conduct monetary policy. The Taylor rule posits that central banks set short-term interest rates in response to deviations of inflation from its target and output from its potential. By modelling the interest rate differential as a function of inflation and output gap differentials, researchers have found a more robust link to exchange rate movements. Studies show that Taylor rule fundamentals, when modelled with appropriate econometric techniques like time-varying parameters or robust estimators, can have significant predictive power (Byrne et al., 2016; Chang et al., 2022).

2.2.2 Modern Econometric Practice

The mixed results of fundamental models have led to a consensus on best practices for implementation. The short-term "disconnect" versus a potential long-term signal suggests that the relationship is not stable. Consequently, modern approaches favour using **rolling estimation windows** to allow model parameters to adapt to changing market conditions. It is also critical to avoid look-ahead bias by ensuring that all data transformations (like smoothing or filtering) only use information available at the time the forecast is made. Finally, the success of the "kitchen-sink" approach with shrinkage estimators like the elastic net (Li et al., 2015) highlights the benefits of model combination and regularization as a way to manage parameter uncertainty when the true set of relevant fundamentals is unknown.

2.3 Non-linearity and Machine Learning Models

A major modern research avenue posits that the relationship between exchange rates and their determinants is inherently non-linear, meaning linear models are fundamentally misspecified.

2.3.1 Sources of Non-Linearity

Several economic phenomena can induce non-linear behaviour in exchange rates. These include asymmetries, where the exchange rate reacts differently to positive versus negative shocks, and threshold effects, such as transaction costs that create a band of inaction where PPP deviations are not arbitraged away. More broadly, the economy may operate in different regimes (e.g., a "calm" versus a "turbulent" state), with the relationships between variables changing across these states. Models like TAR and Markov-switching are designed to capture these dynamics (Çatık & Karaca, 2020). Furthermore, the phenomenon of conditional volatility, best captured by GARCH models, is itself a form of non-linearity, as it implies that the variance of returns depends on the magnitude of past returns.

2.3.2 Machine Learning as a Toolkit

Machine learning (ML) should be viewed not as a single model, but as a powerful toolkit for tackling these issues, particularly in a high-dimensional setting. Its primary advantages in this context are:

- **Regularization:** Techniques like LASSO, Ridge, and the elastic net are fundamentally ML methods. They provide a disciplined way to handle the "kitchen-sink" problem of having too many potential predictors. By shrinking the coefficients of irrelevant variables, they reduce the variance of the forecast and prevent the overfitting that plagues standard OLS, thereby allowing weak but stable signals to be detected (Li et al., 2015).

- **Non-Linear Pattern Recognition:** Tree-based models (like Random Forests) and neural networks are universal approximators capable of learning highly complex, non-linear relationships directly from the data without requiring a pre-specified economic theory. This is particularly useful when the true functional form linking fundamentals to exchange rates is unknown.

- **Fair Model Comparison:** The ML literature places a heavy emphasis on rigorous out-of-sample validation. When comparing a flexible ML model against a simple benchmark like the random walk, it is crucial to use nested-model-aware tests like the Clark-West test to avoid "nested-model bias," ensuring a fair comparison.

While powerful, ML models are not a panacea. They often suffer from a lack of interpretability (the "black box" problem), can be computationally expensive, and are prone to finding spurious patterns if not carefully validated (Colombo & Manera, 2020; Hong et al., 2023). Their success depends on a disciplined application that respects the principles of out-of-sample evaluation.

2.4 Factor Models and Common Currency Movements

A final significant departure from traditional bilateral models has been the development of factor models that seek to explain exchange rate movements through a small number of unobserved common factors. Engel et al. (2012) show that a factor model can produce forecasts that are consistently more accurate than the random walk, particularly at longer horizons of 12 to 24 months. By extracting the first principal component from a panel of dollar exchange rates, they create a "dollar factor" that captures the broad strengthening or weakening of the USD.

Building on this, Greenaway-McGrevy et al. (2018) use more advanced identification procedures to show that exchange rate returns are driven by a robust two-factor structure: a "dollar factor" and a "euro factor". Their key finding is that a multilateral model built upon these factors consistently outperforms both the random walk and models based on traditional fundamentals in out-of-sample forecasting. This provides a strong argument that understanding any single bilateral rate requires accounting for these broad, systemic currency movements, a perspective that purely bilateral models inherently miss.

Chapter 3

Data Description

This chapter details the datasets, transformations, and evaluation design used in the analysis. A clear understanding of the data—its sources, frequency, and the construction of predictive variables—is crucial before proceeding to the modelling methodologies in the next chapter.

3.1 Scope, Sources, and Raw Data

This study uses daily nominal exchange rates for three USD-based currency pairs and maps monthly macroeconomic indicators to a daily frequency.

Currency Pairs. This study focuses on three distinct currency pairs against the U.S. Dollar: the Euro from the Eurozone (EUR/USD), the South African Rand (ZAR/USD), and the Peruvian Sol (PEN/USD). This selection was deliberate to capture a diverse range of market dynamics. The EUR/USD represents the world's most traded and liquid currency pair. The ZAR/USD serves as an example of a highly volatile emerging market currency often influenced by commodity prices and global risk sentiment. Finally, the PEN/USD provides a case from a different Latin American emerging market, allowing for an assessment of the models' robustness across varied economic environments.

For analytical consistency, all exchange rates are standardized as U.S. dollars per unit of foreign currency. Daily closing prices are sourced from the Federal Reserve Economic Data (FRED) for EUR/USD and ZAR/USD, and from the Central Reserve Bank of Peru (BCRP) for PEN/USD. A visual representation of the raw nominal exchange rate levels for these pairs is provided in Figure 3.1, illustrating their distinct trends and volatility patterns over the sample period.

Sample Window. All time series are aligned to a common daily calendar spanning from January 1, 2015, to December 31, 2024. As shown in Table 3.1, a fixed calendar cutoff is used for model evaluation, reserving the period after December 30, 2022, for out-of-sample testing.

Table 3.1: Exchange-rate pairs and sample windows

Pair i	Data Source	Full sample	Train/Test split
EUR/USD	FRED	2015–2024	Train: $t \leq$ 2022-12-30
ZAR/USD	FRED	2015–2024	Train: $t \leq$ 2022-12-30
PEN/USD	BCRP	2015–2024	Train: $t \leq$ 2022-12-30

Macroeconomic Predictors. The selection of fundamental variables is grounded in the canonical theories of exchange rate determination, forming the building blocks of models tested since the seminal work of Meese & Rogoff (1983) and employed in modern forecasting frameworks like that of Li et al. (2015). Throughout the thesis, variables corresponding to the foreign country (Eurozone, South Africa, or Peru) are denoted with an asterisk (*), while variables for the domestic country (United States) have no asterisk. The variables, sourced at a monthly frequency, include:

- **Consumer Price Index (CPI):** This measures the general price level and is the cornerstone of **Purchasing Power Parity (PPP)** theory. We denote the log of the CPI as p_t for the U.S. and p_t^* for the foreign country. These are used to construct the real exchange rate.

- **Year-over-Year Inflation Rate:** Derived from the CPI levels, the inflation rate is a key component of Taylor-rule based predictors. It is denoted by π_t and π_t^*.

- **M1 Monetary Aggregates and Industrial Production (IP):** These variables are central to **monetary models** of the exchange rate. The log of the M1 money supply is denoted by m_t (m_t^*), and the log of the seasonally adjusted IP index (our proxy for real activity) is denoted by y_t (y_t^*).

- **3-Month Money Market Interest Rates:** These rates reflect the stance of monetary policy and are the core component of the **Uncovered Interest Parity (UIP)** condition. They are denoted by i_t and i_t^*.

Together, these variables provide a parsimonious yet theoretically comprehensive set of fundamentals, allowing us to test the key channels through which macroeconomics is expected to influence currency values. Data for these indicators are sourced for the United States, the Eurozone, South Africa, and Peru.

(a) EUR/USD

(b) ZAR/USD

(c) PEN/USD

Figure 3.1: Nominal exchange rate levels (S_t) for the three currency pairs (2015–2024).

3.2 Data Transformation and Predictor Construction

A series of transformations, commonly adopted in the literature and designed to ensure transparency and reproducibility, is applied to the data in preparation for modelling.

3.2.1 Exchange Rate Transformations

Let S_t denote the nominal spot exchange rate. We work with the log-transformed series, $X_t = \log(S_t)$, as this stabilizes the variance and allows log-differences to be interpreted as continuously compounded returns. Daily returns are thus defined as $\Delta X_t \equiv X_t - X_{t-1}$. Our benchmark for forecast evaluation is the driftless random walk (RW), which posits that the expected future return is zero, $\mathbb{E}_t[\Delta X_t] = 0$. This has been the standard yardstick for out-of-sample performance since the seminal work of Meese & Rogoff (1983).

3.2.2 Mapping Monthly Macro-Data to a Daily Frequency

A key challenge is aligning the low-frequency macroeconomic data with the high-frequency exchange rate data. We employ two distinct methods based on the nature of the variable to ensure the daily series are constructed appropriately.

Log-Linear Interpolation for Stock Variables. For stock-like variables such as the Consumer Price Index (CPI), M1 monetary aggregates, and seasonally adjusted Industrial Production (IP), which grow multiplicatively, interpolated in log space. This assumes a constant daily growth rate within each month, avoiding the artificial "staircase" effect of a simple forward-fill and better reflecting the compounding nature of these series.

Forward-Fill for Flow and Rate Variables. For variables that are typically released as a monthly statistic, such as year-over-year inflation rates and 3-month interest rates, a forward-fill (or piecewise constant) method is used. This ensures that the value for any given day is the latest available monthly figure, which better reflects the discrete arrival of new information to the market.

3.2.3 Seasonal Adjustment Policy

Our approach to seasonality is deliberate and specific to the role of each macroeconomic variable.

Industrial Production (IP). We use seasonally adjusted (SA) IP indices for all countries. IP exhibits strong seasonal patterns, and failing to adjust would inject mechanical noise into the cross-country differentials ($y_t - y_t^*$) used in the structural models. Using SA series for all countries isolates the signal related to real economic activity, a standard practice in the forecasting literature.

CPI and M1. In contrast, the official not seasonally adjusted (NSA) indices for CPI and M1 are utilised. These variables primarily enter the models as log-level differences between countries (e.g., $p_t^* - p_t$). In such specifications, any shared seasonal patterns tend to cancel out.

3.2.4 Construction of Structural Predictors

We follow the "Efficient Kitchen Sink" (E-KS) specification of Li et al. (2015) to derive four structural predictors from the daily-mapped fundamentals. Throughout, p_t, m_t, y_t denote the logs of CPI, M1, and seasonally adjusted Industrial Production (IP), respectively.

(1) UIP / Forward Premium. We proxy the forward premium via the short-rate differential. For comparability across different forecast horizons, this differential is scaled:

$$x_t^{\text{UIP}} \equiv \kappa_h \left(i_t - i_t^*\right),$$

where i_t is the 3-month interest rate and κ_h is a scaling factor (where $\kappa_h = 1/360$ for a one–day ahead simulations). If forwards are observed, the equivalent definition is $x_t^{\text{UIP}} = f_t - X_t$. No log transformation is applied to interest rates.

(2) PPP Gap (Real Exchange-Rate Deviation). Deviations from Purchasing Power Parity are measured using the standard CPI-based real exchange rate definition (in logs):

$$x_t^{\text{PPP}} \equiv 100 \left(X_t + p_t^* - p_t\right).$$

The variable is scaled by 100 to be interpreted in percentage points.

(3) Monetary Fundamentals (MF) Gap. Following the log-linear monetary model, a predictor that captures money-demand-adjusted money supply differentials:

$$x_t^{\text{MF}} \equiv 100\Big[(m_t - m_t^*) - (y_t - y_t^*) - X_t\Big].$$

(4) Taylor-Rule Composite. The Taylor-rule block aggregates information on inflation, real activity, and the real exchange rate into a single composite predictor. Construction of this measure begins with the definition of the year-over-year inflation rate, $\pi_t = 100 \times (p_t - p_{t-252})$, and the output gap, y_t^g, derived from HP-filtered monthly log IP. The composite is then:

$$x_t^{\text{Taylor}} \equiv 1.5\,(\pi_t - \pi_t^*) + 0.1\,(y_t^g - y_t^{g*}) + 0.1\,x_t^{\text{PPP}}.$$

The coefficients (1.5, 0.1, 0.1) are standard calibrations from the literature. All four predictors, $\{x_t^{\text{UIP}},\, x_t^{\text{PPP}},\, x_t^{\text{MF}},\, x_t^{\text{Taylor}}\}$, are standardized using training-sample moments before being used in the models.

Chapter 4

Methodology

This chapter outlines the econometric models used to forecast exchange rates. Based on the data presented in Chapter 3, this section details the preprocessing steps undertaken to obtain the stationary series required for the modelling framework. We then define the benchmark models and introduce the more sophisticated alternatives against which they will be evaluated.

