Seminar in Forecasting: Team 4

MRF-ARCH: A MACHINE LEARNING APPROACH TO FORECAST USD/GBP TAIL RISK IN A HIGH DIMENSIONAL DATASET

Sven van Holten Charria
Maurizio Raina

Erasmus School of Economics
Rotterdam, Nederlands

22 April 2024

Supervisor: Prof. Dr. Robin Lumsdaine

Abstract

This paper explores the Macroeconomic Random Forest (MRF) technique, a Machine learning method, to forecast volatility in GBP/USD exchange rates. A novel approach is introduced by integrating MRF time-varying parameters into the ARCH(p) model, resulting in the MRF-ARCH(p) model. Our analysis reveals that the MRF-ARCH(5) produces volatility predictions exhibiting strong performances, especially in times of heightened uncertainty. Furthermore, the model allows for disentangling volatility dynamics by interpreting the MRF-ARCH(1) time-varying parameters.

1 INTRODUCTION

In recent years, machine learning has been tackling increasingly complex problems, such as video recommendations and self-driving cars. Now, we have reached a point where governmental institutions are adapting machine learning to better understand macroeconomic issues, such as forecasting inflation using Quantile Random Forests and understanding the underlying drivers of inflation dynamics (Lenza et al., 2023). Continuing this trend, this paper aims to challenge traditional forecasting methods by adapting a state-of-the-art machine learning method to contribute to the existing literature.

The proposed research examines the forecasting ability of Machine Learning (ML) methods on non-linear macroeconomic data. Goulet Coulombe et al. (2022) studies the reasons driving ML-based forecast models to outperform simple univariate benchmark models when predicting macroeconomic variables in data-rich environments. Four different model features are considered to explore this higher predictive power: non-linearity, regularization, hyperparameter selection, and loss functions. The paper finds that non-parametric non-linearity in ML models is the main driver for this better forecasting and that the superior predictive power of ML techniques is particularly accentuated in situations of high uncertainty, such as during periods of significant macroeconomic volatility and in data-rich environments. In the macroeconomic context, notable non-linearities include structural breaks, smooth transition models and regime switches (Goulet Coulombe, 2024a). Moreover, Medeiros et al. (2021) demonstrates that when handling high-dimensional datasets for inflation forecasting, the Random Forest (RF) method outperforms both benchmark models and other machine learning (ML) techniques. They attribute the model's predictive superiority to its capacity to accommodate flexible non-linear specifications, which is an important trait given the non-linear nature of macroeconomic data.

Literature demonstrates the potential of non-linear ML methods in volatile and non-linear settings with heightened economic uncertainty. Goulet Coulombe et al. (2021) focuses on the non-linearity enveloping the 2020 COVID-19 crisis and its impact on economic activity. The research finds superior predictive abilities by simultaneously considering the data of multiple economically related countries. In addition, the authors construct a monthly mirror of the Federal Reserve Bank of St. Louis (FRED) dataset for the UK, which is employed for the empirical section of this research. The paper concludes that the accuracy of RF models exceeds benchmark ML and

non-ML models, attributing the superiority to the properties of the Random Forest: capturing non-linearities by partitioning a non-linear sample into two smaller linear samples. This splitting enables the model to avoid incorporating irrelevant shocks, such as employment and housing price crashes, into the full model. Instead, it focuses on predictors with higher predictive potential. This research uses this property along with other linearity measures to capture all (non-)linearities in exchange rates, which are inherently influenced by macroeconomic factors.

Machine learning methods do not produce all-round strict improvements over more traditional time-series econometric models. Most importantly, ML models often lack the ability to readily interpret model parameters and understand the roles of variables in accurate prediction, leading to what is commonly referred to as the "black box" paradigm of ML methods. Recently, Goulet Coulombe (2024a) addressed this challenge by introducing the Macroeconomic Random Forest (MRF). In this adaptation of the RF algorithm, the splitting rule is augmented to minimize the Sum of Squared Residuals (SSR) of a pre-specified linear model for each leaf split. In doing this, the fitted linear coefficients are observed for each time period, giving rise to Generalized Time-Varying Parameters (GTVP).

Building strictly upon this machine learning approach to macroeconomic modeling, the primary objective of this research is to extend the capabilities of the MRF model to forecast macroeconomic volatility accurately. This application of machine learning, fitting time-varying parameters to model macroeconomic volatility, has not yet been explored in academic literature. This makes it scientifically relevant to investigate whether the MRF model exhibits forecasting superiority for the latent volatility process, rather than solely focusing on forecasting. Hence, the central research question of the paper is formulated as follows: *"To what extent can the Macroeconomic Random Forest yield accurate dynamic volatility forecasts for USD/GBP exchange rates?"*

The proposed volatility forecasting is implemented through a first-order Autoregressive Conditional Heteroskedasticity (ARCH) model on residuals of the MRF (Engle, 1982). The GTVPs are integrated into the volatility forecast and effectively capture the latent volatility process. Therefore, our contribution to the literature centers on robust volatility estimations achieved by leveraging estimated ARCH parameters for out-of-sample (OOS) volatility forecasting. To evaluate the model's effectiveness, an empirical study is conducted on forecasting USD/GBP exchange rates over a seven-year period, considering two significant non-linear macroeconomic events: Brexit and COVID-19.

The researchers propose naming this novel methodological addition the *MRF-ARCH(p)*. Given the broad scope of the research question, this research focuses on specific aspects of forecasting exchange rate volatility using the MRF-ARCH(p) and its performance with traditional benchmark models:

i *Tail-risk forecasting:* Modeling exchange rate levels and their respective volatility opens the door to estimating tail risk. In this research, tail risk metrics serve as a mean to evaluate the predictive performance of MRF-ARCH(p). Backtesting Value at Risk (VaR) properties and comparing different VaR models via appropriate loss functions indirectly allows us to assess the volatility model.

ii *Time-varying ARCH(p) parameters:* The novel approach of estimating out-of-sample GTVPs using an expanding window presents an opportunity to observe the ARCH parameters at the window margin: each OOS observation is used to estimate an ARCH volatility forecast. Hence, the ARCH-parameters dynamics (ω and α) can be observed over time. Our research further exploits this possibility by studying the volatility dynamics following the Brexit referendum.

Previous literature extensively explores methods for forecasting volatility using different models from the ARCH family. Studies by (Sejnowski & Rosenberg, 1987) and Kristjanpoller and Minutolo (2018) focus on integrating the GARCH model proposed by Bollerslev (1986), with ML techniques, such as Artificial Neural Networks (ANN) and Principal Component Analysis (PCA). This hybrid approach combines GARCH's ability to capture volatility clustering with the nonlinear modeling capacity of the ANN model and the dimension reduction ability of PCA, resulting in superior predictive power. Our research extends this investigation by introducing the MRF-ARCH(p) model for volatility forecasting, with its main advantage lying in its ability to model the ARCH(p) model with time-varying parameters. However, a limiting factor to our research is the unavailability of high-frequency data as a comparison benchmark. Therefore, this research does not utilize long memory models, as suggested by Corsi (2009), to obtain a volatility proxy for the latent process. Instead, the volatility forecasts are assessed against the implied statistical properties of Value-at-Risk (VaR), a tail risk metric. First, the VaR models (and indirectly the volatility forecasts) are assessed through a Likelihood Ratio (LR) test for unconditional and conditional coverage, as well as independence (Christoffersen, 1998). Second,

the models that satisfy these properties are assessed for their performance using a Regulatory Loss Function (RLF) (Lopez, 1998). Finally, the robustness of the MRF-ARCH(p) forecasts is evaluated through an auxiliary experiment that compares the predictive abilities of the model trained on ex-post corrected data and real-time data.

This research paper is structured as follows: Section 2 provides the relevant methodological background necessary to understand the methods used in this research. Appendix A provides necessary mathematical derivations for this section. Section 3 discusses the data to be used in the research, with descriptive statistics tables found in Appendix B. Next, Section 4 describes and explains the methods with accompanying mathematical derivations and definitions. Appendix C provides the necessary mathematical derivations for this section. Section 5 presents the results of the research, while detailed tables and figures on the obtained results are in Appendix D and Appendix E. Finally, Section 6 presents the conclusions drawn and the limitations of the research.

2 THEORETICAL FRAMEWORK

This section starts by providing an overview of the mechanics and methodology behind the MRF model, upon which the MRF-ARCH(p) is based. Next, the transformations from AR(p) to the ARCH(p) volatility model are given. Finally, the tail-risk forecast is discussed.

2.1 Regression Tree

A regression tree is used to split a dataset into smaller subsets based on covariate domain restrictions. The resulting subsets at the end of the trees, also called *leaves*, can be used for estimation purposes. More specifically, a single tree partitions the domains of selected covariates from the full covariate set $s_1, ..., s_p \in \mathbf{S}_t$ such that a single split in covariate domain s_i results in: $\{s_i | s_i \in \mathbb{R}\} \rightarrow \{s_i | s_i \leq c\} \cup \{s_i | s_i > c\}$, for the splitting point $c \in \mathbb{R}$. Each split generates two *children leaves*, where further splits can occur. This process iterates until a stopping criterion is met, producing *terminal leaves*. Each combination of domain sub-spaces within terminal leaves defines a unique, non-overlapping region. Let $R_1, ..., R_k$ represent these regions. Given a dataset $\mathcal{I}$, observations $i \in \mathcal{I}$ are filtered into terminal leaves based on their respective domains. The prediction within each leaf is computed as the average value of the predictive variable: $\hat{y}_{R_s} = E[y_i | i \in R_s]$ (Breiman et al., 1984).

The decision to split covariate s_i at point c within a leaf is made to produce the greatest reduction in the total Sum of Squared Residuals (SSR) in the resulting regions of its children leaves. In mathematical notation, the regression tree solves the optimization problem 2.1 at each non-terminal leaf, creating two children leaves R_1 and R_2.

$$\min_{s \in \mathbf{S}, c \in \mathbb{R}} \left[\sum_{\{i: x_i \in R_1(s,c)\}} (y_i - \hat{y}_{R_1})^2 + \sum_{\{i: x_i \in R_2(s,c)\}} (y_i - \hat{y}_{R_2})^2 \right] \tag{2.1}$$

The splits in the regression tree are determined through a greedy, top-down local optimization function. For every non-terminal leaf in the regression tree, the best possible split is determined conditionally on the split occurring on that leaf, instead of considering other leaves in the tree. Hence, this decision is made locally without considering the global structure of the tree, saving computational time but introducing high levels of variance into regression tree predictions. This method (intentionally) introduces high variance into the creation of regression trees, as each slightly different starting point can result in significantly different terminal leaves. The RF model exploits this characteristic of high variance through the usage of bootstrap aggregation (Medeiros et al., 2021).

2.2 Bootstrap Aggregation and Random Forests

The issue of high prediction variance and susceptibility to over-fitting in regression trees is assessed in Breiman (2001) by incorporating bootstrap aggregation. The resulting RF algorithm entails repeating the following procedure B times and averaging the results: selecting a subset of covariates $s_1, ..., s_q \in \mathbb{S}$, running the regression tree algorithm on this covariate subset, and computing the average prediction for each leaf. This aggregation of regression tree predictions over a random forest of B trees prevents over-fitting and significantly reduces the model's variance. Moreover, this randomization ensures that each tree concentrates on different aspects of the data, yielding significant advantages when working with nonlinear data (Medeiros et al., 2021).

Take, for instance, a (non-linear) regime change, best captured through covariate s_i at time $t = t^*$. Given there is a significant potential improvement in SSR reduction, the RF algorithm would, in most cases, choose to split the domain on s_i at t^*, resulting in leaves with corresponding subdomains $\{s_i | s_i \leq t^*\}$ and $\{s_i | s_i > t^*\}$. Hence, the effect of the non-linearity is reduced in the children leaves. After 'enough' splitting iterations that capture non-linearities, the

resulting leaves are expected to be linear. However, the vanilla RF algorithm fails to capitalize on the resulting linearity, leaving behind valuable information that can be used in prediction. Hence, this research considers a more comprehensive augmented RF model, which combines the strengths of both linear and non-linear models. These strengths include reduced proneness to over-fitting, minimal necessity for hyperparameter tuning, and the ability to handle extensive datasets (Goulet Coulombe et al., 2021).

2.3 Macroeconomic Random Forests

Previous literature introduces a method to capture linearity in the terminal leaves through a *hybrid-linear RF* model, in which Ordinary Least Squares (OLS) regressions are performed on observations within each terminal leaf (Medeiros et al., 2021). A recent methodological improvement implements this same linear model in the splitting rule, enabling the observation of the linear coefficients β_t within different regression trees over time (Goulet Coulombe, 2024a). The proposed general model is defined in Equation 2.2.

$$y = X_t\beta_t + \epsilon_t \text{ , with } \beta_t = \mathcal{F}(\mathbf{S}_t) \tag{2.2}$$

where $\mathcal{F}$ denotes the RF structure, $\mathbf{S}_t$ denotes the full covariate set, and $X_t \subset \mathbf{S}_t$. The general model is integrated into the splitting rule together with a ridge regression penalty parameter. However, the added benefit of the ridge regression is not clear given the low number of carefully chosen covariates included in Equation 2.2. Hence, the ridge penalty is approached cautiously, with the intention of refraining from its usage. These considerations form the basis for the augmented splitting proposed in the optimization problem in Equation 2.3.

$$\min_{j\in\mathcal{J}^-,c\in\mathcal{R}}\left[\min_{\beta_1}\sum_{\{i:x_i\in R_1(j,c)\}}(y_t - X_t\beta_1)^2 + \lambda||\beta_1||_2 + \min_{\beta_2}\sum_{\{i:x_i\in R_2(j,c)\}}(y_i - X_t\beta_2)^2 + \lambda||\beta_2||_2\right] \tag{2.3}$$

where each parent leaf minimizes over the random subset of predictors $\{j_1, ..., j_n\} \in \mathcal{J}^-$ for every bootstrap sample, $J^{-1} \subset \mathbf{S}_t$. The time-varying β_t, defined as the GTVP, allows for observing the coefficients of the general model over time.

2.4 Volatility model derivations

The conditional standard deviation σ_t is referred to as the volatility. A common definition of a volatility model, proposed by Andersen et al. (2006), is defined in Equation 2.4, 2.5, and 2.6.

$$\varepsilon_t = z_t \sigma_t, \quad z_t \overset{\text{iid}}{\sim} D \tag{2.4}$$

$$y_t = \mu_t + \varepsilon_t \tag{2.5}$$

$$\sigma_t^2 = \mathbb{V}[y_t | \mathcal{I}_{t-1}] = \mathbb{V}_{t-1}[y_t] \tag{2.6}$$

where volatility models differ in its specification of the conditional mean μ_t and conditional variance σ_t^2. For instance, the conditional variance of the p-th order ARCH model is $\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2$ (Engle, 1982). Note for shorthand notation the information set at time t is denoted as follows: $\mathbb{E}[X|\mathcal{I}_t] = \mathbb{E}_t[X]$.

This research links the MRF GTVPs to the coefficients in volatility models through the conditional mean $\mu_t = X_t \beta_t$. By applying this conditional mean in Equation 2.5, it becomes possible to define the shocks as: $\varepsilon_t = y_t - X_t \beta_t = z_t \sigma_t$. An important feature of the MRF is its ability to model OOS β_{t+h}, allowing for the fitting of residuals $\hat{\varepsilon}_{t+h} = y_{t+h} - X_t \hat{\beta}_{t+h}$ for the OOS horizon h. Hence, both in-sample and OOS residuals are observed from the MRF estimation, creating an opportunity to forecast volatility through modeling the residuals.

The specifics of the transformation from AR(p) to ARCH(p) is now described. Under the assumption of stability, the ARCH(p) model with intercept for the conditional volatility of a dependent variable can be rewritten as an AR(p) model without intercept for the squared standardized errors in Equation 2.8 (Lawrance, 2023). The stability assumption, equivalent to each of the p (imaginary) roots of the AR(p) polynomial being outside of the unit circle, can be tested with an Augmented Dickey Fuller (ADF) test (Elliott et al., 1996). Under this assumption, the unconditional variance is constant over time and can be expressed by means of a convergent geometric series: $\bar{\sigma}^2 = \omega / (1 - \sum_{i=1}^{p} \alpha_i)$. Furthermore, the ε_t's are assumed to be stable and have a zero-mean. The necessary standardization to obtain an ARCH(p) model from an AR(p) model is as follows (Lawrance, 2023):

$$\tilde{\varepsilon}_t^2 = \varepsilon_t^2 - \bar{\sigma}_{\varepsilon^2}^2 \text{ and } \tilde{z}_t^2 \quad = z_t^2 - 1 \tag{2.7}$$

Using these results, the ARCH(p) can be rewritten:

$$\begin{aligned} \sigma_t^2 &= w + \textstyle\sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 \\ &= \bar{\sigma}_{\varepsilon^2}^2 \left(1 - \textstyle\sum_{i=1}^{p} \alpha_i\right) + \textstyle\sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 \\ &= \bar{\sigma}_{\varepsilon^2}^2 - \textstyle\sum_{i=1}^{p} \alpha_i \bar{\sigma}_{\varepsilon^2}^2 + \textstyle\sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 \end{aligned} \quad \left| \quad \begin{aligned} &= \bar{\sigma}_{\varepsilon^2}^2 + \textstyle\sum_{i=1}^{p} (\alpha_i \varepsilon_{t-i}^2 - \alpha_i \bar{\sigma}_{\varepsilon^2}^2) \\ &= \bar{\sigma}_{\varepsilon^2}^2 + \textstyle\sum_{i=1}^{p} \alpha_i (\varepsilon_{t-i}^2 - \bar{\sigma}_{\varepsilon^2}^2) \\ &= \bar{\sigma}_{\varepsilon^2}^2 + \textstyle\sum_{i=1}^{p} \alpha_i \tilde{\varepsilon}_{t-i}^2 \end{aligned} \right. \tag{2.8}$$

The following algebraic manipulation allows to rewrite the squared residual term at time t in terms of two conditional volatility terms:

$$\varepsilon_t^2 = z_t^2 \sigma_t^2 = \sigma_t^2 \left(1 - 1 + z_t^2\right) = \sigma_t^2 + \sigma_t^2 \left(z_t^2 - 1\right) = \sigma_t^2 + \sigma_t^2 \tilde{z}_t^2 \tag{2.9}$$

Combining the results of Equations 2.7, 2.8 and 2.9, and under the assumption of stability, the AR(p) on the standardized squared errors is shown to be equivalent to the ARCH(p):

$$\tilde{\varepsilon}_t^2 = \sum_{i=1}^{p} \alpha_i \tilde{\varepsilon}_{t-i}^2 + \sigma_t^2 \tilde{z}_t^2 = \sum_{i=1}^{p} \alpha_i \tilde{\varepsilon}_{t-i}^2 + \tilde{v}_t \tag{2.10}$$

This equivalence allows to easily derive the $ARCH(p)$ parameters from the standardized AR(p) estimation:

$$\alpha_{i,ARCH} = \alpha_{i,AR} = \alpha_i \tag{2.11}$$

$$w_{ARCH} = \bar{\sigma}_{\varepsilon^2}^2 \left(1 - \sum_{i=1}^{p} \alpha_i\right) \tag{2.12}$$

Finally, the AR(p) errors $\tilde{v}_t$ are expressed in terms of the ARCH(p) errors z_t through substitution in Equation 2.9:

$$\tilde{v}_t = \sigma_t^2 \tilde{z}_t^2 = \left(\bar{\sigma}_{\varepsilon}^2 + \sum_{i=i}^{p} \alpha_i \tilde{\varepsilon}_{t-i}^2\right) \tilde{z}_t^2 \tag{2.13}$$

If the errors in Equation 2.4 are assumed $z_t \sim \mathcal{N}(0,1)$, then $z_t^2 \sim \chi^2(1)$. Checking the residual distribution allows for model diagnostics on the assumed distribution.

This section concludes with a note on the choice of volatility model. The ARCH(p) model is a special case of the more extensive GARCH(p,q) family, which incorporates an additional (lagged) moving average volatility component (Bollerslev, 1986). In Appendix A, the equivalence of GARCH(p,q) to ARMA(p,q) is derived. Furthermore, under additional assumptions, the ARMA process can be written as a AR(∞). However, this estimation process is cumbersome due to identification issues and the finite approximation of the AR-order, which introduces bias in the

estimated AR coefficients alongside a stricter set of assumptions. In contrast, when estimating the ARCH(p), all coefficients are identified, and there is no bias. This research leaves the resolution of the MRF-GARCH(p,q) and its comparative performance to the MRF-ARCH(p) for future research.

2.5 Value at Risk forecast and assessment

This section introduces the concept of tail risk and the testing framework used to assess volatility forecast models. Tail risk refers to the likelihood of extreme events occurring, quantified in this research through parametric Value at Risk (VaR). VaR is a risk measure used to quantify the maximum loss over a specific time horizon at a given confidence level, frequented as a stress-test metric for market risk in the banking sector (Linsmeier & Pearson, 2000). For the purpose of this research, we define the $VaR_t(1-q,h)$ forecast on the dependent variable at time t, over a horizon of h months for a quantile q, as: $VaR_t(1-q,h) = \mathbb{E}t[y_{t+h,h}] + z_q\hat{\sigma}_{t+h|t}$, where $\mathbb{E}_t[y_{t+h,h}]$ represents the expected value of the dependent variable h periods ahead. The term $z_q\hat{\sigma}_{t+h|t}$ is defined as the product of the estimation of potential loss at a specified confidence level under a specific quantile distribution, and the h-step-ahead volatility forecast.