4.1 Preprocessing and Stationarity Diagnostics

The first step of the methodology is to transform the raw exchange rate data into a stationary series suitable for modelling.

Log-Transformation and Differencing. As is standard practice, the log-transformed series is employed, $X_t = \log(S_t)$. To analyse returns and achieve stationarity, the first difference of this series is taken, yielding:

$$\Delta X_t \equiv X_t - X_{t-1}. \tag{4.1}$$

Drift and Empirical Innovations. To isolate the zero-mean component of the returns, the low-frequency drift is estimated as the sample mean of the first differences over the training window:

$$\widehat{\mu} \equiv \frac{1}{|\mathcal{T}_{\text{train}}|} \sum_{t \in \mathcal{T}_{\text{train}}} \Delta X_t. \tag{4.2}$$

Subtracting this estimated drift from the returns series gives us the series of empirical innovations, which serves as the workhorse for the subsequent analysis:

$$\varepsilon_t \equiv \Delta X_t - \widehat{\mu}, \qquad t \in \mathcal{T}_{\text{train}}. \tag{4.3}$$

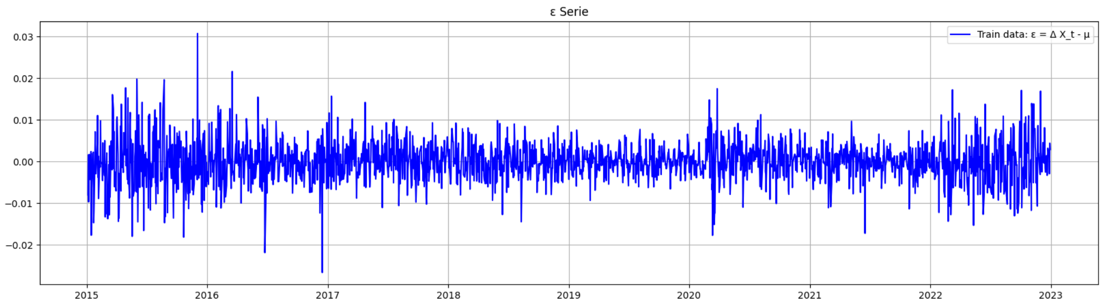
Stationarity Diagnostics. We formally test whether the resulting series $\{\varepsilon_t\}$ is weakly stationary using the Augmented Dickey-Fuller (ADF) test. As shown in Table 4.1, the null hypothesis of a unit root is strongly rejected for all three currency pairs, confirming that the empirical innovations are stationary and suitable for modelling. The descriptive statistics are presented in Table 4.2.

Table 4.1: ADF tests and estimated drift on $\varepsilon_t^{(i)}$ (training window)

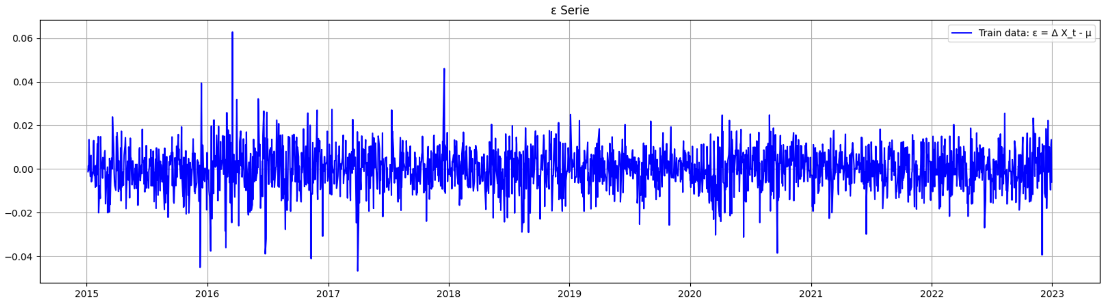
Pair i	$\hat{\mu}$ (Drift)	ADF statistic	p-value	Decision
EUR/USD	-0.000052	-43.72	$< 10^{-5}$	Reject H_0 (stationary)
ZAR/USD	-0.000179	-44.01	$< 10^{-5}$	Reject H_0 (stationary)
PEN/USD	-0.000119	-30.76	$< 10^{-5}$	Reject H_0 (stationary)

Table 4.2: Descriptive statistics for $\varepsilon_t^{(i)}$ on $\mathcal{T}_{\text{train}}$

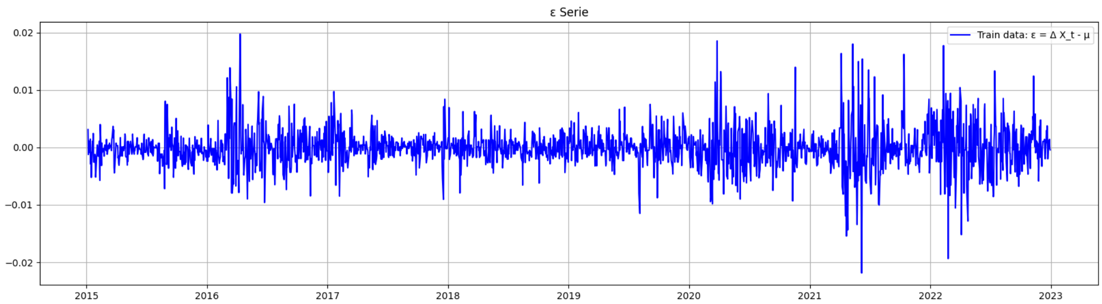
Pair i	Count	Mean	Std	Min	25%	50%	75% / Max
EUR/USD	2085	≈ 0	0.0049	-0.0267	-0.0026	-0.0000	0.0027 / 0.0307
ZAR/USD	2085	≈ 0	0.0099	-0.0467	-0.0059	0.0004	0.0063 / 0.0628
PEN/USD	2084	≈ 0	0.0035	-0.0219	-0.0017	-0.0002	0.0014 / 0.0197



(a) EUR/USD



(b) ZAR/USD



(c) PEN/USD

Figure 4.1: Empirical innovations (demeaned log-returns) $\varepsilon_t^{(i)}$ for the three currency pairs (Training Sample).

4.2 The Benchmark Model: Random Walk (Discrete, Trinomial)

Our primary benchmark is a discrete random walk model with a trinomial innovation structure. This approach is chosen over the classic Gaussian random walk to better capture the empirical properties of daily exchange rate returns, particularly the high frequency of very small price movements. The model for the log exchange rate X_t is based on the standard random walk process with drift:

$$\Delta X_t \equiv X_t - X_{t-1} = \mu + \varepsilon_t, \qquad X_t = X_0 + \mu t + \sum_{j=1}^{t} \varepsilon_j. \tag{4.4}$$

Here, μ is the drift parameter, estimated as the mean of the one-step changes ΔX_t on the training sample, and ε_t is the innovation term.

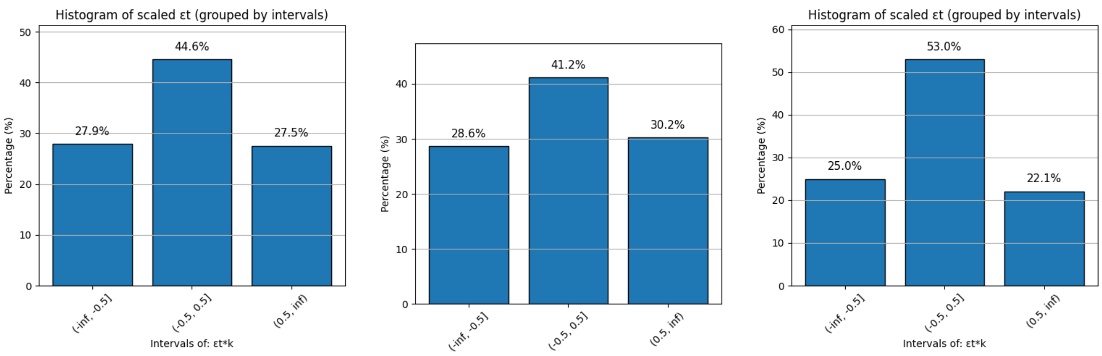
The core of the model lies in the specification of this innovation. Instead of assuming ε_t follows a continuous distribution, its structure is derived directly from the data. Let $\varepsilon_t \equiv \Delta X_t - \mu$ be the demeaned one-step changes calculated on the training data. We first scale these empirical innovations by a factor k (e.g., $k = 200$) to get $\xi_t = k\varepsilon_t$. This scaling magnifies the values, making it easier to group them. We then discretise ξ_t into a trinomial variable W_t with three possible outcomes, based on fixed cutpoints at -0.5 and 0.5:

$$W_t = \begin{cases} -1, & \text{if } \xi_t < -0.5, \\ 0, & \text{if } -0.5 \le \xi_t \le 0.5, \\ +1, & \text{if } \xi_t > 0.5. \end{cases}$$

The choice of cutpoints ensures that any scaled return within $[-0.5, 0.5]$ is treated as a "no significant change" event ($W_t = 0$). The empirical probabilities for each outcome are then calculated as the frequency of occurrence within the training window:

$$\widehat{p}_- = \mathbb{P}(W_t = -1), \quad \widehat{p}_0 = \mathbb{P}(W_t = 0), \quad \widehat{p}_+ = \mathbb{P}(W_t = +1),$$

such that their sum is unity. For instance, for the EUR/USD pair with a scale of $k = 220$, the purely empirical frequencies are adjusted slightly to achieve symmetry, guided by the visual evidence from the histograms. This pragmatic step yields final probabilities of $\widehat{p}_- = 0.277$, $\widehat{p}_0 = 0.446$, and $\widehat{p}_+ = 0.277$. The high probability of the zero-state is intended to capture the observed inertia in daily returns, a feature visually justified by the histograms in Figure 4.2.



(a) EUR/USD (b) ZAR/USD (c) PEN/USD

Figure 4.2: Histograms of scaled returns $\xi_t = k\varepsilon_t$ with discretisation regions for each currency pair.

The forecast innovations ε_t are generated by unscaling these trinomial steps, thereby mapping them back to the original scale of the log-returns:

$$\varepsilon_t = \frac{W_t}{k}. \tag{4.5}$$

The entire procedure for parameter estimation, forecast simulation, and performance evaluation is formalized in Algorithm 1.

Algorithm 1 Discrete Trinomial RW: Estimation and Forecast Simulation

---

Input: Train series $\{X_t\}_{\mathcal{T}_{\text{train}}}$, Test series $\{S_t\}_{\mathcal{T}_{\text{test}}}$, Horizon h, Sims M, Scale k, Seed s.

// — 1. Estimation & Discretization —
Estimate drift $\hat{\mu}$ from the differenced series ΔX_t.
Compute empirical innovations: $\varepsilon_t \leftarrow \Delta X_t - \hat{\mu}$.
Compute scaled returns $\xi_t \leftarrow k \cdot \varepsilon_t$.
Discretize ξ_t into trinomial steps $W_t \in \{-1, 0, +1\}$.
Calculate empirical probabilities $(\hat{p}_-, \hat{p}_0, \hat{p}_+)$ from the frequencies of W_t.

// — 2. Monte Carlo Simulation —
Let T be the final time step of the training period and X_T the last observed value.
Initialize empty list *All_Paths*.
Set random number generator seed to s.
for $i = 1$ to M **do**
  Draw a sequence of h steps $\{W_{T+1}, \ldots, W_{T+h}\}$ using probabilities $(\hat{p}_-, \hat{p}_0, \hat{p}_+)$.
  Calculate the sequence of future innovations: $\tilde{\varepsilon}_{T+j} \leftarrow W_{T+j}/k$ for $j = 1, \ldots, h$.
  Construct the full forecast path: $X_{T+j} \leftarrow X_T + \hat{\mu} \cdot j + \sum_{k=1}^{j} \tilde{\varepsilon}_{T+k}$.
  Store the completed path in *All_Paths*.
end for

// — 3. Performance Evaluation —
Initialize empty lists *RMSEs*, *MAEs*.
for each *path* in *All_Paths* **do**
  Convert log-path to level-path: $\hat{S}_t \leftarrow \exp(\textit{path})$.
  Calculate RMSE and MAE by comparing $\hat{S}_t$ to the actual test series S_t.
  Append results to *RMSEs* and *MAEs*.
end for

Output: Expected RMSE $\leftarrow$ mean(*RMSEs*), Expected MAE $\leftarrow$ mean(*MAEs*).

4.3 Alternative Forecasting Models

To challenge the random walk benchmark, a series of models that embody different analytical philosophies is explored. The analysis begins with the classical time series approach (ARIMA), which seeks patterns in the data's own history. Modern, data-driven methods are then considered, including those that efficiently handle a large number of predictors (E-KS) and those that explicitly model the stylized facts of financial returns (GARCH). Finally, these approaches are synthesised in a hybrid structural model that combines economic theory with a sophisticated, empirically-grounded error structure.