The VaR definition allows for different specifications for a given level q and horizon h, depending on the error distribution, volatility model, and forecast. These different VaR specifications are evaluated based on the properties of their violations. A VaR violation occurs when the observed value y_{t+h} exceeds the VaR forecast $VaR_t(1-q,h)$ (Christoffersen, 1998). A VaR model is deemed valid if it satisfies the following three properties: i) *unconditional coverage*, where the fraction of VaR violations is q; ii) *independence*, indicating that VaR breaches are independent and identically distributed across the sample; iii) *conditional coverage*, denoting the joint satisfaction of independence and unconditional coverage. These properties can be backtested via Likelihood Ratio (LR) tests. Under the null hypothesis of no property rejection, LR follows a $\chi^2(d)$ distribution, where the degrees of freedom d are 1 for independence and unconditional coverage tests, and 2 for conditional coverage. Rejection of the null hypothesis H_0 in any of these three tests can indicate that the VaR model is mis-specified (Christoffersen, 1998).

Concluding the evaluation of VaR models, the Regulatory Loss Function (RLF) is introduced to quantify their performance: $RLF_t = (y_t - VaR_t(1-q,h))^2$ for each violation at time t, and 0 for no violation (Lopez, 1998). The RLF offers an additional evaluative metric for com-

paring competing VaR models. When there are two valid models, the one that generates more conservative VaR forecasts is preferred.

3 DATA

The dataset employs a mirrored US and UK macroeconomic covariate set to better explain the USD/GBP exchange rate. The merged dataset includes the full FRED-MD database (McCracken & Ng, 2016) and its mirrored UK counterpart, the UK-MD database (Goulet Coulombe et al., 2021), providing monthly data from March 1998 to December 2021 (287 observations). Specifically, the FRED-MD dataset comprises a total of 126 monthly macroeconomic variables, while its UK counterpart includes 110 variables. This dataset has been extensively used in the literature for forecasting macroeconomic variables, such as by RF and MRF models (Goulet Coulombe, 2024a; Goulet Coulombe et al., 2021, 2022; Medeiros et al., 2021). These variables are classified into 8 main groups, covering a broad spectrum of US and UK macroeconomics, providing a comprehensive view of time-varying macroeconomic conditions.

The covariates are transformed according to the original FRED-MD authors' recommendations in McCracken and Ng (2016), and all transformations are documented in Appendix B. Next, the Augmented Dickey Fuller test is employed to test for unit roots and assess the stationarity of time-series data after the recommended transformations. For each co-variate, if the null hypothesis of a unit root is not rejected, it is differenced once. In the given dataset, only one differencing cycle is required. Notably, the dependent variable (USD/GBP exchange rate) is directly sourced from FRED due to the unfavorable log-difference transformation in the FRED-MD dataset. The first-differenced dependent variable is a convenient linear transformation of the original series, and its stationarity property makes it more appropriate for the usage in MRFs general model (Eq. 2.2). Therefore, for our research, the model for the stationary dependent variable is developed using the MRF approach. Appendix B contains comprehensive statistics, such as summary statistics, and the p-values of statistical tests for normality and unit roots, specifically the Jarque-Bera test and the Augmented Dickey-Fuller test, respectively.

An expanding window approach is used due to the limited number of observations and the advantage of training the MRF model on data containing significant non-linearities. For example, training the MRF model on the 2008 recession and creating domain splits in the regression trees to identify various non-linearities in the data has been found to help detect similar future non-

linearities, such as those experienced during the COVID-19 pandemic (Goulet Coulombe et al., 2021). Therefore, to better support the abilities of the MRF-ARCH(p), it is exposed to periods of high USD/GBP exchange rate volatility and non-linearities. Figure 1 displays notable exchange rate fluctuations, such as during the Brexit referendum, after which the GBP depreciated by 15% against the USD following the UK's decision to leave the European Union, resulting in high volatility (Plakandaras et al., 2017). Notably, the graph of the first difference series shown in Figure 2 exhibits similar traits of high volatility during the same periods as Figure 1.

Figure 1: Exhange rate USD/GBP

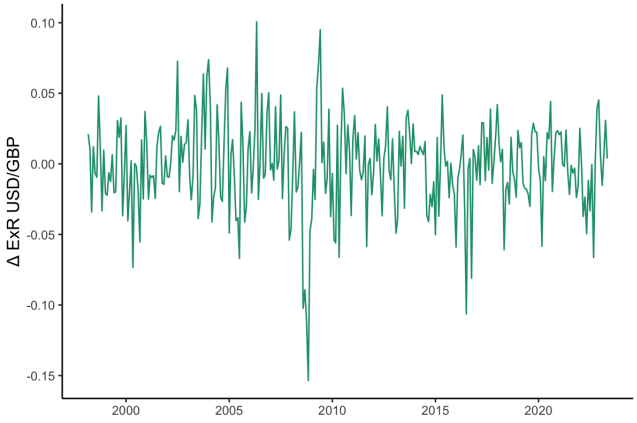


Figure 2: Δ Exchange rate USD/GBP

Lastly, to align with robustness checks conducted in relevant literature, MRF forecasts are executed on *'vintages'* of US FRED-MD data (Medeiros et al., 2021). These vintages, also published by the Federal Reserve Bank of St. Louis, present real-time values of macroeconomic variables at the time of publication with no ex-post corrections (McCracken & Ng, 2016). This approach avoids the use of future corrections to the estimated macroeconomic variables, which are not yet known for each expanding window. The robustness check is carried out using real-time ('vintage') data versus Ex-Post corrected data, both consisting of 134 covariates spanning from December 2014 to January 2022 (86 observations).

4 METHODOLOGY

The methodology follows the three-step MRF method proposed by Chinn et al. (2023), and extends their research by including the ARCH(p) component. Their work highlights two main findings: first, machine learning (ML) methods demonstrate greater accuracy with short samples containing only the most informative predictors; and second, predictive power increases when

variables are pre-selected before factor extraction.

Each component of the three-step method— variable pre-selection, factor extraction, and MRF estimation— is discussed below. The next section outlines the hyperparameter tuning process. The methodology concludes with the addition of this paper to literature: the ARCH estimation and a brief summary of performance metrics.

4.1 Variable pre-selection

In the first step, a co-variate set is constructed through pre-selection using the Least Angle Regression (LARS) algorithm. LARS has been shown to outperform t-statistics and other correlation-based selection algorithms (Efron et al., 2004). The LARS algorithm starts with an empty covariate set (X) and estimates ($\widehat{\mu}$) set to 0, selecting the variable most correlated with the residual vector to build up the estimates $\widehat{\mu} = X\widehat{\beta}$, where $\widehat{\beta}$ represents the vector of covariate estimates. Next, the residuals are recalculated using skewed (equiangular) projection, and the algorithm takes the largest possible step towards the previously selected variable until another covariate has as much correlation with the residual vector as the previously chosen one. For each added covariate to X, the estimate at step k is defined and updated as follows: $\hat{\mu_k} = \hat{\mu_{k-1}} + \hat{\gamma_k} x_k$, where $\hat{\gamma_k}$ represents the length of step k. This process continues until a stopping rule is met. LARS benefits from its low computational cost, as it only requires m steps (where m is the number of covariates used), and its ability to easily calculate the step size. Another advantage of LARS is its capability to select predictors that are partially correlated with others already in the active set.

4.2 Factor extraction

Alongside LARS, the linear dimensionality reduction technique of Principal Component Analysis (PCA) is found to be effective at summarizing and orthogonalizing the various variables due to its simplicity and low computational requirements (Efron et al., 2004). PCA transforms the data into linear and orthogonal combinations of the variables, ensuring that the data within different vectors are uncorrelated. Each of these vectors is referred to as a factor. Moreover, these factors are constructed such that the first factor explains the most variance, followed by subsequent factors in descending order.

In addition, Goulet Coulombe et al. (2022) suggest the inclusion of Moving Average Factors

(MAFs) to extract relevant data from lags through a weighted average across the lagged variables. MAFs can be seen as a smooth index of the corresponding economic indicators, with a function similar to that of moving PCAs. They argue for their inclusion to allow for a greater number of meaningful splits in the MRFs regression trees. MAFs looking back one year are included for similar reasons, and to better capture seasonality. Hence, the covariate set comprises both standard PCAs and MAFs. Finally, the covariate set is lagged with lag k, resulting in the full covariate set $\mathbf{S}_t$.

4.3 Macroeconomic Random Forest

Goulet Coulombe (2024a) finds that the Factor-Augmented Autoregressive linear specification is optimal for capturing both linearity and non-linearity in macroeconomic data across various dates. Since this research emphasizes the MRF-ARCH(p) model rather than MRF forecasts, the same linear specification is assumed in the estimation. The linear equation considers two lags of the dependent variable and a single lag of the two PCA factors with the highest eigenvalues. Therefore, the chosen (initial) linear specification for the general model described in Equation 2.2 is defined in Equation 4.1.

$$y_t = \omega_t + \phi_{1,t} y_{t-1} + \phi_{2,t} y_{t-2} + \gamma_{1,t} F_{1,t} + \gamma_{2,t} F_{2,t} + \epsilon_t \tag{4.1}$$

meaning the GTVPs are defined as $\beta_t = \{\omega_t, \phi_{1,t}, \phi_{2,t}, \gamma_{1,t}, \gamma_{2,t}\}$.

Given the time-dependent nature of macroeconomic data, it's important that observations within the terminal leaves maintain a sequential order to ensures that the linear regression captures time dependencies. For this, the methodology employs *block bootstrapping*, where blocks of a predetermined minimum consecutive length are filtered into the leaves. This length should be optimized through hyperparameter tuning. Preserving this time dependency is especially important to retain the auto-regressive properties captured by the ARCH model (Goulet Coulombe, 2024a).

4.4 Hyperparameter Tuning

The R package 'Macroeconomic Random Forests' (Goulet Coulombe, 2024b) is used to apply the MRF function, and the R package 'rugarch' (Galanos, 2023) is used to apply all (G)ARCH and

ARMA-GARCH models. A seed is set for replicability across MRF iterations. Hyperparameter tuning is conducted to adjust the numerous parameters in the MRF function that cannot be directly inferred from the context of macroeconomic forecasting. Goulet Coulombe (2024a) argues that the gains from hyperparameter tuning in the MRF are minimal and focuses on a small number of variations in the Ridge-lambda (RL) and Random Walk Regularization (RWR). In contrast, Medeiros et al. (2021) take a different approach by brute-forcing many parameters in their Three-Step method, including the number of variables selected by LARS, the number of bootstrap samples, RL, RWR, block size, chosen components in the linear equation, and the number of included factors. Due to computational constraints, our research is limited in its ability to conduct extensive global hyperparameter tuning. Instead, the tuning in this study is achieved using a local optimization method, which involves three consecutive steps, each focusing on a smaller number of parameters.