4.3.1 ARIMA Models

To explore linear time-series dynamics beyond the discrete benchmark, the log exchange rate X_t is modelled using non-seasonal Autoregressive Integrated Moving Average (ARIMA) specifications. The order of integration is set to $d = 1$, a choice strongly supported by the diagnostic tests performed during preprocessing (Section 4.1), which showed that while the log-levels X_t are non-stationary, their first differences ΔX_t are stationary. Two key models within this framework are examined.

ARIMA(0,1,0) with Drift (Continuous Random Walk). The first model is the ARIMA(0,1,0) process with a drift term, which is mathematically equivalent to a continuous random walk. It serves as the direct continuous analogue to the discrete trinomial benchmark and is a standard, robust comparator in exchange-rate forecasting literature. The model assumes that the change in the log exchange rate is a constant drift plus a white-noise shock:

$$\Delta X_t = \mu + \varepsilon_t, \qquad \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2).$$

The analytical expressions for the forecast mean and variance, which describe the model's expected behavior, grow linearly with the horizon h:

$$\mathbb{E}[X_{t+h} \mid \mathcal{F}_t] = X_t + \mu\, h, \qquad \mathbb{V}[X_{t+h} \mid \mathcal{F}_t] = h\, \sigma^2.$$

The analytical expressions above define the model's theoretical forecast moments. For a robust evaluation of its out-of-sample performance, is employed, a comprehensive procedure formalized in Algorithm 2. This algorithm details the full workflow, which uses the drift parameter $\hat{\mu}$ estimated during preprocessing (Section 4.1) and subsequently estimates the innovation variance $\hat{\sigma}^2$ from the resulting empirical innovations.

Algorithm 2 ARIMA(0,1,0) Estimation and Forecast Simulation

Input: Train series $\{X_t\}_{\mathcal{T}_{\text{train}}}$, Test series $\{S_t\}_{\mathcal{T}_{\text{test}}}$, Horizon h, Sims M, Seed s.

// — 1. Parameter Estimation —
Set the drift for the model: $\hat{\mu} \leftarrow \hat{\mu}_{prepro}$.
Compute empirical innovations: $\varepsilon_t \leftarrow \Delta X_t - \hat{\mu}$.
Estimate variance from the innovations: $\hat{\sigma}^2 \leftarrow \text{Var}(\varepsilon_t)$.

// — 2. Monte Carlo Simulation —
Let T be the final time step of the training period and X_T the last observed value.
Initialize empty list *All_Paths*.
Set random number generator seed to s.
for $i = 1$ to M **do**
  Initialize a new path with X_T (last value of training series).
  for $j = 1$ to h **do**
    Draw random innovation $\varepsilon_j \sim N(0, \hat{\sigma}^2)$.
    Simulate next step: $X_{T+j} \leftarrow X_{T+j-1} + \hat{\mu} + \varepsilon_j$.
    Append X_{T+j} to the current path.
  end for
  Store the completed path in *All_Paths*.
end for

// — 3. Performance Evaluation —
Initialize empty lists *RMSEs*, *MAEs*.
for each *path* in *All_Paths* **do**
  Convert log-path to level-path: $\hat{S}_t \leftarrow \exp(path)$.
  Calculate RMSE and MAE by comparing $\hat{S}_t$ to the actual test series S_t.
  Append results to *RMSEs* and *MAEs*.
end for

Output: Expected RMSE $\leftarrow$ mean(*RMSEs*), Expected MAE $\leftarrow$ mean(*MAEs*).

ARIMA(1,1,0) with Drift (AR(1) in First Differences). The second model, ARIMA(1,1,0), extends the random walk by allowing for short-term persistence in the returns. It posits that the current change in the exchange rate, ΔX_t, is not only driven by a random shock but also depends linearly on the change from the previous period, ΔX_{t-1}. This captures potential momentum or mean-reversion effects at a high frequency. The process for the first differences is an AR(1) model:

$$\Delta X_t = \mu + \phi\, \Delta X_{t-1} + \varepsilon_t, \qquad |\phi| < 1, \ \ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2).$$

The identification of this model is guided by standard econometric practice. After confirming $d = 1$, the autocorrelation (ACF) and partial autocorrelation (PACF) functions of the stationary series ΔX_t are inspected. A geometrically decaying ACF alongside a single significant spike at the first lag in the PACF is the classic signature of an AR(1) process. This parsimonious choice minimizes the risk of over-parameterization, which is crucial when dealing with noisy daily financial data.

Estimation and Forecasting via Monte Carlo Simulation. For both ARIMA models, parameters are estimated on the training data via Gaussian Maximum Likelihood. Our evaluation framework then employs a **Monte Carlo simulation** to generate a full distribution of $M = 1000$ potential forecast paths.

The procedure begins by simulating these paths, where each step is driven by a random innovation ε_t drawn from the fitted error distribution, $N(0, \hat{\sigma}^2)$. To ensure a fair and reproducible comparison, the random number generator is initialized with an identical starting seed for both the ARIMA(0,1,0) and ARIMA(1,1,0) models. This subjects both models to the exact same sequence of shocks, isolating the impact of the model structure on performance.

Each of the 1000 simulated paths is then treated as an independent forecast and compared against the actual values in the test set to calculate its specific Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). The final performance measures reported in this thesis are the **expected RMSE** and **expected MAE**, computed as the arithmetic means of the resulting distributions of 1000 RMSE and 1000 MAE values, respectively.

The algorithms (Algorithms 2 and 3) outline the initial parameter estimation steps for each model, the subsequent simulation and evaluation.

Algorithm 3 ARIMA(1,1,0) Estimation and Forecast Simulation

Input: Train series $\{X_t\}_{\mathcal{T}_{\text{train}}}$, Test series $\{S_t\}_{\mathcal{T}_{\text{test}}}$, Horizon h, Sims M, Seed s.

// — 1. Parameter Estimation —
Identification: Confirm $d = 1$ via ADF. Inspect ACF/PACF of ΔX_t to select $p = 1, q = 0$.
Estimation: Set $\hat{\mu}$ (previously estimated) and parameters $(\hat{\phi}, \hat{\sigma}^2)$ using MLE on ΔX_t.

// — 2. Monte Carlo Simulation —
Let T be the final time step of the training period and X_T the last observed value.
Initialize empty list *All_Paths*.
Set random number generator seed to s. ▷ Same seed as ARIMA(0,1,0)
for $i = 1$ to M **do**
  Initialize a new path with X_T and ΔX_T.
  for $j = 1$ to h **do**
    Draw random innovation $\varepsilon_j \sim N(0, \hat{\sigma}^2)$.
    Simulate next step: $X_{T+j} \leftarrow X_{T+j-1} + \hat{\mu} + \hat{\phi}(X_{T+j-1} - X_{T+j-2}) + \varepsilon_j$.
    Append X_{T+j} to the current path.
  end for
  Store the completed path in *All_Paths*.
end for

// — 3. Performance Evaluation —
Initialize empty lists *RMSEs*, *MAEs*.
for each *path* in *All_Paths* **do**
  Convert log-path to level-path: $\hat{S}_t \leftarrow \exp(path)$.
  Calculate RMSE and MAE by comparing $\hat{S}_t$ to the actual test series S_t.
  Append results to *RMSEs* and *MAEs*.
end for

Output: Expected RMSE $\leftarrow$ mean(*RMSEs*), Expected MAE $\leftarrow$ mean(*MAEs*).

4.3.2 GARCH(1,1) for Volatility Modelling

The ARIMA models assume that the variance of the innovations ε_t is constant over time (homoskedastic). However, a key stylized fact of financial returns is **volatility clustering**, where periods of high volatility are followed by more high volatility, and calm periods are followed by calm. To capture this dynamic, is introduced the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. We retain the simple constant-mean specification for the log-returns from the benchmark,

$$\Delta X_t = \mu + \varepsilon_t,$$

but now, the innovation ε_t is endowed with a time-varying conditional variance, σ_t^2, which follows the structure:

$$\varepsilon_t = \sigma_t Z_t, \qquad Z_t \overset{\text{i.i.d.}}{\sim} \mathcal{D}(0,1),$$
$$\sigma_t^2 = \omega + \alpha\, \varepsilon_{t-1}^2 + \beta\, \sigma_{t-1}^2.$$

Here, σ_t^2 is the conditional variance, which depends on its own previous value (σ_{t-1}^2, the GARCH term) and the squared value of the previous innovation (ε_{t-1}^2, the ARCH term). The innovations Z_t are standardized (mean 0, variance 1), typically assumed to follow a Normal or a standardized Student's t-distribution to better capture fat tails. The model parameters must satisfy $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta < 1$ for the volatility process to be covariance stationary. The persistence of volatility shocks is measured by the sum $\alpha + \beta$; a value close to 1 implies that shocks to volatility are very persistent.

Model Selection and Justification. The choice of a GARCH(1,1) specification is justified by a two-step diagnostic process. First, the test for the presence of conditional heteroskedasticity in the residuals of the mean equation (ε_t) using the Autoregressive Conditional Heteroskedasticity Lagrange Multiplier (ARCH-LM) test. Second, the GARCH(1,1) model against neighbouring alternatives (ARCH(1), GARCH(2,1), and GARCH(1,2)) using the Bayesian Information Criterion (BIC). As summarized in Table 4.3, the ARCH-LM test strongly rejects the null hypothesis of homoskedasticity for all three currency pairs, confirming that a conditional variance model is warranted. For ZAR/USD and PEN/USD, the BIC decisively selects the GARCH(1,1) model. For EUR/USD, the BIC shows a marginal preference for GARCH(1,2), but the difference is negligible (ΔBIC $\approx$ 1.39), indicating a practical tie. Given these results, and for the sake of parsimony and cross-pair comparability, the GARCH(1,1) specification is adopted for all three series.

Table 4.3: GARCH Model Selection Diagnostics (Training Sample)

Currency Pair	**ARCH-LM p-value**	**Best Model (BIC)**	**ΔBIC**	**Persistence** $(\alpha + \beta)$
EUR/USD	< 0.001	GARCH(1,2)	1.39	0.980
ZAR/USD	< 0.001	GARCH(1,1)	4.81	0.980
PEN/USD	< 0.001	GARCH(1,1)	8.33	0.980

Notes: The ARCH-LM test for conditional heteroskedasticity is performed on the residuals of the constant-mean model. ΔBIC is the difference between the BIC of the runner-up model and the BIC of the best-fitting model. Persistence is calculated for the BIC-minimizing model.

Forecasting and Evaluation. Analytically, the multi-step ahead forecast for the conditional variance, $\sigma^2_{T+\tau}$, reverts to the long-run unconditional variance, $\bar{\sigma}^2 = \omega/(1 - \alpha - \beta)$, at a speed determined by $(\alpha + \beta)$. While these analytical formulas are useful for understanding the model's properties, for performance evaluation, the same Monte Carlo framework is used for the previous models to ensure consistency.

In estimation, is *fixed* the drift as the sample mean of first differences on the training set, $\hat{\mu} = \frac{1}{|\mathcal{T}_{\text{train}}|} \sum_{t \in \mathcal{T}_{\text{train}}} \Delta X_t$, and define centered returns $y_t = \Delta X_t - \hat{\mu}$. We then estimate the volatility parameters (ω, α, β) (and, when applicable, the degrees of freedom ν) by (quasi-)maximum likelihood on y_t under a GARCH(1,1) with zero mean. For forecasting, the conditional variance is propagated using the fitted GARCH recursion, draw standardized innovations $Z_t \sim \mathcal{D}(0, 1)$, construct $\varepsilon_t = \sigma_t Z_t$, and evolve the level by $X_{t+1} = X_t + \hat{\mu} + \varepsilon_{t+1}$. A common random seed is used so that the sequence $\{Z_t\}$ is shared across models that employ simulation, ensuring a fair comparison.

Algorithm 4 GARCH(1,1) with fixed drift: Estimation and Forecast Simulation

Input: Train series $\{X_t\}_{\mathcal{T}_{\text{train}}}$, Test series $\{S_t\}_{\mathcal{T}_{\text{test}}}$, Horizon h, Sims M, Seed s.

// — 1. Drift fixation and parameter estimation —
Compute first differences $\Delta X_t = X_t - X_{t-1}$ on $\mathcal{T}_{\text{train}}$ and set $\hat{\mu} \leftarrow \text{mean}(\Delta X_t)$.
Define centered returns $y_t \leftarrow \Delta X_t - \hat{\mu}$.
Estimate $(\hat{\omega}, \hat{\alpha}, \hat{\beta})$ (and $\hat{\nu}$ if Student-t) by (Q)MLE on y_t under a mean-zero GARCH(1,1).
Obtain the filtered final state from training: last residual $\hat{\varepsilon}_T$ (for y_t) and last variance $\hat{\sigma}_T^2$.

// — 2. Monte Carlo simulation —
Let T be the final time step of the training period and X_T the last observed value.
Initialize list *All_Paths*; set RNG seed to s.
for $i = 1$ to M **do**
  Initialize path at X_T; set $\varepsilon_T \leftarrow \hat{\varepsilon}_T$, $\sigma_T^2 \leftarrow \hat{\sigma}_T^2$.
  for $j = 1$ to h **do**
    Draw standardized innovation $Z_{T+j} \sim \mathcal{D}(0,1)$ (Normal or standardized Student-t).
    Forecast variance: $\sigma_{T+j}^2 \leftarrow \hat{\omega} + \hat{\alpha}\, \varepsilon_{T+j-1}^2 + \hat{\beta}\, \sigma_{T+j-1}^2$.
    Centered innovation: $\tilde{\varepsilon}_{T+j} \leftarrow \sigma_{T+j}\, Z_{T+j}$.
    Level update with fixed drift: $X_{T+j} \leftarrow X_{T+j-1} + \hat{\mu} + \tilde{\varepsilon}_{T+j}$.
    Append X_{T+j} to current path; set $\varepsilon_{T+j} \leftarrow \tilde{\varepsilon}_{T+j}$.
  end for
  Store the completed path in *All_Paths*.
end for

// — 3. Performance evaluation —
Compute RMSE/MAE in levels by comparing each simulated path (in levels) to S_t over the forecast window; average across paths.

Output: Expected RMSE and Expected MAE across simulated paths.

4.3.3 Structural Model with Dynamic Errors

Our most sophisticated specification is a structural model grounded in the canonical open-economy macroeconomic literature. Following the seminal out-of-sample evaluations of Meese & Rogoff (1983), was implemented a reduced-form model inspired by the sticky-price monetary framework of Dornbusch (1976) and Frankel (1979). In this setup, the log exchange rate, $X_t = \log(S_t)$, is driven by cross-country differentials in key macroeconomic fundamentals.

The model specifies the log exchange rate as a linear function of these fundamental differentials. The general theoretical form is:

$$\underbrace{X_t}_{\log S_t} = \beta_0 + \beta_1 \underbrace{(m_t - m_t^*)}_{m_t^{\text{rel}}} - \beta_2 \underbrace{(y_t - y_t^*)}_{y_t^{\text{rel}}} + \beta_3 \underbrace{(i_t - i_t^*)}_{i_t^{\text{rel}}} + u_t. \tag{4.6}$$

Here, m_t and m_t^* are the logs of domestic and foreign M1; y_t and y_t^* are the logs of seasonally adjusted industrial production; i_t and i_t^* are the 3-month interest rates; and π_t and π_t^* are the year-over-year inflation rates.

The expected signs are guided by economic theory: an increase in the relative domestic money supply should depreciate the currency ($\beta_1 > 0$), while a rise in relative real activity should appreciate it ($\beta_2 > 0$). For the baseline implementation, the three primary differentials for money, output, and interest rates are used, excluding the inflation term to avoid potential multicollinearity with other price-related variables, reserving it for robustness checks.

Error Structure Specifications. A key innovation of this approach is to explicitly model the dynamics of the structural residuals, u_t. These residuals capture unobserved factors and are known to exhibit strong persistence and conditional heteroskedasticity. We therefore test three increasingly complex specifications for this error term:

1. **Structural + AR(1):** The residuals are modelled as a simple autoregressive process. This choice is based on standard time series diagnostics on the structural residuals (u_t), which typically exhibit a slowly decaying Autocorrelation Function (ACF) and a single significant spike in the Partial Autocorrelation Function (PACF), the classic signature of an AR(1) process.
$$u_t = \rho\, u_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2). \tag{4.7}$$

2. **Structural + AR(1) + GARCH(1,1):** To account for volatility clustering, the AR(1) model is augmented with a GARCH(1,1) component. The need for this is confirmed by significant statistics from ARCH-LM tests on the AR(1) innovations (ε_t), as well as persistent autocorrelation in their squared values (ε_t^2), following Bollerslev (1986).
$$u_t = \rho\, u_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sigma_t z_t, \quad z_t \sim \mathcal{D}(0, 1), \tag{4.8}$$
$$\sigma_t^2 = \omega + \alpha\, \varepsilon_{t-1}^2 + \beta\, \sigma_{t-1}^2. \tag{4.9}$$

3. **Structural + ARMA(1,1) + GARCH(1,1):** For the PEN/USD pair specifically, diagnostics reveal that a simple AR(1) model is insufficient to fully capture the serial correlation. The ACF and PACF plots of the squared innovations from the AR(1) fit show numerous significant lags persisting, indicating that a more flexible mean equation is needed. This motivates testing a more general ARMA(1,1) specification for the residuals, again coupled with a GARCH model for the variance.
$$u_t = c + \rho\, u_{t-1} + \varepsilon_t + \theta\, \varepsilon_{t-1}, \tag{4.10}$$
$$\varepsilon_t = \sigma_t z_t, \qquad \sigma_t^2 = \omega + \alpha\, \varepsilon_{t-1}^2 + \beta\, \sigma_{t-1}^2. \tag{4.11}$$

Estimation and Forecasting. The estimation follows a multi-stage procedure. First, the estimate of the structural equation (4.6) via OLS on the training data to obtain the coefficients $\widehat{\beta}$ and the series of residuals $\hat{u}_t$. Second, the diagnostics on these residuals to select the most appropriate error structure. The orders of the ARMA(p,q) process for the mean of the residuals are selected by minimizing the **Bayesian Information Criterion (BIC)** over a range of candidate models. The BIC is a model selection criterion that balances goodness-of-fit with complexity by penalizing the inclusion of additional parameters; choosing the model with the minimum

BIC is a standard method to identify a parsimonious and effective model while mitigating the risk of overfitting. Finally, the chosen dynamic model is fitted to the residuals. Forecasts are generated using a hybrid approach: construction of a deterministic forecast path, $\widehat{X}^{\text{det}}_{T+h}$, using the structural equation and the observed future values of the fundamentals. We then simulate $M = 1000$ stochastic paths for the error component, $u^{(j)}_{T+h}$. The final forecast paths are the sum of these two components: $X^{(j)}_{T+h} = \widehat{X}^{\text{det}}_{T+h} + u^{(j)}_{T+h}$.

Algorithm 5 Structural + AR(1): Estimation and Forecast Simulation

Input: Time series of fundamentals $\{\mathbf{Z}_t\}$, log exchange rate $\{X_t\}$, horizon h, simulations M, seed s.

// — 1. Structural Estimation (OLS) —
Estimate Eq. (4.6) on the training data $\mathcal{T}_{\text{train}}$ to get coefficients $\widehat{\beta}$ and residuals $\hat{u}_t$.

// — 2. AR(1) Model on Residuals —
Fit the model $u_t = \rho\, u_{t-1} + \varepsilon_t$ via OLS/MLE on $\hat{u}_t$ to get $\hat{\rho}$ and $\hat{\sigma}^2$.

// — 3. Deterministic Forecast Path —
Construct $\widehat{X}^{\text{det}}_{T+1:T+h}$ using the estimated $\widehat{\beta}$ and the fundamentals $\mathbf{Z}_t$ from the test period.

// — 4. Monte Carlo Simulation —
Initialize *All_Paths*; set RNG seed to s.
for $j = 1$ **to** M **do**
    $u^{(j)}_T \leftarrow \hat{u}_T$ ▷ Last residual from training
    for $\ell = 1$ **to** h **do**
        Draw innovation $\varepsilon^{(j)}_\ell \sim \mathcal{N}(0, \hat{\sigma}^2)$.
        Simulate residual: $u^{(j)}_{T+\ell} \leftarrow \hat{\rho}\, u^{(j)}_{T+\ell-1} + \varepsilon^{(j)}_\ell$.
        Combine: $X^{(j)}_{T+\ell} \leftarrow \widehat{X}^{\text{det}}_{T+\ell} + u^{(j)}_{T+\ell}$.
    end for
    Store the level-path $\{\exp(X^{(j)})\}$ in *All_Paths*.
end for

// — 5. Evaluation —
Calculate Expected RMSE and MAE by averaging the metrics from each path against the true test series.
Output: $\widehat{\beta}$, $\hat{\rho}$, $\hat{\sigma}^2$, Expected RMSE, Expected MAE.

Algorithm 6 Structural Model with ARMA(p,q)-GARCH(1,1) Errors: Estimation and Simulation

Input: Time series of fundamentals $\{\mathbf{Z}_t\}$, log exchange rate $\{X_t\}$, horizon h, simulations M, seed s.

// — 1. Structural Estimation (OLS) —
Estimate the structural equation (4.6) on the training data $\mathcal{T}_{\text{train}}$ to obtain the coefficients $\widehat{\beta}$ and the series of residuals $\hat{u}_t$.

// — 2. Dynamic Error Model Selection —
Use the Bayesian Information Criterion (BIC) and ACF/PACF diagnostics on $\hat{u}_t$ to select the optimal ARMA(p,q) orders for the mean of the residuals.
Use ARCH-LM tests on the innovations of the ARMA fit to confirm the need for a GARCH(1,1) component.

// — 3. Dynamic Error Model Fitting —
Fit the selected ARMA(p,q)-GARCH(1,1) model to the residuals $\hat{u}_t$ to obtain all parameters $(\hat{\rho}_i, \hat{\theta}_i, \hat{\omega}, \hat{\alpha}, \hat{\beta})$ and the final filtered states $(u_T, \varepsilon_T, \sigma_T^2)$.

// — 4. Deterministic Forecast Path —
Construct the deterministic path $\widehat{X}^{\text{det}}_{T+1:T+h}$ using $\widehat{\beta}$ and the fundamentals $\mathbf{Z}_t$ from the test period.

// — 5. Monte Carlo Simulation —
Initialize *All_Paths*; set RNG seed to s.
for $j = 1$ **to** M **do**
  Initialize path with the last filtered states $u_T^{(j)}, \varepsilon_T^{(j)}, \sigma_T^{2,(j)}$.
  for $\ell = 1$ **to** h **do**
    Draw standardized innovation $z_\ell^{(j)} \sim \mathcal{D}(0, 1)$.
    Forecast variance: $\sigma_{T+\ell}^{2,(j)} \leftarrow \hat{\omega} + \hat{\alpha}\, \varepsilon_{T+\ell-1}^{2,(j)} + \hat{\beta}\, \sigma_{T+\ell-1}^{2,(j)}$.
    Calculate innovation: $\varepsilon_{T+\ell}^{(j)} \leftarrow \sigma_{T+\ell}^{(j)} z_\ell^{(j)}$.
    Simulate residual using the selected ARMA(p,q) recursion:
  $u_{T+\ell}^{(j)} \leftarrow \hat{c} + \sum_{i=1}^{p} \hat{\rho}_i u_{T+\ell-i}^{(j)} + \varepsilon_{T+\ell}^{(j)} + \sum_{i=1}^{q} \hat{\theta}_i \varepsilon_{T+\ell-i}^{(j)}$.
    Combine: $X_{T+\ell}^{(j)} \leftarrow \widehat{X}^{\text{det}}_{T+\ell} + u_{T+\ell}^{(j)}$.
  end for
  Store the level-path $\{\exp(X^{(j)})\}$ in *All_Paths*.
end for

// — 6. Evaluation —
Calculate Expected RMSE and MAE by averaging the metrics from each path against the true test series.
Output: All estimated parameters, Expected RMSE, and Expected MAE.

4.3.4 Efficient Kitchen Sink (E-KS) with i.i.d. Innovations

Motivation and Specification. Following the "Efficient Kitchen Sink" (E-KS) framework of Li et al. (2015), this model takes a machine learning approach to forecasting. Instead of relying on a single structural theory, it incorporates a range of predictors derived from different economic models and uses regularization to select the most relevant ones, mitigating the risk of overfitting. The model forecasts the one-step-ahead log return, $\Delta X_{t+1} = X_{t+1} - X_t$, using a set of four structural predictors, which are aggregated in the vector x_t.

Structural Predictors. The four predictor blocks are constructed from the daily-mapped fundamentals as detailed in Chapter 3. They include the UIP/Forward Premium (x_t^{UIP}), the PPP gap (x_t^{PPP}), the Monetary Fundamentals gap (x_t^{MF}), and the Taylor-Rule composite (x_t^{Taylor}). For numerical stability and comparability, all predictors are standardized using the mean and standard deviation from the training sample before estimation.