Some parameters in the MRF function can be inferred from previous research or recommendations from the package author (Goulet Coulombe, 2024b). For instance, *mtry.frac*, which determines the fraction of $\mathbf{S}_t$ considered in each split, is set to 1/3. This low value ensures sufficient randomization across the regression trees, though excessively low values may result in the creation of inefficient trees. Setting *mtry.frac* to 1/3 has been found to be optimal for tuning similar macroeconomic data. Another parameter, *minsize*, serves as a stopping criterion, dictating the minimum number of observations within a leaf to be considered for a further split or classified as a terminal leaf. It is set to 15, as recommended by the author for monthly data. The *block size* length is set to 12, allowing for sufficient block length to capture time-dependent relationships. The number of bootstrap samples (B) is set to 100, as suggested by the author for an adequate mean prediction. While a higher value, such as 200-300, could potentially yield more credible results, it is computationally infeasible and is left for future research. Finally, the number of PCA and MAF factors is each set to 5, in line with previous literature

The hyperparameter tuning is performed on Mean Squared Prediction Error. The first tuning step involves a grid search on different values of variables selected by LARS (30, 40, 50), Random Walk Regularization (RWR) (0, 0.5, 0.95), and Ridge-lambda (RL) (0.1, 0.5). The second tuning step utilizes the best-performing parameters from the first step and further tunes the number of lags (3, 5, 8) and whether to include Moving Average Factors (MAFs). In this step, the number of PCA and MAF factors is each set to 5, aligning with previous literature.

The third step uses the best-performing parameters from the second step and explores additional tuning options for the number of PCA and MAF factors (5, 10, 15).

4.5 MRF-ARCH(p) volatility estimation

This section outlines the procedure for estimating an ARCH model on the residuals based on the time-varying parameter estimates $\hat{\beta}_t$ obtained from the MRF estimation. This novel technique utilizes theoretical results concerning the equivalence of ARCH(p) model exchange rates, conditional volatility, and AR(p) models for the standardized squared shocks under the stability assumption, as demonstrated in Subsection 2.4 (Lawrance, 2023).

Recall that an expanding window is used for forecasting, generating h-steps-ahead forecasts for the dependent variable in each window. This produces a GTVP estimate $\hat{\beta}_t$ of dimension $[1 \text{x} k]$ and a residual series $\varepsilon_t = Y - X\hat{\beta}_t$ of dimension $[t+h \text{x} 1]$. The residual vector is squared and demeaned to obtain $\tilde{\varepsilon}^2$, which forms the basis for the AR(p) to ARCH(p) transformation as detailed in Subsection 2.4. Following this, an AR(p) estimation without intercept is conducted on the $\tilde{\varepsilon}^2$ vector to identify ARCH parameters. After a preliminary Augmented Dickey-Fuller (ADF) test for unit roots, the fitted AR(p) coefficients are used to derive the corresponding ω and α_i estimates of the ARCH(p) model for each expanding window. The outcome is a time-varying estimation of the ω_t and $\alpha_{i,t}$ MRF-ARCH(p) parameters.

The coefficients of the AR(p) model are estimated using the R package '*arima*' (R Core Team, 2024), which fits based on Maximum Likelihood, while the (skewed) t-distributions are fitted using the R package '*fGarch*' (Wuertz et al., 2024). The choice of an appropriate order p is based on the tail-risk framework outlined in Subsection 2.5. To compare different orders of AR processes in terms of their volatility forecasts, the $VaR(h, 1-q)$ models for horizons $h = 1, 2, 3$, autoregressive order $p \in 1, ..., 12$, over 3 different quantile levels q, are assessed. Tail risk analysis typically considers $q \in (0.01, 0.05, 0.1)$. However, in this monthly data application, considering the 1% level is deemed inappropriate due to the out-of-sample size being just 97, which can reduce the power of tests in the VaR assessment framework introduced in Subsection 2.5 (Røynstrand et al., 2012). Therefore, all VaR model assessments consider the quantile values $q \in (0.05, 0.1, 0.15)$.

Furthermore, different quantile distributions other than the $N(0, 1)$ are also considered. According to the literature, a Student-$t(\nu)$ or skewed Student-$t(\nu, \xi)$ distribution is often more

appropriate due to the fatter tails than normal in financial assets (Kuester et al., 2005). The distribution parameters ν, denoting the degrees of freedom, and ξ, denoting the skewness, are fitted on in-sample residual. In summary, the autoregressive order p that satisfies the conditions outlined by Christoffersen et al. (2001) and minimizes the RLF across different quantile levels and distributions is preferred for the MRF-ARCH(p) forecasts. For this, the RLF at time t is defined as: $RLF = (T - T_0 + 1)^{-1} \sum_{t=T_0}^{T} (y_t - VaR_t(1-q,h))^2$.

4.6 Point forecasting: exchange rate change and volatility

The forecasting of the MRF-ARCH(p) model is based on the fitted parameters at time t and their forecasted value at $t+h$ (Francq & Zakoian, 2019). Equation 4.2 presents the h-step ahead forecast for the MRF model. Equation 4.3 presents the h-step ahead forecast for volatility using the MRF-ARCH(p) model, illustrating its recursive nature. Equations C.2, C.3, and C.4 present the explicit h-step ahead specification of the ARCH forecast for fixed horizons 1, 2, and 3, respectively, depending on the chosen order p. Due to the dependence of this recursive formula on the choice of p and h, a general formula cannot be derived for all values of horizon h.

$$\hat{y}_{t+h} = \mathbb{E}(y_{t+h}|I_t, X_t) = X_t \hat{\beta}_{t+h} \tag{4.2}$$

$$\begin{aligned} \hat{\sigma}^2_{t+h|t} &= \mathbb{E}(\sigma^2_{t+h}|I_t) = \mathbb{E}(\hat{\omega}_t + \sum_{i=1}^{p} \hat{\alpha}_i z^2_{t+h-i} \sigma^2_{t+h-i}|I_t) \\ &= \hat{\omega}_t + \sum_{i=1}^{p} \hat{\alpha}_i \mathbb{E}(\sigma^2_{t+h-i}|I_t) = \hat{\omega}_t + \sum_{i=1}^{p} \hat{\alpha}_i \hat{\sigma}^2_{t+h-i|t} \end{aligned} \tag{4.3}$$

4.7 Benchmark Models and Comparison

The MRF model and MRF-ARCH(p) model are compared to forecasting and volatility benchmark models. Subsection 4.7.1 outlines the benchmark choice and comparison measurements used to assess MRF forecasting accuracy. Furthermore, Subsection 4.7.2 introduces the volatility benchmarks and comparison procedure to assess MRF-ARCH(p).

4.7.1 Exchange rate forecast comparison

The predictive accuracy of the MRF model is compared to two different benchmarks. These compared models are used for Out-of-Sample (OOS) forecasting at time t across forecasting

horizons 1, 2 and 3.

First, an AR model of order p is used. For a time series X_t at time t, this model is defined as $X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t$. To select the optimal p, AR models of order 1 up to 10 are built for each forecasting horizon. The AR model that results in the lowest MSPE is chosen for comparison, indicating superior OOS predictive performance. Typical model performance measures such as the Akaike Information Criterion or Bayesian Information Criterion are not used for order selection to prevent model over-fitting, in line with relevant literature (Goulet Coulombe, 2024a; Ledolter & Abraham, 1981). In addition, a Random Walk (RW) model without trend is included for comparison. This model is derived from setting an AR(1) model with a parameter equal to one and no intercept. Forecast values for forecasting horizon h at time t are defined as $\widehat{X}_{t,t+h} = X_t$.

Forecast error metrics are analyzed to assess the OOS predictive accuracy between models. Let T be the total length of the sample, and T_0 be the index of the last observation in the estimation sample, such that $T - T_0 + 1$ is the size of the OOS period. According to standard practice, the Mean Squared Prediction Error (MSPE) is used as a measure of forecast accuracy. For horizon h and forecasted values at time t, the MSPE performance metric is defined as: $MSPE_h = (T - T_0 + 1)^{-1} \sum_{t=T_0}^{T} f_{t,h}^2$. Furthermore, Mean Absolute Error (MAE), defined as $MAE_h = (T - T_0 + 1)^{-1} \sum_{t=T_0}^{T} \mid f_{t,h} \mid$ and Mean Absolute Deviation (MAD), defined as $MAD_{t,h} = median(f_{t,h} - median(f_{t,h}))$, are also considered. MAE is less sensitive to outliers in the data and provides a better interpretation of results, as all errors are equally scaled. Furthermore, MAD uses the median as a central point of reference rather than the mean, making this metric robust to outliers.

The Diebold-Mariano (DM) test is introduced to assess whether the differences in the OOS performance of the MRF model against the AR and RW models are statistically significant. The null hypothesis of this Equal Predictive Ability test is that both compared models have equal performance. Additionally, to account for potential large Type I Errors in horizons greater than one, our research considers a small sample correction, introducing a new test statistic and student-t distributed critical values (Harvey et al., 1997). Therefore, this study uses a 'corrected' one-sided DM test with a squared error loss function.

4.7.2 Volatility forecast comparison

The predictive accuracy of the MRF-ARCH(p) model is compared to three different benchmark models. Firstly, an ARCH(p) model provides a simpler version of the MRF-ARCH(p), where the conditional mean $\mu_t = \mu$ is constant (Engle, 1982). Following Equation 2.5, the residual series on which the ARCH benchmark is modeled is: $\hat{\epsilon}_t = y_t - \mu_{y_t}$. The choice of autoregressive order p follows the same method as outlined for MRF-ARCH(p) in Subsection 4.5. The next benchmark is the GARCH of order *(1-1)*, with that order being commonly evaluated as sufficient and efficient for describing a wide range of time series correlation dynamics (Andersen et al., 2006). From now on, GARCH denotes GARCH(1,1). Finally, the ARMA(1,1)-GARCH(1,1) model further extends the GARCH model by allowing for a time-varying conditional mean (μ_t), modeled by an ARMA(1,1) model. This model is defined as: $\Delta y_t = \sum_{j=1}^{p} \phi_j \Delta y_{t-j} + \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{i-j}$. From now on, ARMA-GARCH denotes ARMA(1,1)-GARCH(1,1). The GARCH model is then estimated on these ε_{i-j}^2 residuals, similarly to the MRF-ARCH(p) model.

To conclude, the implied VaR is assessed and compared across models via the two-step process outlined in Subsections 2.5 and 4.5. A total of 108 VaR models are considered (4 benchmarks, 3 quantile distributions, 3 quantile levels, 3 horizons).

5 Results

The presentation of the results starts with the selection of variables and the extraction of factors in Subsection 5.1, followed by an exploration of the results of hyperparameter tuning in Subsection 5.2. The MRF and MRF-ARCH forecasting is divided into three sections: forecasting exchange rates (Subsection 5.3), forecasting volatility (Subsection 5.4), and the MRF-ARCH(1) volatility GTVP interpretation and comparison to ARCH(1) (Subsection 5.5).