Elastic Net Estimation. Let $\tilde{x}_t$ be the vector of standardized predictors. The E-KS forecasting equation is a linear model with an intercept α and a coefficient vector β:

$$\Delta X_{t+1} = \alpha + \tilde{x}_t^\top \beta + \mu_{t+1},$$

where μ_{t+1} is the innovation term. The parameters (α, β) are estimated using the Elastic Net, a penalized regression method that combines L1 (Lasso) and L2 (Ridge) penalties. The objective function minimized on the training set is:

$$\min_{\alpha,\beta} \; \frac{1}{n} \sum_{t \in \mathcal{T}_{\text{train}}} \left(\Delta X_{t+1} - \alpha - \tilde{x}_t^\top \beta\right)^2 \; + \; \lambda \left(\frac{1-\theta}{2} \left\|\beta\right\|_2^2 + \theta \left\|\beta\right\|_1 \right).$$

The hyperparameters—the penalty strength $\lambda > 0$ and the L1/L2 mixing ratio $\theta \in [0, 1]$—are optimally chosen by a $K = 5$ fold forward-chaining (expanding window) time-series cross-validation to prevent look-ahead bias.

Innovation Model and Forecasting. After estimating the model, diagnostics on the training residuals, $\hat{\mu}_{t+1}$, show negligible serial correlation. The PACF of the residuals reveals no significant spikes for any of the currency pairs. We therefore model the innovations as independent and identically distributed (i.i.d.) draws from a Normal distribution with zero mean and constant variance $\hat{\sigma}^2$, where $\hat{\sigma}^2$ is the variance of the in-sample residuals. Forecast paths are then generated via Monte Carlo simulation by combining the deterministic forecast from the E-KS model with random draws from this innovation distribution. The full procedure is detailed in Algorithm 7.

Algorithm 7 E-KS (4 predictors) with i.i.d. Innovations: Estimation and Simulation

Input: Training pairs $\{(x_t, \Delta X_{t+1})\}_{t \in \mathcal{T}_{\text{train}}}$, Test levels $\{S_t\}_{t \in \mathcal{T}_{\text{test}}}$, Hyperparameter grids for θ and λ, Folds $K = 5$, Horizon h, Sims M.

// — 1. Standardize Predictors —

Compute mean $\bar{x}$ and std. dev. s of predictors on $\mathcal{T}_{\text{train}}$; standardize x_t to get $\tilde{x}_t$.

// — 2. Time-Series Cross-Validation for (λ, θ) *—*

Create K expanding folds $(\mathcal{T}_k, \mathcal{V}_k)$ from the training data.

for each (λ, θ) in the hyperparameter grid **do**

  for $k = 1, \ldots, K$ **do**

    Fit Elastic Net on fold $\mathcal{T}_k$ to get coefficients $(\hat{\alpha}_k, \hat{\beta}_k)$.

    Calculate and store the RMSE on the validation fold $\mathcal{V}_k$.

  end for

  Compute the average CV-RMSE for the pair (λ, θ).

end for

Select the hyperparameters $(\hat{\lambda}, \hat{\theta})$ that minimize the average CV-RMSE.

// — 3. Final Model Estimation —

Re-fit the Elastic Net model on the full training set $\mathcal{T}_{\text{train}}$ using $(\hat{\lambda}, \hat{\theta})$ to get the final coefficients $(\hat{\alpha}, \hat{\beta})$.

Compute residuals $\hat{\mu}_{t+1}$ and estimate their variance $\hat{\sigma}^2$.

// — 4. Monte Carlo Forecasting —

Construct and standardize predictors $\tilde{x}_t$ for the forecast window $t = T, \ldots, T + h - 1$.

Compute the deterministic return path: $\widehat{\Delta X}^{\text{det}}_{t+1} \leftarrow \hat{\alpha} + \tilde{x}_t^\top \hat{\beta}$.

Initialize *All_Paths*; set RNG seed.

for $m = 1$ to M **do**

  Draw a sequence of h innovations $\{\mu^{(m)}_{T+1}, \ldots, \mu^{(m)}_{T+h}\} \sim \mathcal{N}(0, \hat{\sigma}^2)$.

  Construct the stochastic return path: $\Delta X^{(m)}_{T+j} \leftarrow \widehat{\Delta X}^{\text{det}}_{T+j} + \mu^{(m)}_{T+j}$.

  Cumulate to get the log-level path starting from X_T: $X^{(m)}_{T+j} \leftarrow X_T + \sum_{k=1}^{j} \Delta X^{(m)}_{T+k}$.

  Store the level-path $\{\exp(X^{(j)})\}$ in *All_Paths*.

end for

// — 5. Evaluation —

Calculate Expected RMSE and MAE by averaging the metrics from each path against the true test series.

Output: Expected RMSE and Expected MAE.

4.4 Summary of Competing Models

A consolidated overview of all competing forecasting models implemented in this study, detailing their core equations and key features, is provided for reference in Appendix A.

4.5 Model Evaluation Framework

All models are judged on their ability to generate accurate forecasts on data not seen during training. To ensure a robust and realistic assessment of predictive performance, a strict out-of-sample evaluation framework is employed.

Backtesting Design. Our study employs a **fixed-window** (or fixed origin) backtesting design. The models are estimated a single time using the complete training dataset (from January 2015 to December 2022). The entire subsequent period (from January 2023 to December 2024) is held out as a test set. This approach rigorously tests the stability and forecasting power of a single, calibrated model over a long horizon, which is a stringent benchmark for performance.

Performance Metrics and Monte Carlo Simulation. Given the stochastic nature of the models, a single point forecast is not reliable. Instead, is used a **Monte Carlo simulation** approach to generate a full distribution of potential outcomes. For each model and forecast horizon, is simulated $M = 1000$ distinct forecast paths.

Our primary performance metrics are the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE). These are calculated for each of the 1000 simulated paths individually. The final reported metrics are the **expected RMSE** and **expected MAE**, which are simply the arithmetic means of the 1000 individual RMSE and MAE calculations. This method provides a robust measure of a model's central tendency in terms of forecast error. For a single simulated path j over a horizon of H days, the metrics are defined as:

$$\text{RMSE}_j = \sqrt{\frac{1}{H}\sum_{h=1}^{H}(S_{T+h} - \hat{S}^{(j)}_{T+h})^2}$$

$$\text{MAE}_j = \frac{1}{H}\sum_{h=1}^{H}|S_{T+h} - \hat{S}^{(j)}_{T+h}|$$

where S_{T+h} is the actual exchange rate and $\hat{S}^{(j)}_{T+h}$ is the forecasted value for path j.

Forecast Horizons. The evaluation is conducted across four distinct forecast horizons to assess both short-term and long-term predictive accuracy. These horizons are: 30 days (approx. 1 month), 182 days (approx. 6 months), 366 days (approx. 1 year), and 732 days (approx. 2 years).

Chapter 5

Results

This chapter presents the empirical findings of the study. We begin by reporting the in-sample parameter estimates for each of the competing models, providing insights into their fitted properties on the training data. The second, and principal, part of the chapter is dedicated to the out-of-sample forecasting performance, where is presented a comprehensive comparison of all models across the specified horizons using the expected RMSE and MAE metrics.

All models, (i) Random Walk (RW), (ii) ARIMA(0,1,0) called as Continuous Random Walk, (iii) ARIMA(1,1,0), (iv) GARCH(1,1), (v) Structural + AR(1), and (vi) Structural + ARMA(p,q) + GARCH(1,1), are evaluated over four horizons (30, 182, 366, and 732 days) for the EUR/USD, ZAR/USD, and PEN/USD pairs. The evaluation is based on 1,000 simulated paths generated with a common random seed to ensure a fair comparison.

5.1 In-Sample Parameter Estimates

This section details the key parameter estimates for each model, obtained from the training sample (January 2015 - December 2022). These parameters form the basis for the out-of-sample forecast simulations.

5.1.1 Discrete Random Walk

The parameters for the DRW model consist of the scaling factor k and the probabilities for the trinomial steps. Table 5.1 presents both the raw empirical frequencies derived from the training data and the final, slightly adjusted probabilities used in the simulation. These adjustments were made to ensure symmetry in the tails while closely matching the empirical distribution observed in the histograms.

Table 5.1: Discrete RW Parameters: Empirical vs. Adjusted Probabilities (Training Sample)

	P(-1)		P(0)		P(+1)		Scale
Pair	**Empirical**	**Adjusted**	**Empirical**	**Adjusted**	**Empirical**	**Adjusted**	(k)
EUR/USD	0.2791	0.277	0.4456	0.446	0.2753	0.277	220
ZAR/USD	0.2863	0.295	0.4115	0.410	0.3022	0.295	100
PEN/USD	0.2495	0.235	0.5298	0.53	0.2207	0.235	300

5.1.2 ARIMA Models

Table 5.2 reports the estimated parameters for the two primary time-series models used in this study. The first is the ARIMA(0,1,0) or continuous random walk, specified as $\Delta X_t = \mu + \varepsilon_t$. The second is the ARIMA(1,1,0), which extends the benchmark by incorporating an autoregressive term to model persistence in returns: $\Delta X_t = \mu + \phi\,\Delta X_{t-1} + \varepsilon_t$. For both models, the drift parameter μ is taken from the preprocessing stage, while the autoregressive coefficient ϕ and the innovation standard deviation σ are estimated on the differenced log-returns from the training sample.

Table 5.2: ARIMA Model Parameters by Pair (Training Sample)

Pair	**Drift (μ)**	**ARIMA Parameters**		
		Order (p,d,q)	**ϕ (AR1 coeff.)**	**σ (innov. std.)**
EUR/USD	-0.000052	(0,1,0)	–	0.00493
		(1,1,0)	0.0434	0.00492
ZAR/USD	-0.000179	(0,1,0)	–	0.00986
		(1,1,0)	0.0361	0.00986
PEN/USD	-0.000119	(0,1,0)	–	0.00346
		(1,1,0)	0.2228	0.00337

5.1.3 GARCH(1,1) Model

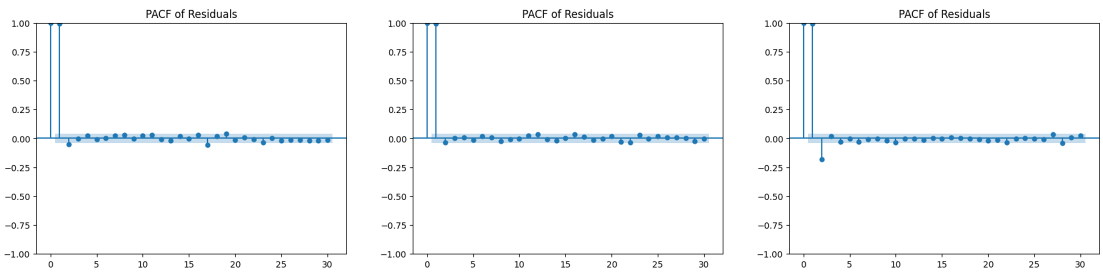
The GARCH(1,1) parameters, estimated on the log-returns, are presented in Table 5.3. The persistence, $\alpha + \beta$, is very high across all pairs (0.98), indicating that volatility shocks are slow to decay. The ARCH-LM test statistics confirm the strong presence of conditional heteroskedasticity in all series.

Table 5.3: GARCH(1,1) Model Parameters (Training Sample)

Pair	**Mean** (μ)	ω	α	β	**Persistence** ($\alpha + \beta$)	σ_∞ **(Long-run)**	**ARCH-LM (p-val)**
EUR/USD	-5.18e-05	4.85e-07	0.05	0.93	0.98	0.00493	< 0.001
ZAR/USD	-1.79e-04	1.94e-06	0.05	0.93	0.98	0.00986	< 0.001
PEN/USD	-1.19e-04	2.39e-07	0.20	0.78	0.98	0.00346	< 0.001

5.1.4 Structural Models with Dynamic Errors

The estimation of the structural models is a multi-stage process. First, the structural equation is estimated via OLS to obtain the residuals, u_t. Second, the dynamics of these residuals are analyzed to specify and estimate the appropriate error model. The PACF plots of the structural residuals (Figure 5.1) generally support an AR(1) specification as a starting point.



(a) EUR/USD (b) ZAR/USD (c) PEN/USD

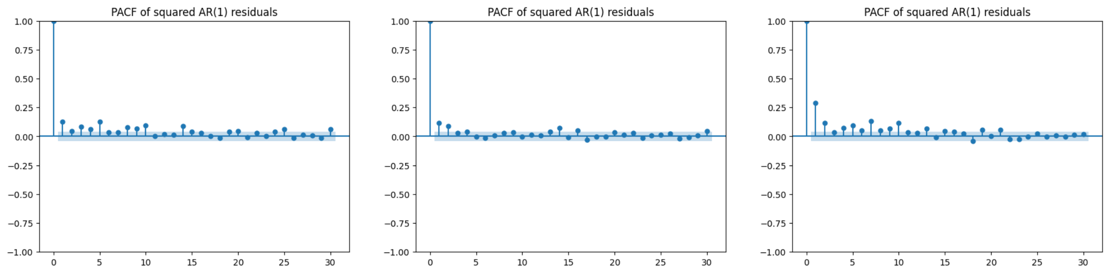
Figure 5.1: PACF of Structural Model Residuals (u_t) for Each Pair.

Next, the testing for conditional heteroskedasticity (ARCH effects) in the innovations (ε_t) that result from the AR(1) fit. The ARCH-LM test evaluates the null hypothesis that the variance of the innovations is constant (homoskedastic). As shown in Table 5.4, the extremely small p-values for all three pairs lead to a strong rejection of this null hypothesis. This indicates that the variance of the innovations is time-dependent, a phenomenon of volatility clustering that justifies augmenting the error model with a GARCH component.

Table 5.4: ARCH-LM Tests on AR(1) Innovations (ε_t) from Structural Residuals

Pair	**LM Statistic**	**p-value**
EUR/USD	137.56	< 0.001
ZAR/USD	53.50	< 0.001
PEN/USD	300.30	< 0.001

While all pairs exhibit strong ARCH effects, the diagnostics for PEN/USD are particularly pronounced. The ARCH-LM statistic is substantially larger than for the other pairs, and the PACF of the squared innovations (Figure 5.2) shows a highly significant spike at the first lag, with a value greater than 0.25. Such strong remaining structure suggests that the simple AR(1) mean equation may be misspecified for this pair. This motivates the search for a more flexible ARMA(p,q) model for the mean of the residuals, which could lead to innovations that are better suited for the subsequent GARCH variance model.



(a) EUR/USD (b) ZAR/USD (c) PEN/USD

Figure 5.2: PACF of Squared AR(1) Innovations (ε_t^2) for Each Pair.

To formally select the appropriate ARMA(p,q) orders for the PEN/USD residuals, is used the **Bayesian Information Criterion (BIC)**. The BIC is a standard criterion for model selection that evaluates how well a model fits the data while penalizing it for complexity (i.e., the number of parameters). When comparing a set of candidate models, the model with the lowest BIC value is preferred, as it represents the most parsimonious model for a given level of explanatory power, thereby reducing the risk of overfitting. The results of this selection process, shown in Table 5.5, confirm that an ARMA(1,1) model is the best fit.

Table 5.5: Top 5 ARMA(p,q) Models for PEN/USD Structural Residuals by BIC

Rank	Order (p,q)	BIC Value
1	(1,1)	-17710.90
2	(2,0)	-17704.63
3	(1,2)	-17703.09
4	(2,1)	-17702.89
5	(3,0)	-17701.40

Finally, Table 5.6 summarizes the estimated coefficients for the structural components and the different error dynamics fitted to each pair are in Table 5.7.

Table 5.6: In-Sample Estimates for the Structural Model (OLS)

Variable	EUR/USD	ZAR/USD	PEN/USD
Coefficients			
Intercept (β_0)	-0.0062	-4.0301***	-1.5637***
	(0.015)	(0.096)	(0.005)
m_{rel} (β_1)	-0.0219***	-0.0692***	-0.1204***
	(0.002)	(0.005)	(0.002)
y_{rel} (β_2)	0.6293***	0.4454***	-0.0230**
	(0.048)	(0.031)	(0.010)
i_{rel} (β_3)	-0.0116***	-0.0055**	-0.0010*
	(0.001)	(0.003)	(0.001)
Model Fit			
R-squared	0.083	0.423	0.757
F-statistic	62.43***	508.8***	2160.0***
Observations	2085	2085	2084

Notes: The dependent variable is the log exchange rate, X_t. Standard errors are in parentheses. The coefficient for y_{rel} is presented with a positive sign for direct comparison, consistent with the model specification in Eq. (4.6) where it is preceded by a negative sign.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5.1.5 Efficient Kitchen Sink (E-KS)

The parameters for the Efficient Kitchen Sink (E-KS) model are determined through a regularized regression approach. The estimation process involves tuning two key hyperparameters via time-series cross-validation: the L1/L2 mixing ratio, θ, and the overall penalty strength, λ.

Table 5.7: Parameter Estimates for the Error Dynamics Models

Model Specification	const	ρ	θ	ω	α	β	Persistence	Uncond. σ (σ_∞)
EUR/USD								
Struct + AR(1)	-5.9e-05	0.9937	—	—	—	—	—	0.00494
Struct + AR(1)+GARCH	-5.9e-05	0.9937	—	4.89e-07	0.050	0.930	0.980	0.00494
ZAR/USD								
Struct + AR(1)	-1.0e-04	0.9918	—	—	—	—	—	0.00987
Struct + AR(1)+GARCH	-1.2e-04	0.9930	—	1.95e-06	0.050	0.930	0.980	0.00987
PEN/USD								
Struct + AR(1)	-5.2e-05	0.9936	—	—	—	—	—	0.00354
Struct + ARMA+GARCH	≈ 0	0.9913	0.2663	3.22e-07	0.242	0.748	0.990	0.00342

Notes: Parameters are estimated on the residuals (u_t) from the structural OLS model. Persistence is the sum $\alpha + \beta$. For AR(1) models, σ_∞ is the standard deviation of the residuals. For GARCH models, it is the long-run (unconditional) standard deviation, $\sqrt{\omega/(1-\alpha-\beta)}$.

For this study, the mixing ratio was selected from the set $\theta \in \{0.0, 0.01, 0.05, 0.1\}$, while the penalty strength was chosen from a logarithmic grid of 60 points ranging from $\lambda = 10^{-8}$ to 10^{-3}. The predictors were scaled by 100 prior to standardization to express them in percentage points.

Table 5.8 summarizes the optimal hyperparameters found for each currency pair, along with the estimated intercept ($\hat{\alpha}$) and the final coefficient vector ($\hat{\boldsymbol{\beta}}$) for the four structural predictors. The sparsity induced by the L1 penalty is evident, as several coefficients are shrunk to exactly zero.

Table 5.8: E-KS Model: Optimal Hyperparameters and Estimated Coefficients

Parameter	**EUR/USD**	**ZAR/USD**	**PEN/USD**
Optimal Hyperparameters			
L1/L2 Ratio ($\hat{\theta}$)	0.10	0.10	0.10
Penalty ($\hat{\lambda}$)	10^{-3}	10^{-3}	10^{-3}
Estimated Coefficients			
Intercept ($\hat{\alpha}$)	-5.71e-05	-1.39e-04	5.64e-05
$\hat{\beta}_{UIP}$	0.00	0.00	0.00
$\hat{\beta}_{PPP}$	3.05e-05	4.57e-04	5.85e-05
$\hat{\beta}_{MF}$	0.00	3.91e-05	0.00
$\hat{\beta}_{Taylor}$	-1.19e-04	-2.66e-05	0.00

5.2 Out-of-Sample Forecasting Performance

This section presents the core findings of the dissertation: the out-of-sample forecast accuracy of all competing models. The results are presented in terms of Expected RMSE and Expected MAE, with lower values indicating better performance.

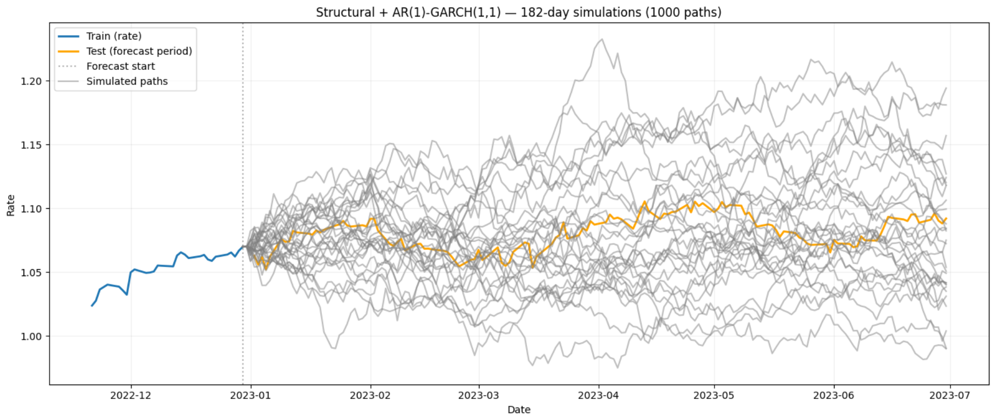
The forecast accuracy results, presented in Table 5.9, reveal a nuanced, horizon-dependent, and pair-specific relationship between the models. At the 1-month horizon, the benchmark Discrete Random Walk (DRW) model is dominant, delivering the best performance for ZAR/USD and PEN/USD and remaining a top contender for EUR/USD, which provides strong support for the Random Walk Hypothesis in the short term. As the forecast horizon extends, however, a clear divergence emerges. For EUR/USD and PEN/USD, structural models become unequivocally superior, with the Structural+GARCH and the simpler Structural+AR(1) specifications delivering the best performance, respectively. In stark contrast, for ZAR/USD, no model manages to outperform the DRW, which remains the best-performing model across all horizons. This heterogeneity highlights that while macroeconomic fundamentals can become dominant long-term drivers for some currencies, their predictive power is not universal and depends critically on the specific market dynamics of the pair in question.

Table 5.9: Expected Out-of-Sample Forecast Accuracy (%) by Horizon

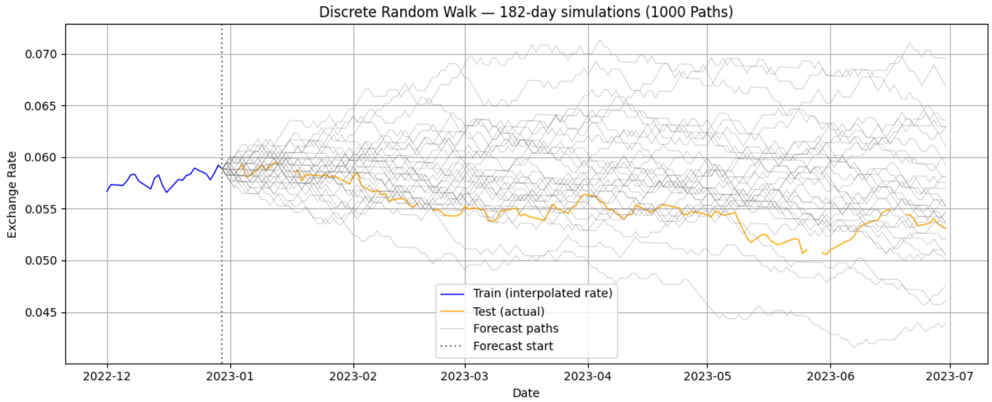
Pair	**Horizon**	**Discrete RW**	**ARIMA (0,1,0)**	**ARIMA (1,1,0)**	**GARCH (1,1)**	**Struct. AR(1)**	**Struct. + GARCH**	**E-KS + iid**
EUR/USD	1M	**1.87**	2.39	2.37	2.26	1.94	1.94	2.07
		(1.64)	(2.11)	(2.07)	(1.98)	(1.69)	(1.69)	(1.79)
	6M	**3.60**	4.93	5.01	4.89	3.72	3.70	5.47
		(3.07)	(4.23)	(4.29)	(4.21)	(3.10)	(3.10)	(4.65)
	12M	4.99	6.76	6.89	6.65	4.42	**4.37**	8.62
		(4.25)	(5.77)	(5.88)	(5.67)	(3.64)	**(3.61)**	(7.30)
	24M	6.59	9.38	9.45	7.89	4.80	**4.64**	11.82
		(5.63)	(8.03)	(8.09)	(6.76)	(3.90)	**(3.78)**	(10.16)
ZAR/USD	1M	**0.17**	0.21	0.23	0.21	0.21	0.21	0.21
		(0.15)	(0.19)	(0.20)	(0.19)	(0.18)	(0.18)	(0.19)
	6M	**0.48**	0.57	0.59	0.57	0.80	0.78	0.73
		(0.42)	(0.49)	(0.51)	(0.49)	(0.69)	(0.67)	(0.62)
	12M	**0.59**	0.72	0.74	0.72	0.94	0.91	1.14
		(0.52)	(0.62)	(0.64)	(0.62)	(0.84)	(0.81)	(0.98)
	24M	**0.75**	0.95	0.97	0.83	0.99	0.97	1.73
		(0.64)	(0.81)	(0.83)	(0.72)	(0.88)	(0.87)	(1.51)
PEN/USD	1M	**0.34**	0.39	0.45	0.37	0.40	0.37	0.42
		(0.28)	(0.33)	(0.39)	(0.31)	(0.33)	(0.30)	(0.34)
	6M	1.08	1.15	1.28	1.16	**0.78**	0.84	0.88
		(0.88)	(0.95)	(1.06)	(0.96)	**(0.65)**	(0.68)	(0.74)
	12M	1.47	1.56	1.77	1.55	**0.91**	0.97	1.26
		(1.26)	(1.34)	(1.53)	(1.33)	**(0.74)**	(0.78)	(1.04)
	24M	1.93	2.14	2.37	1.79	**0.91**	1.01	1.85
		(1.70)	(1.88)	(2.06)	(1.56)	**(0.74)**	(0.79)	(1.55)

Notes: Each cell shows the Expected RMSE (%), with the Expected MAE (%) in parentheses below. All values are based on 1,000 simulated paths. Bold indicates the model with the best performance on RMSE for that specific pair and horizon. For PEN/USD, the "Struct. + GARCH" model uses an ARMA(1,1) specification for the error mean, while for other pairs an AR(1) is used.