5.1 Variable Pre-selection & Factor Extraction

The pre-selection of covariates for forecasting the GBP/USD exchange rate is conducted using the LARS algorithm over an expanding window from 191 to 286 months. Applying the three-step method alongside the initial hyperparameter tuning yields the outcomes presented in Table D.1. The table indicates that the lowest LARS value (n=30) results in the best outcomes with the lowest relative MSPE across various horizons and hyperparameter iterations. This observation

aligns with existing literature, which suggests the use of LARS for MRF and indicates that selecting approximately 10% of the covariates from the FRED-MD dataset optimizes MSPE (Chinn et al., 2023). In the third iteration, further model improvement is observed for n=28 across a range of better performing hyperparameter settings, leading to the lowest MSPE in more than 60% of cases across different horizons.

Figure E.1 illustrates the frequency with which each covariate is selected by the LARS algorithm across the expanding window. A clear pattern emerges, indicating the algorithm's preference for certain variables, which are consistently chosen as the window expands. These primarily include other exchange rates (JPY, CAD), mortgage rates, and industrial indices. Nine covariates are chosen for every expanding window. In the variable pre-selection process, approximately half of the chosen variables are related to US economic indicators, while the other half are drawn from UK data. This balance reflects equal country macroeconomic contributions in explaining the exchange rate.

5.2 Hyperparameter Tuning

Three iterations of hyperparameter tuning are conducted as specified in the methodology in Subsection 4.4. Table D.1, D.2, and D.3 present the results for the three different hyperparameter tuning iterations, respectively. The most significant results are outlined below.

First, the performance of MRFs is found to be highly sensitive to the configuration of the RL and RWR, within and across horizons. The optimal combination in the first iteration is RL: 0.5, RWR: 0.95. However, given the sensitivity of MRF, this configuration is assumed to perform best overall, across any other configuration of the other hyperparameters. This assumption is likely incorrect, but is maintained for this report, given its minor influence on prediction ability in previous literature (Goulet Coulombe, 2024a). More specifically, Table D.1 shows that for the 1 and 2-month forecasting horizons, an RL and RWR configuration of 0.1 and 0 yields the lowest MSPE. For longer horizons, 3, 4, 6, and 12 months ahead, an RL of 0.5 and RWR of 0.95 yield the lowest MSPE. This trend of higher horizons preferring covariate shrinkage is extended in Table D.2, as higher horizons favor a higher number of RL. Second, including MAFs strictly improves the model's MSPE. This effect is strongest for two and four-month horizons. Third, a higher number of lags (8 lags) is shown to be more beneficial for shorter forecasting horizons (1 and 2 months), while a lower number of lags (3 and 5 lags) is shown to be more beneficial for

longer horizons (4 months). Finally, Table D.3 present how a small number of factors (5 PCA and 5 MAF factors) better performance for horizons 1 and 3, while for horizon 2 including a higher amount of factors (10 PCA and 10 MAF factors) gives the lowest MSPE. Considering all tuning iterations, the following final specification is chosen: (*RL = 0.5, RWR = 0.95, lags = 8, PCA factors = 5, and MAF factors = 5*). This specification is better suited for lower horizons, motivated by the intention to maintain a fair comparison across all horizons, and as a higher weight is placed on lower horizons due to their stronger relevance for ARCH(p) forecasting and financial applications.

To conclude, the variation in optimal RL and RWR settings, along with the number of chosen lags, aligns with econometric theory. This diagnostic suggests that including more lags and variables enables the model to capture immediate trends, which is beneficial for short-term forecasting. Conversely, long-term forecasts prioritize overarching trends and seasonality. Thus, using fewer lags and variables simplifies the model, allowing it to focus on these longer-term patterns and reducing the introduction of noise from excessive data.

Referring back to Subsection 2.3, the tuning results deviate from the original hypothesis of keeping the RL at a low value due to the already parsimoniously constructed *general model*. This observation indicates that each of the covariates is dense in predictive ability, can efficiently dictate splitting rules individually, and therefore is a good choice to include in the general model.

5.3 Exchange Rate Forecast

The forecasting results based on optimal parameter settings are plotted in Figure E.2. An initial screening of Figure E.2a suggests a strong ability of the MRF, fitted using the three-step approach and optimal hyperparameters, to forecast for shorter horizons. On one hand, periods of high volatility appear to be well-captured due to the MRFs strong ability to model non-linear relationships. On the other hand, periods of low volatility also seem to be captured effectively due to the linear specification in the general model of the MRF and its advantage in modeling linear relationships (Goulet Coulombe, 2024a).

The benchmarks introduced in Subsection 4.7.1 include the AR(p) model and the RW model. Following the methodology outlined in the same section, the AR(p) model is found to produce the best results for order 3. The AR(3) model is selected for exchange rate forecasting due to its superior performance in minimizing the MSPE for the short-term horizons (1 and 2), as

outlined in Table D.4. Table D.4 displays that lower-ordered AR models perform better on longer horizons, a pattern consistent with econometric principles. Forecasting further into the future with an AR model increases parameter uncertainty. In the trade-off between efficiency and bias, efficiency becomes more crucial for OOS forecasts compared to the in-sample fit. Therefore, to prevent over-fitting and especially for longer horizons, parsimonious models tend to yield superior performance (Ledolter & Abraham, 1981).

The predictive accuracy of the compared models is assessed using the metrics introduced in Subsection 4.7.1: MSPE, MAE, and MAD. To aid interpretation, these metrics are standardized relative to those of the RW model (Medeiros et al., 2021). Figure 3 displays the standardized metrics for the first four horizons. Across these horizons, the MRF three-step method exhibits substantially lower MSPE, ranging from 83% up to 91% lower than the RW. Additionally, the figure demonstrates lower MAE values, indicating better forecasting performance, and lower MAD values, a measure robust to outliers in the data. The AR(3) benchmark shows lower MSPE, MAE, and MAD values for all horizons compared to the RW model metrics, though it is consistently outperformed by the MRF.

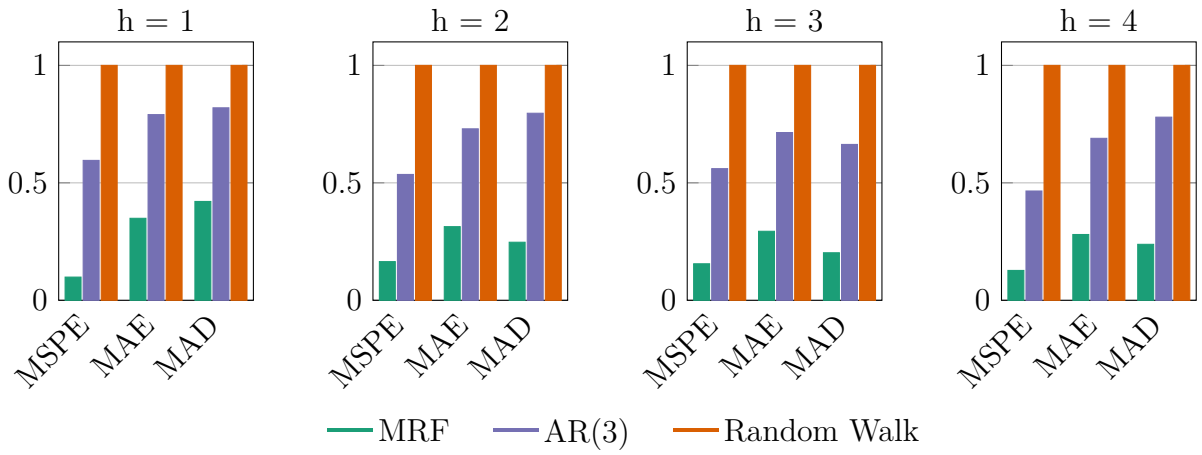


Figure 3: Standardized MSPE, MAE, and MAD for horizons 1, 2, 3, and 4

Notes: Values are standardized for each horizon according to values of the RW.

More formally, these differences in predictive performance are tested using a one-sided DM test with the settings outlined in Subsection 4.7.1. Results for all four considered forecasting horizons are provided in Table 1. MRF significantly outperforms both the AR benchmark and RW benchmark for all four horizons considered. Moreover, AR significantly outperforms the RW benchmark. This finding is consistent with the results of Chinn et al. (2023).

Table 1: DM Test Results for h = 1, 2, 3, and 4.

Horizon 1	MRF	AR	RW
MRF	-	**0.000**	**0.000**
AR	1.000	-	**0.003**
RW	1.000	0.997	-

Horizon 3	MRF	AR	RW
MRF	-	**0.000**	**0.000**
AR	1.000	-	**0.000**
RW	1.000	1.000	-

Horizon 2	MRF	AR	RW
MRF	-	**0.000**	**0.000**
AR	1.000	-	**0.001**
RW	1.000	0.999	-

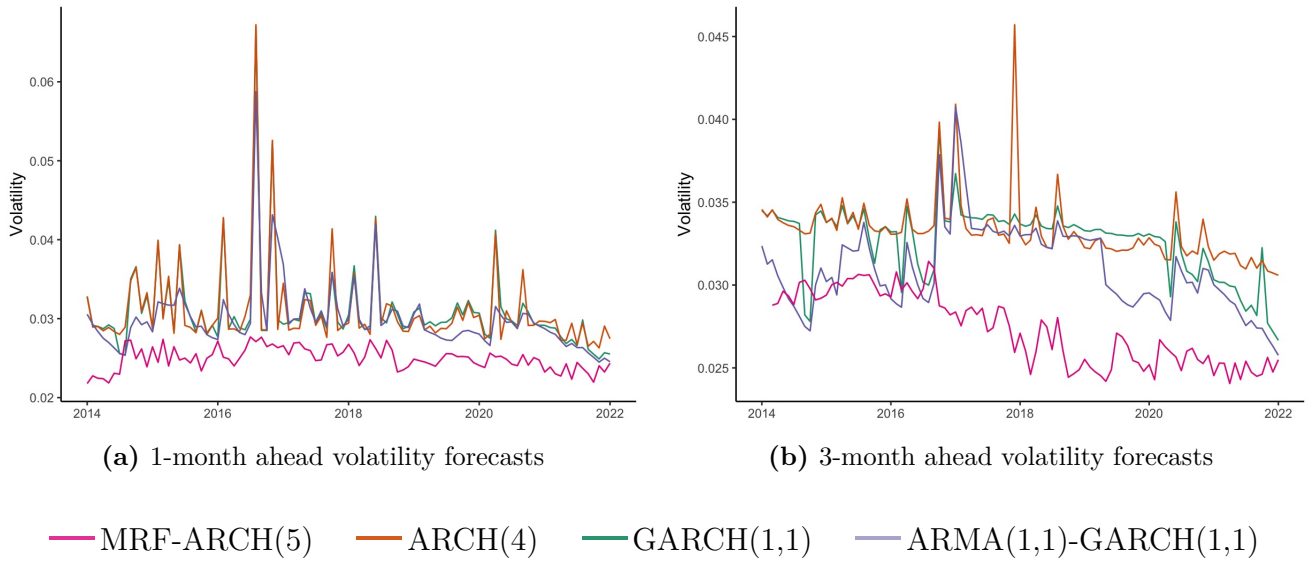
Horizon 4	MRF	AR	RW
MRF	-	**0.000**	**0.000**
AR	1.000	-	**0.000**
RW	1.000	1.000	-

Notes: Result P-values for one sided DM-test with small sample correction suggested in Harvey et al. (1997). P-values in bold indicate significance at a 5% level.