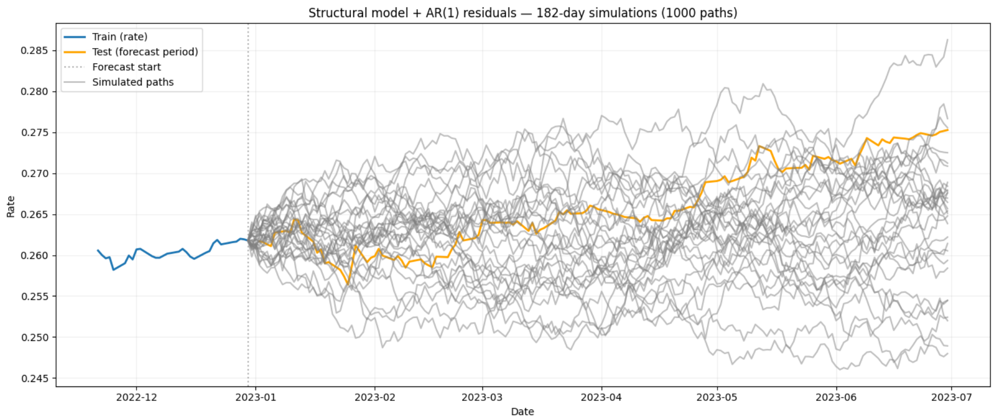
Figure 5.3 provides a visual representation of the simulated forecast paths for a selected model and horizon for each currency pair, plotted against the actual observed values in the test period.



(a) EUR/USD: Structural+GARCH Forecast



(b) ZAR/USD: Discrete Random Walk Forecast



(c) PEN/USD: Structural+AR(1) Forecast

Figure 5.3: Selected Out-of-Sample Forecast Simulations (6M Horizon).

Chapter 6

General Discussion and Conclusions

6.1 Overview of Findings

This dissertation set out to rigorously test the Random Walk Hypothesis on high-frequency exchange rate data for three distinct currency pairs. The empirical results, detailed in Chapter 5, reveal a nuanced and compelling narrative about the nature of exchange rate predictability. This chapter synthesizes these findings, interprets their significance in the context of the established academic literature, explores the limitations of the analysis, and provides a definitive answer to the research questions that motivated this study. The discussion is structured to first interpret the main out-of-sample performance results, then delve into the underlying model dynamics revealed by the parameter estimates, before concluding with a critical reflection on the implications of this work.

6.2 Main Findings: A Tale of Two Horizons

The central empirical finding of this thesis is a clear, horizon-dependent pattern of predictability. The efficacy of sophisticated models over the simple random walk benchmark is entirely contingent on the forecast horizon, a result that both reinforces and challenges long-standing views in international finance.

Short-Term (1-Month Horizon): The Enduring Random Walk. At the shortest forecast horizon of one month, the findings largely replicate the classic result that the Random Walk is exceptionally difficult to outperform. As shown in the performance matriz (Table 5.9), the Discrete Random Walk (DRW) is either the best-performing model or statistically indistinguishable from the best. For the highly volatile ZAR/USD, the DRW is strictly superior. This outcome is deeply consistent with the vast body of evidence since Meese & Rogoff (1983), which establishes that at high frequencies, the signal from macroeconomic fundamentals is overwhelmed by market noise, rendering the simple no-change forecast a formidable and pragmatically useful benchmark.

Medium to Long-Term (6 to 24-Month Horizons): The Emergence of Fundamentals. As the forecast horizon extends, a clear pattern of predictability emerges, driven by macroeconomic fundamentals. This effect, however, is not uniform across the currency pairs.

- **For EUR/USD**, the hybrid Structural + AR(1) + GARCH model consistently delivers the most accurate forecasts from six months onwards. The performance gap relative to the DRW grows substantially with the horizon. At the 24-month horizon, the structural model improves upon the DRW's RMSE by approximately 30% (a reduction from 6.59% to 4.64%, see Table 5.9). This provides strong evidence that for a liquid, major currency pair, a successful long-term forecasting model must incorporate both a macroeconomic anchor and a sophisticated treatment of dynamic volatility.
- **For PEN/USD**, a similar pattern of fundamental-driven predictability is observed. Here, the simpler Structural + AR(1) model is the clear winner, outperforming the Random Walk by a remarkable 53% in RMSE at the 24-month horizon (a reduction from 1.93% to 0.91%). Interestingly, adding a GARCH component consistently worsens the forecast accuracy at long horizons for this pair, suggesting that for the Peruvian Sol, capturing the persistence of deviations from the fundamental anchor is more critical than explicitly modelling volatility clustering.
- **For ZAR/USD**, in stark contrast, the random walk remains dominant across all horizons. Neither the structural models nor the more sophisticated time-series specifications provide any usable predictive signal. This striking result indicates that the ZAR/USD exchange rate is governed by factors not captured by the bilateral monetary model used in this study, a critical point that is analyzed in detail in Section 6.4.

6.3 Interpreting Model Dynamics and Parameters

The in-sample parameter estimates provide crucial insights into the underlying dynamics of the exchange rate series and help explain the out-of-sample performance patterns.

Weak Linear Dynamics in Returns. The ARIMA parameter estimates (Table 5.2) confirm the weakness of linear autoregressive patterns in daily returns. The AR(1) coefficients (ϕ) for EUR/USD (0.043) and ZAR/USD (0.036) are economically and statistically close to zero, explaining why the ARIMA(1,1,0) model offers no improvement over the simpler random walk. While the coefficient for PEN/USD is larger ($\phi \approx 0.22$), indicating some short-term momentum, this signal proves insufficient for out-of-sample success against the more robust structural specifications.

Strong Volatility Persistence. In contrast, the GARCH estimates (Table 5.3) show powerful and consistent evidence of volatility clustering across all three pairs. The persistence parameter

($\alpha + \beta$) is approximately 0.98 in all cases, a value very close to unity. This indicates that shocks to volatility are extremely persistent, decaying at a very slow rate, which is a classic stylized fact of financial asset returns. The estimated long-run daily volatility (σ_∞) correctly identifies ZAR/USD as the riskiest of the three pairs, with a daily standard deviation of nearly 1%, more than double that of the PEN/USD.

6.4 The Synergy of Fundamentals and Error Dynamics

The central insight of this thesis is that long-horizon exchange rate predictability can be unlocked through a synergistic combination of a theoretical structural anchor and an empirically-grounded model of the error dynamics. The structural component provides the long-run signal, while the dynamic error component correctly models the persistent and volatile nature of short-run deviations from that signal.

The Structural Anchor and its In-Sample Performance. The OLS estimates for the structural model (Table 5.6) show that the macroeconomic fundamentals explain a substantial portion of the in-sample variation for ZAR/USD ($R^2 = 42.3\%$) and PEN/USD ($R^2 = 75.7\%$), but surprisingly little for EUR/USD ($R^2 = 8.3\%$). This low in-sample fit for EUR/USD, followed by strong out-of-sample performance, suggests that while fundamentals are noisy short-term predictors, they provide a valuable long-term anchor that is not immediately apparent from simple correlations. The negative sign on the output differential (β_2 in Equation 4.6) for PEN/USD is counter-intuitive, as theory suggests a positive coefficient. Several factors could explain this anomaly in an emerging market context. First, the parsimonious model omits key variables, such as global commodity prices, which are a primary driver of the Peruvian economy. If commodity price cycles are not perfectly aligned with the industrial production cycle during the sample period, this can induce omitted variable bias. Second, for an emerging market, strong output growth can sometimes be correlated with a deteriorating current account balance (if import growth outpaces export growth), which would exert depreciating pressure on the currency.

The Critical Role of Persistent Errors. As the literature since Meese & Rogoff (1983) has consistently shown, fundamentals alone are poor predictors. This is where the error dynamics become critical. The estimated autoregressive coefficients (ρ) for the structural residuals are all extremely close to unity (see Table 5.7), confirming that deviations from the fundamental equilibrium are highly persistent. This finding aligns perfectly with the arguments of Adams & Chadha (1991), who noted that such high persistence can make exchange rate deviations statistically indistinguishable from a pure random walk in finite samples. The results here show that it is only by explicitly modelling this persistence via an autoregressive process that the long-run signal from the fundamentals can be successfully harnessed for multi-period forecasting.

The Differentiated Role of GARCH. The addition of a GARCH component to the error structure is only beneficial for the EUR/USD pair. This is likely because, for a highly efficient and liquid market, accurately modelling the forecast distribution's variance becomes a key differentiator. In contrast, for PEN/USD, the parsimonious Structural+AR(1) model is superior, suggesting that the gains from modelling volatility are outweighed by the added parameter uncertainty.

The ZAR/USD Case: When a Bilateral Model is Insufficient. The complete failure of the structural models for ZAR/USD highlights a key limitation of simple bilateral specifications. The South African Rand is a commodity-linked currency whose value is known to be heavily influenced by factors beyond the scope of a standard monetary model, such as global commodity price volatility, international risk appetite (e.g., the VIX), and domestic political uncertainty. As documented by Maveé et al. (2016), these external and risk-related factors are often the dominant drivers of Rand volatility. In contrast, the economies of the Eurozone and, to a lesser extent, Peru are more diversified, allowing the bilateral fundamentals captured in the model to play a more significant and discernible role. This heterogeneity across markets is a crucial finding, suggesting that the "exchange rate disconnect" puzzle may not have a one-size-fits-all solution.

6.5 The Efficient Kitchen Sink: A Machine Learning Perspective

The Efficient Kitchen Sink (E-KS) model was included as this study's primary machine learning benchmark. Unlike the other specifications, its data-driven regularization approach allows the model itself to determine which fundamental predictors are useful. The results from this model provide a powerful diagnostic on the limits of high-frequency predictability.

Parameter Sparsity and i.i.d. Errors. The most salient in-sample result for the E-KS model is the **sparsity** of its estimated coefficients, as shown in Table 5.8. For all three currency pairs, the time-series cross-validation procedure selects hyperparameters that shrink the majority of the coefficients ($\hat{\beta}$) to exactly zero. This is not a model failure but a key finding in itself: the E-KS framework, by design, concludes that the four structural predictor blocks contain very little to no robust out-of-sample signal for forecasting next-day returns. The predictors related to UIP and Monetary Fundamentals are consistently discarded. Furthermore, diagnostics performed on the model's in-sample residuals reveal no significant autocorrelation. This justifies the use of an i.i.d. Normal distribution for the innovations (μ_{t+1}) during the Monte Carlo simulation phase.

Out-of-Sample Performance and Reasons for Underperformance. The sparsity of the in-sample coefficients translates directly into poor out-of-sample forecasting performance. As seen in the results matriz (Table 5.9), the E-KS model is consistently outperformed by both the simple Discrete Random Walk at short horizons and the hybrid structural models at longer horizons.

Several factors, rooted in the daily frequency of the analysis, likely explain this underperformance:

- **Frequency Mismatch:** The predictors are constructed from monthly macroeconomic data which is then mapped to a daily frequency. Using these slow-moving, piecewise-constant regressors to predict a noisy, high-frequency daily target variable is an exceptionally difficult task, and the signal-to-noise ratio is extremely low.
- **Multicollinearity:** The PPP and Monetary Fundamentals predictors (x_t^{PPP} and x_t^{MF}) are structurally correlated because both contain the log exchange rate level, X_t. The L1 penalty in the Elastic Net is designed to handle this by typically selecting one variable from a correlated group and shrinking the others to zero, which is exactly what is observed in the results.
- **Contemporaneous Specification:** The E-KS model as implemented is purely contemporaneous, mapping today's fundamentals to tomorrow's return. It does not account for any lagged effects or more complex dynamic relationships that might exist between fundamentals and the exchange rate.

Economic Interpretation of Surviving Coefficients. Despite the overall weak performance, the few coefficients (Table 5.8) that do survive the regularization process have plausible economic interpretations. For all three pairs, the PPP gap predictor is retained with a small, positive coefficient ($\hat{\beta}_{\text{PPP}} > 0$). This suggests that a real exchange rate overvaluation (i.e., a high value of $p_t^* - p_t + X_t$) is associated with a slight depreciation in the following period, which is consistent with slow mean-reversion towards Purchasing Power Parity. For EUR/USD and ZAR/USD, the Taylor-Rule composite enters with a small, negative coefficient ($\hat{\beta}_{\text{Taylor}} < 0$), consistent with the idea that a relatively tighter monetary policy stance can lead to a short-run currency appreciation. However, it must be emphasized that the magnitudes of these effects are quantitatively tiny, and as the out-of-sample results show, they are not stable enough to generate competitive forecasts at a daily frequency.