5.4 Volatility forecast

The volatility forecast uses the residual vector resulting from the MRF specifications outlined in Subsection 5.2 as the input. Following the procedure detailed in Subsection 4.5, the ARCH(p) is estimated using an AR(p). Tables D.5, D.6, and D.7 report the VaR model results for 1, 2, and 3-month ahead forecasts respectively. Across all horizons, quantile levels, and quantile distributions considered, MRF-ARCH(5) outperforms the other autoregressive orders p and is chosen as the 'best autoregressive order'. From now on, MRF-ARCH denotes MRF-ARCH(5). An analogous procedure for the ARCH(p) benchmark is reported in Tables D.8, D.9, and D.10. In this case, an ARCH(4) benchmark is chosen and is denoted as ARCH.

Now the focus shifts to comparing the predictive volatility performance of the MRF-ARCH relative to the benchmark models introduced in Subsection 4.7.2: the ARCH model, the GARCH model, and the ARMA-GARCH model. Figure 4 presents a comparison of the volatility forecasts of the four models for horizons 1 and 3 (horizon 2 is shown in Figure E.3). Across all horizons, MRF-ARCH exhibits lower volatility forecasts. This can be attributed to the model's construction. The benchmarks only rely on lagged values of exchange rates as data inputs. In contrast, MRF-ARCH incorporates a large set of pre-processed and non-linearly modeled covariates (Goulet Coulombe et al., 2021). As a result, the residuals used for estimating volatility in MRF-ARCH are of a lower magnitude, as less of the exchange rate change remains unexplained by the model. Not only do the specifications of the conditional mean and volatility differ, but also the information set $\mathcal{I}_t$ on which they are conditioned.



(a) 1-month ahead volatility forecasts

(b) 3-month ahead volatility forecasts

Figure 4: Volatility Forecasts of MRF-ARCH, ARCH, GARCH, and ARMA-GARCH for h = 1 and 3

The volatility forecasts for both the ARCH and GARCH models show significantly higher levels for the shorter-term horizon. This pattern can be attributed to the nature of these models, which, under stationarity, emphasize short-term fluctuations and are more sensitive to immediate shocks. In contrast, the ARMA-GARCH volatility forecast shares some similarities with the MRF-ARCH model, albeit with consistently larger magnitude changes across horizons. The sharp fluctuations of the ARMA-GARCH model suggest a high susceptibility to shocks in the underlying series through the ARMA specification for the unconditional mean. This time-varying conditional mean allows for more complex (partial) auto-correlation dynamics in exchange rates to be captured in the volatility forecast. Moving to higher horizons, the ARCH and GARCH specifications lose their volatile structure and trend towards their unconditional mean, resulting in flatter forecasts. This is due to the low estimates of the volatility parameters in these models, causing the h-step-ahead forecast to rapidly and exponentially approach its unconditional mean. This trend becomes apparent when assuming stability and expressing the ARCH and GARCH forecasts as deviations from the unconditional variance (Andersen et al., 2006). Looking at the graphs, the MRF-ARCH model presents a more stable forecast that is less susceptible to extreme shocks, due to the MRFs strengths in non-linear modeling and the pre-processing of the high-dimensional covariate set.

Table 2: RLF for MRF-ARCH, ARCH, GARCH and ARMA-GARCH VaR models for quantile level = 0.05, 0.1, and 0.15.

Quantile level	Horizon	Quantile distribution	MRF-ARCH	ARCH	GARCH	ARMA-GARCH
0.05	1	normal	0.003952	**0.002536**	0.002902	0.003017
		t	0.004118	**0.003020**	0.003454	0.003343
		skewed-t	0.003691	**0.002548**	0.002916	0.002687
	2	normal	0.003221	**0.001869**	0.002238	0.002366
		t	0.003373	**0.002301**	0.002731	0.002673
		skewed-t	0.002985	**0.001880**	0.002251	0.002057
	3	normal	0.003108	**0.001155**	0.001412	0.001793
		t	0.003253	**0.001536**	0.001822	0.002081
		skewed-t	0.002883	**0.001165**	0.001423	0.001507
0.1	1	normal	0.007989	**0.005709**	0.006384	0.006530
		t	0.008670	**0.007324**	0.008113	0.007698
		skewed-t	0.008324	**0.006856**	0.007619	0.007053
	2	normal	0.006996	**0.004895**	0.005477	0.005861
		t	0.007641	**0.006517**	0.007125	0.007018
		skewed-t	0.007313	**0.006052**	0.006656	0.006380
	3	normal	0.010938	**0.003982**	0.004315	0.005133
		t	0.012059	**0.005577**	0.005908	0.006275
		skewed-t	0.011925	**0.005116**	0.005450	0.005645
0.15	1	normal	NP	**0.000043**	0.010070	0.010264
		t	NP	NP	NP	0.012006
		skewed-t	NP	**0.000043**	0.012318	0.011644
	2	normal	0.010938	**0.008399**	0.009013	0.009593
		t	0.012059	**0.010881**	0.011496	0.011349
		skewed-t	0.011925	**0.010639**	0.011256	0.010986
	3	normal	NP	**0.007441**	0.007760	0.008833
		t	0.011860	**0.009922**	0.010233	0.010598
		skewed-t	0.011725	**0.009680**	0.009991	0.010233

Notes: 'NP' signifies that either the conditional coverage, unconditional coverage, or independence of the Var model is violated. Lowest RLF values for each horizon, quantile distribution and quantile level are marked in bold.

The focus now shifts to assessing the volatility forecasts in terms of the implied VaR. The VaR forecast for the MRF-ARCH model for a quantile level of 0.05, 0.10, and 0.15 is shown in Figure E.4. The results of the 2-step VaR assessment procedure are reported in Table 2. Based on the minimization loss function, the ARCH consistently outperforms the other benchmarks and MRF-ARCH across different quantile levels, distributions, and horizons. To assess the ARCH VaR performance in more detail and across quantile levels, the horizon is fixed to one-month-ahead and the quantile distribution is fixed to normal. The ARCH forecast VaR for a quantile level of 0.1 is showed in Figure 5. Results for quantile level 0.05 and 0.15 are shown in Figure E.5. In line with the volatility forecasts observed in Figure 4, the ARCH VaR model exhibits more volatility and spiked edges. Conversely, the MRF-ARCH produces a more stable VaR forecast, albeit with less conservative thresholds, resulting in worse RLF minimization performance compared to ARCH.

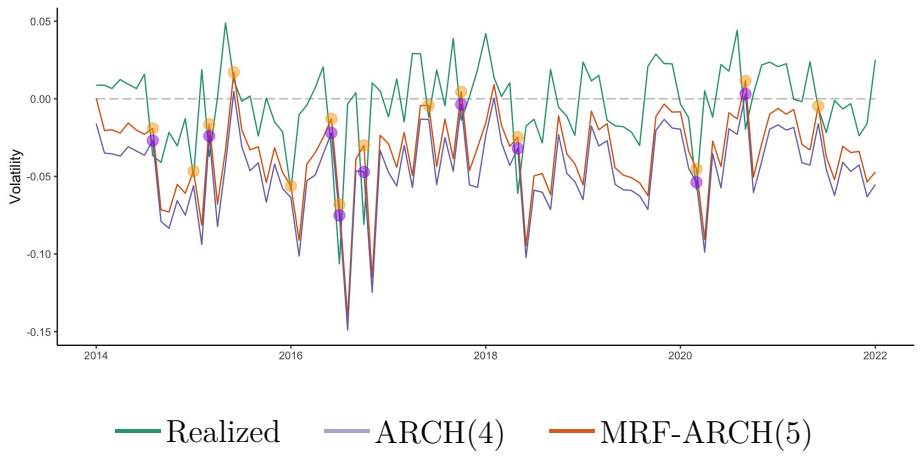
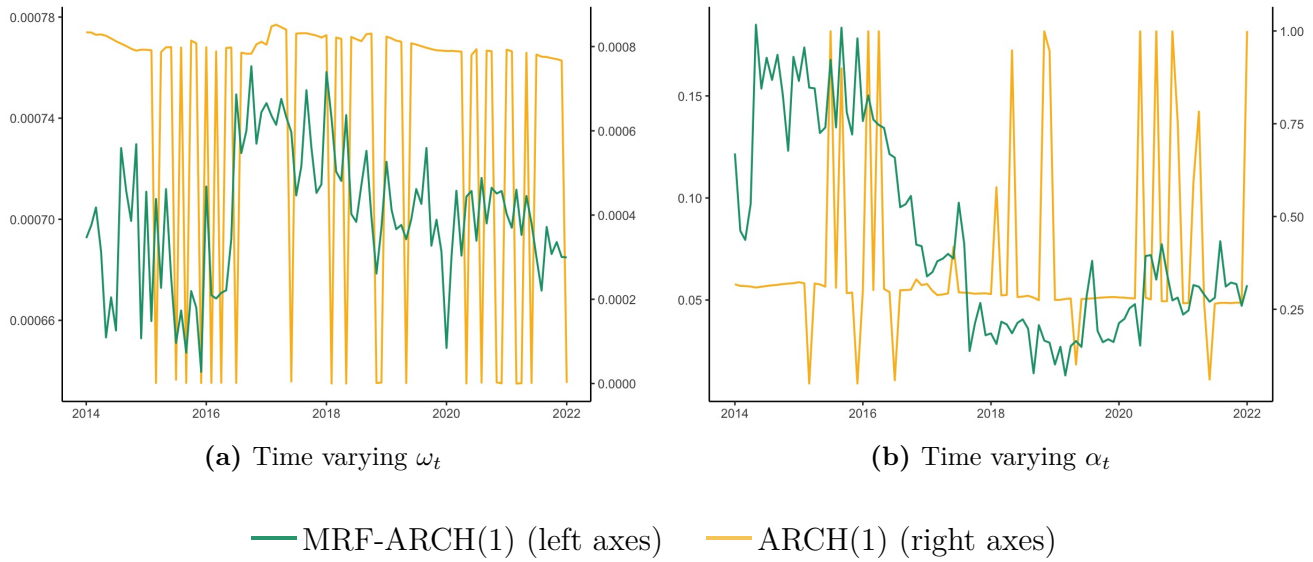


Figure 5: MRF-ARCH and ARCH VaR using normally distributed quantiles for q = 0.1 and h = 1

Figure E.6 compares the 1-month ahead VaR forecast of MRF-ARCH with the other two benchmarks (GARCH and ARMA-GARCH) for quantile levels of 0.05, 0.10 and 0.15 and a standard normal distribution for the z_t quantile. As expected from the 2-step VaR assessment, all volatility models appear to provide valid VaR estimates. Consistent with the volatility forecasts in Figure 4, the MRF-ARCH produces less volatile VaRs than all competing models.