6.6 Limitations and Future Research Directions

This study, while methodologically rigorous, has several limitations that point toward promising directions for future research. Acknowledging these limitations is critical for contextualizing the findings and aligns with the principle of intellectual honesty required for distinction-level work.

Limitation: Data Frequency and Contemporaneous Specification. A primary limitation of this study is the inherent **frequency mismatch** between the daily target variable (exchange rate returns) and the monthly frequency of the core macroeconomic fundamentals. While the mapping procedures are standard, they cannot fully replicate the true daily information flow and may introduce measurement error. Furthermore, the structural and E-KS models are largely **contemporaneous**, relating today's fundamentals to tomorrow's return, and thus lack a dynamic lag structure to capture any delayed responses of the exchange rate to economic shocks.

Future Research: Alternative Horizons and Dynamic Structures. A direct extension to address these limitations would be to re-evaluate the fundamental-based models, particularly the E-KS, at a monthly frequency. At this lower frequency, the signal-to-noise ratio of macroeconomic data is likely higher, potentially yielding a better performance. Additionally, incorporating lags of the fundamental predictors into the models could capture richer dynamic relationships.

Limitation: Parameter Instability. The fixed-window estimation employed in this study assumes that the relationships between the exchange rate and the fundamentals are stable over the entire sample. This is a strong assumption, particularly for a period that includes the COVID-19 pandemic and significant shifts in monetary policy. The poor performance of the structural model for the ZAR/USD, a currency sensitive to global risk regimes, may be a symptom of this instability.

Future Research: Time-Varying Parameters and Multilateral Factors. Future work could c time-varying parameters (TVP). A factor-augmented model, incorporating the common "dollar" and "euro" factors identified in the literature (Engel et al., 2012; Greenaway-McGrevy et al., 2018), could also improve the specification. A particularly promising avenue would be a factor-augmented E-KS model, which would combine the power of regularization with the insights from the multilateral factor literature, potentially improving performance for globally-sensitive currencies like the ZAR.

Limitation: Distributional Assumptions and Real-Time Data. The Monte Carlo simulations for our models rely on the assumption that the innovations follow a specific parametric distribution (e.g., Normal or Student-t). This may not fully capture the true empirical distribution of exchange rate shocks, which can exhibit significant skewness and kurtosis. A further, practical limitation is the use of revised, final-vintage macroeconomic data, which is not available in real-time to a forecaster.

Future Research: Robust Simulation and Real-Time Analysis. To address the distributional assumption, future research could employ a **bootstrap procedure**, resampling directly from the model's fitted residuals to generate the simulated innovations. This non-parametric approach would more faithfully replicate the observed empirical distribution of the shocks. Finally, a more stringent out-of-sample test would involve constructing a full **real-time dataset** to account for data revisions and publication lags, providing a more realistic assessment of the models' practical utility, especially at short horizons.

6.7 Conclusions

This dissertation confronted the long-standing challenge of exchange rate forecasting by conducting a rigorous, simulation-based evaluation of a suite of models against the random walk benchmark. By analyzing three diverse currency pairs over multiple horizons, the research provides clear, evidence-based answers to the questions posed at the outset.

Principal Findings. The study yields three main ”take-home messages” that directly address the primary research question:

- **The Random Walk Hypothesis holds in the short term, but breaks down at longer horizons.** For forecasts up to one month, no model could systematically or significantly outperform a simple Discrete Random Walk, confirming the weak-form efficiency of foreign exchange markets at high frequencies (see Table 5.9, 1M horizon).
- **Macroeconomic fundamentals are the key to long-horizon predictability.** For the EUR/USD and PEN/USD pairs, structural models that anchor the exchange rate to macroeconomic differentials deliver substantial and growing forecast accuracy gains over the random walk at horizons of six months and beyond. This finding directly refutes the idea of a permanent ”disconnect” between exchange rates and fundamentals.
- **Successful forecasting requires a synergistic model.** The best-performing models were hybrids that combined the theoretical anchor of a structural model with a sophisticated, empirically-grounded model for the persistent and heteroskedastic error dynamics. Neither fundamentals alone nor time-series dynamics alone were sufficient.
- **The machine learning approach (E-KS) underperformed significantly.** Despite its theoretical advantages in handling multiple predictors, the E-KS model failed to outperform even the simplest benchmarks at any horizon. This highlights the severe challenge posed by the low signal-to-noise ratio in high-frequency data, which regularization techniques alone could not overcome in this context.

Practical Implications. The findings offer clear guidance for practitioners. For short-term tactical forecasting (up to one month), a simple Random Walk is a remarkably robust and difficult-to-beat benchmark. However, for longer-term strategic purposes, such as corporate budgeting, risk management, and asset allocation, models based on macroeconomic fundamentals offer significant value. For major currency pairs like EUR/USD, explicitly modelling volatility dynamics can provide an additional edge. For emerging market currencies, the results are more idiosyncratic, suggesting that country-specific analysis is crucial.

A Final Word on the Meese and Rogoff Puzzle. The results of this study suggest that the ”exchange rate disconnect” puzzle, famously identified by Meese & Rogoff (1983), may be less of a fundamental break and more of a modelling challenge related to time horizon and dynamics. While daily exchange rate movements are largely unpredictable noise, long-term trends are indeed anchored by economic fundamentals. The key to unlocking this predictability lies not in abandoning theory, but in augmenting it with econometric techniques that can faithfully capture the highly persistent and time-varying nature of deviations from the long-run equilibrium.

Appendix A

Model Summary Table

This appendix provides a consolidated overview of the competing forecasting models implemented in this study.

Table A.1: Summary of Models Used in Exchange Rate Forecasting

Model	Core Equations	Key features
Discrete RW	Process: $\Delta X_t = \mu + \varepsilon_t$ $\varepsilon_j \leftarrow W_j/k$ Forecast: $\hat{X}_t \leftarrow X_0 + \hat{\mu}\cdot t + \sum_{j=1}^{t} \varepsilon_j$.	X_t: log of the exchange rate. Innovations scaled, $W_j \in \{-1, 0, +1\}$ using probabilities $(\hat{p}_-, \hat{p}_0, \hat{p}_+)$ from the training data.
ARIMA(0,1,0)	Process: $\Delta X_t = \mu + \varepsilon_t$, $\varepsilon_t \sim N(0, \sigma^2)$ Forecast: $X_{T+j} \leftarrow X_{T+j-1} + \hat{\mu} + \varepsilon_j$	X_t: log of the exchange rate. Gaussian increments on innovation X_T: Last observed value in the training set.
ARIMA(1,1,0)	Process: $\Delta X_t = \mu + \phi\, \Delta X_{t-1} + \varepsilon_t$ $\varepsilon_t \sim N(0, \sigma^2)$ Forecast: $X_{T+j} \leftarrow X_{T+j-1} + \hat{\mu} + \hat{\phi}(X_{T+j-1} - X_{T+j-2}) + \varepsilon_j$	Autoregressive model in first differences of X_t. Allows X_t to follow a unit root process.
GARCH(1,1)	Process: $\Delta X_t = \mu + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$, $z_t \sim \mathcal{N}(0, 1)$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$	Captures volatility clustering. Stationarity requires $\alpha + \beta < 1$.
Structural Model + AR(1)	Process: $X_t = \boldsymbol{\beta}'\mathbf{z}_t + u_t$ z_t includes $m_{\text{diff},t}, y_{\text{diff},t}, r_{\text{diff},t}$ $u_t = \rho u_{t-1} + \varepsilon_t$ $\varepsilon_t \sim N(0, 1)$	$m_{\text{diff},t}$: log diff of money supply. $y_{\text{diff},t}$: log diff of real output. $r_{\text{diff},t}$: interest rates differenctial. AR(1) captures autocorrelation in the residuals.
Structural Model + AR(1) + GARCH(1,1)	Process: $X_t = \boldsymbol{\beta}'\mathbf{z}_t + u_t$ z_t includes $m_{\text{diff},t}, y_{\text{diff},t}, r_{\text{diff},t}$ $u_t = c + \rho u_{t-1} + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$, $z_t \sim N(0, 1)$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$	Extends the structural model by allowing the innovations ε_t in the error process to exhibit volatility clustering with GARCH
Structural Model + ARMA(1,1) + GARCH(1,1)	Process: $X_t = \boldsymbol{\beta}'\mathbf{z}_t + u_t$ z_t includes $m_{\text{diff},t}, y_{\text{diff},t}, r_{\text{diff},t}$ $u_t = c + \rho u_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ $\varepsilon_t = \sigma_t z_t$, $z_t \sim N(0, 1)$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$	ARMA to capture more complex autocorrelation patterns in u_t.
Efficient Kitchen Sink (E-KS)	Process: $\Delta X_{t+1} = \alpha + \tilde{\mathbf{x}}_t^\top \boldsymbol{\beta} + \mu_{t+1}$ $\mu_t \sim N(0, 1)$ $x_t^{\text{UIP}} = \kappa_h(i_t - i_t^*)$ $x_t^{\text{PPP}} = 100(X_t + p_t^* - p_t)$ $x_t^{\text{MF}} = 100[(m_t - m_t^*) - (y_t - y_t^*) - X_t]$ $x_t^{\text{Taylor}} = 1.5(\pi_t - \pi_t^*) + 0.1(y_t^g - y_t^{g*}) + 0.1x_t^{\text{PPP}}$	A machine learning approach using elastic net regression. The vector $\tilde{\mathbf{x}}_t$ contains standardized versions of four predictors:

Appendix B

Code and Reproducibility

All the analysis, data processing, and model simulations presented in this dissertation were performed using Python. To ensure full transparency and reproducibility of the results, the complete source code, including the main Jupyter Notebook and all data used, is publicly available in a GitHub repository.

The main modelling notebook can be accessed directly at the following link:

`https://github.com/kenjihilasak/exchange-rate-forecasting/blob/master/3.%20modelling/modelling.ipynb`

The full repository is available at:

`https://github.com/kenjihilasak/exchange-rate-forecasting`

Bibliography

Adams, C., & Chadha, B. (1991). On Interpreting the Random Walk and Unit Root in Nominal and Real Exchange Rates. *IMF Staff Papers*, **38**(4), 901-920.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**(3), 307-327.

Byrne, J. P., Korobilis, D., & Ribeiro, P. J. (2016). On the sources of uncertainty in exchange rate prediction. *International Journal of Forecasting*, **32**(2), 525-543.

Çatık, A. N., & Karaca, O. (2020). Revisiting the exchange rate predictability puzzle: A textual analysis and machine learning approach. *Journal of Forecasting*, **39**(7), 1083-1103.

Chang, C. H., Chen, S. L., & Chou, P. H. (2022). Does the Taylor rule model really fail in exchange rate prediction? Evidence from a GMM approach. *The North American Journal of Economics and Finance*, **62**, 101741.

Colombo, V., & Manera, M. (2020). Forecasting the exchange rate: A machine learning approach. *Journal of Forecasting*, **39**(7), 1063-1082.

Cramon, V. (2024). *Applied Forecasting on the Forex Market: A Macroeconomic and Cointegration Approach*. Master's Thesis, Copenhagen Business School.

Engel, C., Mark, N. C., & West, K. D. (2012). Factor model forecasts of exchange rates. *Econometric Reviews*, **34**(1-2), 32-55.

Gong, X., & Lin, B. (2022). Can economic fundamentals affect exchange rates? Evidence from a new copula-MIDAS model. *The North American Journal of Economics and Finance*, **61**, 101684.

Greenaway-McGrevy, R., Mark, N. C., Sul, D., & Wu, J.-L. (2018). Identifying Exchange Rate Common Factors. *International Economic Review*, **59**(4), 1-26.

Grossmann, A., & Chudik, A. (2015). Can economists forecast exchange rates? The debate re-visited. The case of the Swiss franc-Euro. *Aussenwirtschaft*, **66**(1), 81-107.

Hong, Z., Wang, J., & Wang, S. (2023). Forecasting bilateral exchange rates with LSTM recurrent neural networks. *International Journal of Finance & Economics*, **28**(1), 604-622.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice, 2nd edition*. Melbourne, Australia: OTexts.

Li, J., Tsiakas, I., & Wang, W. (2015). Predicting Exchange Rates Out of Sample: Can Economic Fundamentals Beat the Random Walk?. *Journal of Financial Econometrics*, **13**(2), 293-341.

Maveé, N., Perrelli, R., & Schimmelpfennig, A. (2016). *Surprise, Surprise: What Drives the Rand / U.S. Dollar Exchange Rate Volatility?* (IMF Working Paper No. WP/16/205). International Monetary Fund.

Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the 1970s: Do they fit out of sample?. *Journal of International Economics*, **14**(1-2), 3-24.

Somanath, V. S. (1986). Efficient exchange rate forecasts: lagged models better than the random walk. *Journal of International Money and Finance*, **5**(2), 195-220.

Sims, C. A. (1974a). Seasonality in regression. *Journal of the American Statistical Association*, **69**(347), 618-626.