5.5 Time-varying ARCH parameters

The primary exploratory benefit of the proposed MRF-ARCH methodological framework is the opportunity to observe the time-varying ARCH parameters. These findings are illustrated using the MRF-ARCH(1) mode. This low-order model limits the number of volatility parameters to 2, providing a more intuitive interpretation. The first parameter, the intercept ω_t, represents the baseline volatility at time t. The second parameter, α_t, represents the persistence of past shocks. These parameters are then compared to those of a simple ARCH(1) model.



(a) Time varying ω_t **(b)** Time varying α_t

Figure 6: Time-Varying Parameters of ARCH(1) and MRF-ARCH(1) model for h = 1.

Figure 6 presents the comparison between the ω_t and α_t parameters of the MRF-ARCH(1) and ARCH(1) models. The MRF-ARCH(1) model yields a smooth function for both parameters. In contrast, the ARCH(1) model shows erratic and abrupt shifts in these parameters throughout the time series, likely due to identification issues, making its interpretation more challenging. The MRF-ARCH(1) model proposed in this study offers a more sensitive and smooth distribution of volatility shifts between the two parameters over time. Specifically, the figure highlights highly volatile periods following events like the Brexit referendum (23 June 2016) and the COVID-19 crisis, which are selected for further analysis.

During the period following the Brexit referendum, the level of ω_t increased, indicating elevated baseline volatility. This suggests an immediate reaction of financial markets to the potential outcomes of the referendum, leading to increased uncertainty and subsequent higher volatility. In contrast, the terms related to shocks in the previous month, α_t, gradually decreased from the time of the referendum until the making of the withdrawal agreement that took effect from February 2020. As negotiations and plans for Brexit became clearer, the market's reaction to new information became less reactive. This resulted in a reduction in the impact of subsequent shocks on volatility, emphasizing a larger relative baseline level ω_t compared to α_t.

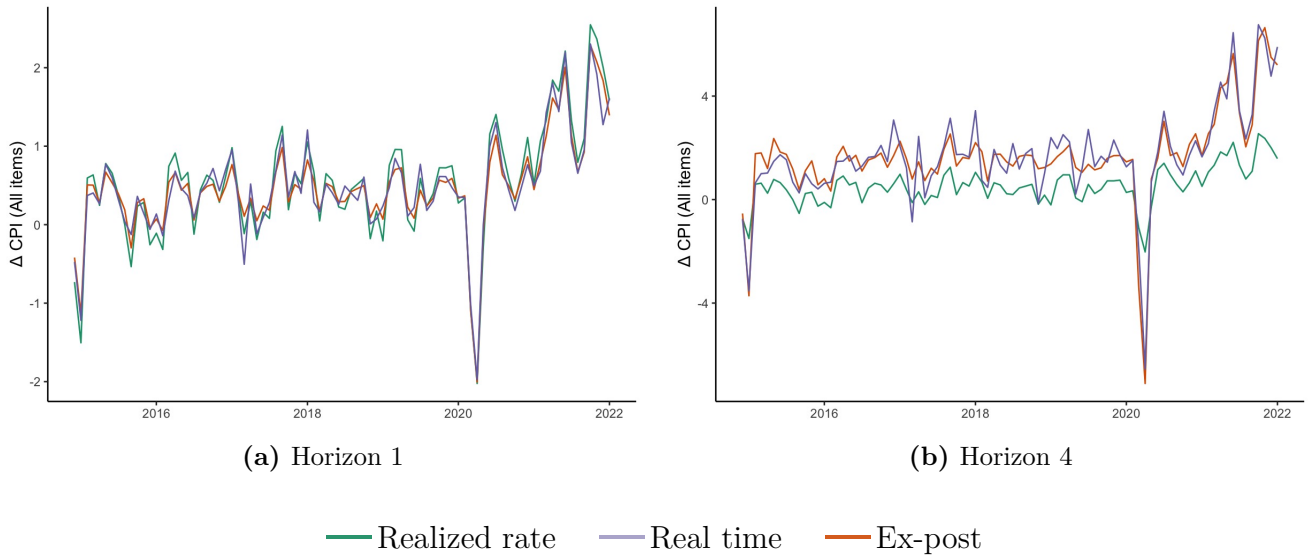
However, the COVID-19 pandemic presented a different scenario, where exchange rate volatility remained stable despite the global economic disruption. This stability can be attributed to the symmetric nature of the shock on both the UK and US economies. The simultaneous dis-

ruptions caused by the pandemic and the similar expansionary monetary and fiscal responses led to a balanced effect on exchange rates, maintaining lower volatility. The Figure shows a spike in ω_t observed in Figure 6a, dropping significantly for March 2020, the initial month of COVID-19 economic turmoil for the US and UK economies. This sudden drop in baseline volatility could potentially be explained by extreme uncertainty and a temporary collapse in global trade due to the restrictions in place to contain the pandemic, which may have caused foreign exchange market participants to freeze their trading activities. Alternatively, this unexpected result could be attributed to the MRF model's robust ability to capture non-linearities during times of unprecedented shocks in macroeconomic variables, a strength highlighted in recent literature on volatility (Goulet Coulombe et al., 2021). Therefore, the MRF model appears to efficiently incorporate the full information set $\mathcal{I}_t$.

5.6 Robustness Check: Real time data

A robustness check is performed on our research results by comparing the performance of the MRF model on two different datasets. The first dataset is the FRED-MD dataset used in our research and the second is a Real-Time dataset outlined in Section 3, containing only data available at the time of the forecast. The Real-Time dataset is not available for the UK dataset compiled in Goulet Coulombe et al. (2021). To assess the performance of both models without data inequalities, the robustness check is conducted using monthly predictions of inflation, following the approach of Medeiros et al. (2021), for which no UK data is required.

Figure 7 presents a comparison of the forecasting results for Horizons 1 and 4. For Horizon 1, 8 lags are included, and for Horizon 4, 5 lags are included, as this setting minimizes the MSPE. The figure shows that for both considered horizons, both the Ex-Post and Real-Time data yield similar forecasts. The individual model performance at forecasting monthly inflation values is not assessed as the purpose of this section is to check comparative performance of the model for different datasets.

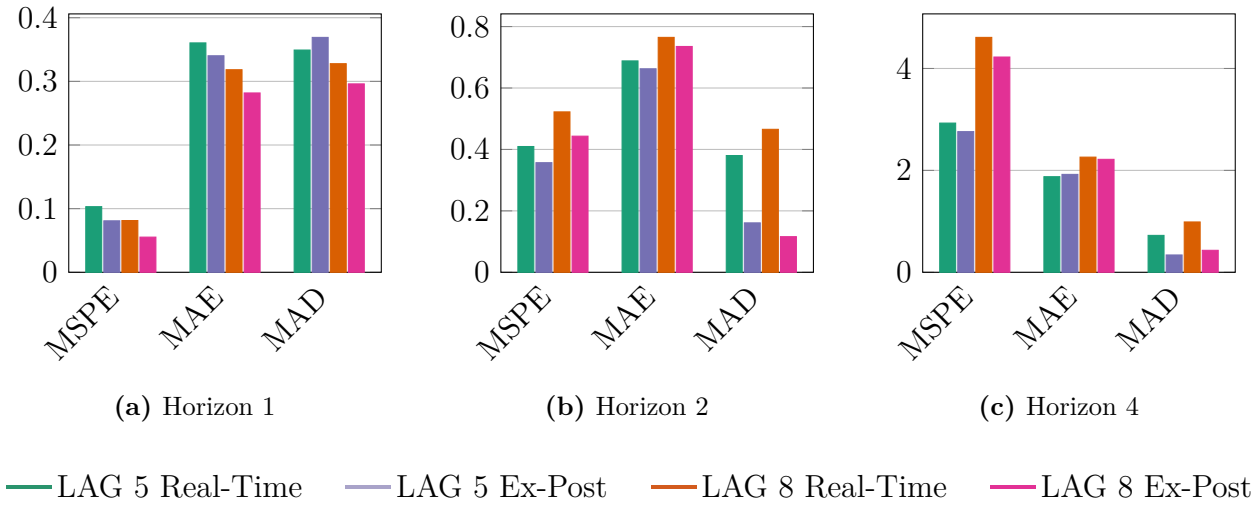


(a) Horizon 1 **(b)** Horizon 4

Figure 7: MRF forecast with Real-Time data and Ex-Post data for h = 1 and 4

Figure 8 displays the results of prediction performance measures for both datasets using different lag settings (5 and 8) for Horizons 1, 2, and 4. For all considered horizons, the MSPE, MAE, and MAD are generally lower for the Ex-Post data compared to those for the Real-Time data with the same number of lags. This suggests a better prediction performance of the MRF in forecasting monthly inflation when using the Ex-Post data. This difference could be due to the Ex-Post containing adjusted and, therefore, more accurate data than the Real-Time dataset.

More formally, a one-sided DM-test with the small sample correction proposed by Harvey et al. (1997) and a squared error loss function is used to evaluate the statistical significance of the differences presented in Figure 8. The results are detailed in Table D.11. Across all forecasting horizons and various numbers of included lags, there are no significant differences in predictive performance between Ex-Post and Real-Time data. The only scenario where Ex-Post data significantly outperforms Real-Time data is when 8 and 5 lags are respectively included for 4-month-ahead forecasts. Therefore, it can be argued that the research methodology is robust to real-time data, and choosing ex-post data does not significantly affect the external validity of the results.



(a) Horizon 1 **(b)** Horizon 2 **(c)** Horizon 4

Figure 8: MSPE, MAE, and MAE for horizon = 1, 2 and 4 and lags = 5 and 8.

Notes: Values are standardized for each horizon according to the ones of a RW performed in the Vintages data.

6 Conclusion

This research investigates the forecasting capabilities of the MRF model for exchange rate volatility, with a specific focus on the first differenced USD/GBP nominal exchange rate. The study follows a three-step methodology proposed by Chinn et al. (2023) to construct the MRF model. In addition to contributing to the existing literature, the study introduces the volatility forecasting model, MRF-ARCH, which is evaluated using tail-risk metrics. The primary finding highlights the interpretability of time-varying MRF-ARCH(1) volatility parameters.

The results demonstrate that the MRF model significantly outperforms the considered benchmarks in making one-month-ahead predictions for the USD/GBP exchange rate. The model exhibits a predictive accuracy improvement ranging between 83% and 91% compared to a Random Walk (RW) model based on MSPE. The robustness of these results is confirmed for longer horizons, extending up to 4 months ahead. Our findings align with those of Goulet Coulombe et al. (2022), indicating that accounting for non-linearities in macroeconomic data significantly enhances predictive accuracy beyond traditional linear models. Furthermore, these results remain robust even when using a 'vintages' dataset based on information available in real time.

A second contribution of this research is the integration of time-varying ARCH parameters. Modeling the MRF-ARCH(1) allows for a nuanced understanding of volatility shifts over time,

distinguishing between the persistence of past shocks and the baseline volatility across the time dimension. Moreover, the forecast sample and dependent variable enable the assessment of the impacts of the Brexit referendum and the COVID-19 pandemic, showing how the time-varying parameters shift during periods of high non-linearity and macroeconomic uncertainty. By addressing identification issues in ARCH(1) and using of a high-dimensional dataset, the MRF-ARCH model demonstrates its relevance as a potential addition for macroeconomic analysis. However, it is important to note that when assessed in terms of RLF minimization, the model is outperformed by a simpler ARCH(4).

This research leaves the opportunity for further exploration into the application of machine learning models in volatility forecasting. Future studies could extend the MRF-ARCH by incorporating a GARCH volatility model specification with time-varying parameters. This MRF-GARCH model has the potential to improve the model's capability to predict more intricate volatility patterns in higher-frequency data, providing interpretable parameters for the MRF-GARCH(1,1). Additionally, the models proposed in this research are based exclusively on monthly macroeconomic observations. The inclusion of daily data could improve the modeling of exchange rate volatility, enabling the examination of how new macroeconomic information releases affect both the volatility level and the time-varying parameters of the volatility model. This inclusion would also significantly increase the sample size, making the testing procedure more powerful (Christoffersen et al., 2001; Røynstrand et al., 2012). To further analyze the performance measures of MRF-based volatility models, other VaR loss functions besides the RLF could be looked at, as the RLF may be perceived as having a conservative bias in the selection of the volatility model (Lopez, 1998; Sarma et al., 2003).

7 Acknowledgement

We wish to express our gratitude to several individuals who have significantly contributed to the progress of this report. First and foremost, we extend our sincere appreciation to our supervisor, Prof. Dr. Lumsdaine, for her guidance and constructive feedback during our weekly sessions. Her comments and suggestions, particularly regarding volatility estimation and hyperparameter tuning, have been critical in making this research possible. Additionally, we would like to acknowledge the authors of Chinn et al. (2023) and Goulet Coulombe (2024a) for generously making the code public with their papers. Their work has served as a reference to build on for developing our code in this novel field. Furthermore, we are grateful for Goulet Coulombe's efforts in creating a publicly available FRED-like UK dataset, which has greatly enhanced the applicability of our research to exchange rates (Goulet Coulombe et al., 2021).

References

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2006). Chapter 15 volatility and correlation forecasting. In *Handbook of economic forecasting* (pp. 777–878, Vol. 1). Elsevier.

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

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Taylor & Francis.

Chinn, M. D., Meunier, B., & Stumpner, S. (2023, June). *Nowcasting world trade with machine learning: A three-step approach* (Working Paper No. 31419). National Bureau of Economic Research.

Christoffersen, P. (1998). Evaluating interval forecasts. *International Economic Review*, *39*(4), 841–862.

Christoffersen, P., Hahn, J., & Inoue, A. (2001). Testing and comparing Value-at-Risk measures. *Journal of Empirical Finance*, *8*(3), 325–342.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174–196.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, *32*(2), 407–499.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, *64*(4), 813–836.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1007.

Francq, C., & Zakoian, J.-M. (2019). *GARCH models: Structure, statistical inference and financial applications*. John Wiley & Sons.

Galanos, A. (2023). *Rugarch: Univariate garch models.* [R package version 1.5-1.].

Goulet Coulombe, P. (2024a). The macroeconomy as a random forest. *Journal of Applied Econometrics*, *39*(3), 401–421.

Goulet Coulombe, P. (2024b). *R package macrorf* [R package version 0.1.1].

Goulet Coulombe, P., Leroux, M., Stevanovic, D., & Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, *37*(5), 920–964.

Goulet Coulombe, P., Marcellino, M., & Stevanović, D. (2021). Can machine learning catch the COVID-19 recession? *National Institute Economic Review*, *256*, 71–109.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*(2), 281–291.

Kristjanpoller, W., & Minutolo, M. C. (2018). A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis. *Expert Systems with Applications*, *109*, 1–11.

Kuester, K., Mittnik, S., & Paolella, M. S. (2005). Value-at-Risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, *4*(1), 53–89.

Lawrance, T. (2023). ARMA and GARCH Connections. *University of Warwick*. Retrieved April 11, 2024, from https://personalpages.manchester.ac.uk/staff/georgi.boshnakov/events/Priestley/Lawrance_slides.pdf

Ledolter, J., & Abraham, B. (1981). Parsimony and its importance in time series forecasting. *Technometrics*, *23*(4), 411–414.

Lenza, M., Moutachaker, I., & Paredes, J. (2023). Forecasting euro area inflation with machine-learning models. Retrieved April 14, 2024, from https://www.ecb.europa.eu/press/research-publications/resbull/2023/html/ecb.rb231017~b910853393.en.html

Linsmeier, T. J., & Pearson, N. D. (2000). Value at Risk. *Financial analysts journal*, *56*(2), 47–67.

Lopez, J. (1998). Testing your risk tests. *Financial Survey*, *20*(3), 18–20.

McCracken, M. W., & Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, *34*(4), 574–589.

Medeiros, M. C., Vasconcelos, G. F., Veiga, Á., & Zilberman, E. (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*, *39*(1), 98–119.

Plakandaras, V., Gupta, R., & Wohar, M. E. (2017). The depreciation of the pound post-Brexit: Could it have been predicted? *Finance Research Letters*, *21*, 206–213.

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved April 22, 2024, from https://www.R-project.org/

Røynstrand, T., Nordbø, N. P., & Strat, V. K. (2012). *Evaluating power of Value-at-Risk backtests* [Master's thesis, Institutt for industriell økonomi og teknologiledelse].

Sarma, M., Thomas, S., & Shah, A. (2003). Selection of Value-at-Risk models. *Journal of Forecasting*, *22*(4), 337–358.

Sejnowski, T. J., & Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems*, *1*(1), 145–168.

Wuertz, D., Chalabi, Y., Setz, T., Maechler, M., & Boshnakov, G. N. (2024). *Fgarch: Rmetrics - autoregressive conditional heteroskedastic modelling* [R package version 4033.92].

A GARCH(1,1), ARMA(1,1), AR(∞) models: equivalence and parameter transformations

Definition of conditional volatility in volatility models is given by Equation A.1, Equation A.2, and Equation A.3,

$$y_t = \mu_t + \epsilon_t \tag{A.1}$$

$$\sigma_t^2 = \mathbb{V}(y_t | I_{t-1}) \tag{A.2}$$

$$\epsilon_t = z_t \sigma_t \tag{A.3}$$

where z_t is a *i.i.d.* standard normal series, and I_{t-1} is the information set available one period before. This allows to make volatility one period ahead forecasts with the information available today. Note that in case the MRF is employed for estimating and forecasting y_t on regressors $X_t \subset S_t$, then $\mu_t = X_t \beta_t$.

The specification of the GARCH(1,1) in Equation A.4 defines Equation A.2 to model for interest rates volatility.

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 \tag{A.4}$$

By adding $\epsilon_t^2 - \sigma_t^2 = \nu_t$ on both sides of Equation A.4 you can derive an ARMA(1,1) model from the GARCH(1,1) model. The procedure to obtain the ARMA(1,1) model for the squared innovation term ϵ_t^2 is shown in Equation A.5. Furthermore note that $\epsilon_{t-1}^2 - \nu_{t-1} = \sigma_{t-1}^2$.

$$\epsilon_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 + \nu_t = \omega + \alpha \epsilon_{t-1}^2 + \gamma \epsilon_{t-1}^2 - \gamma \nu_{t-1} + \nu_t = \omega + (\alpha + \gamma) \epsilon_{t-1}^2 - \gamma \nu_{t-1} + \nu_t \tag{A.5}$$

Hence, by defining $\phi = \alpha + \gamma$, and $\theta = -\gamma$, the GARCH(1,1) on y_t in Equation A.5 can be rewritten as an ARMA(1,1) for the ϵ_t^2 in Equation A.4. The transformation can also be done the other way around from ARMA to GARCH by considering the inverse parameter transformations: $\alpha = \phi + \theta$, and $\gamma = -\theta$.

$$\epsilon_t^2 = \omega + \phi \epsilon_{t-1}^2 + \theta \nu_{t-1} + \nu_t \ , with \ \nu_t = \epsilon_t^2 - \sigma_t^2 \tag{A.6}$$

$$\epsilon_t^2 = \omega + \pi_1 \epsilon_{t-1}^2 + \pi_2 \epsilon_{t-2}^2 + \pi_3 \epsilon_{t-3}^2 + \pi_4 \epsilon_{t-4}^2 + \ \ldots \ + \nu_t a \tag{A.7}$$

An important property of the ARMA(1,1) model in Equation A.6 is being equivalent to the AR(∞) model in Equation A.7 when the MA component is invertible (i.e. when $|\theta| < 1$). Note that the invertibility can be tested via a Augmented Dickey Fuller test to not impose the restriction $|\theta| < 1$ on the AR(∞) parameters. This result and the corresponding parameter transformations can be derived by using the lag AR and MA polynomials as follows:

Equation A.6 can be rewritten into A.8 employing lag polynomials notations, where $\Phi_1(L) = 1L^0 - \phi L^1 = 1 - \phi L$, and $\Theta_1(L) = 1L^0 + \theta L^1 = 1 + \theta L$.

$$\Phi_1(L)\epsilon_t^2 = \Theta_1(L)\nu_t \tag{A.8}$$

Assuming $|\theta| < 1$, the MA component of the ARMA(1,1) is invertible and the ARMA process is equivalent to an AR(∞) with AR polynomial $\Phi_\infty(L) = 1 - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \pi_4 L^4 - ....$ This is defined in Equation A.9, Equation A.10, and A.11.

$$\Phi_\infty(L) = \frac{\Phi_1(L)}{\Theta_1(L)} \tag{A.9}$$

$$\implies \Phi_\infty(L)\Theta_1(L) = \Phi_1(L) \tag{A.10}$$

$$\implies (1 - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \pi_4 L^4 - ...)(1 + \theta L) = 1 - \phi L \tag{A.11}$$

Equation A.11 can be solved Lag-by-Lag to express the AR(∞) coefficients in terms of ARMA(1,1) parameters, as shown in Equation A.12.

$$\implies \begin{cases} \phi = \pi_1 - \theta \implies \pi_1 = \phi + \theta \\ 0 = \pi_2 + \pi_1\theta \implies \pi_2 = -\theta\pi_1 \\ 0 = \pi_3 + \pi_2\theta \implies \pi_3 = -\theta\pi_2 \\ ... \\ \pi_k = -\theta\pi_{k-1} = (-\theta)^{k-1}(\phi + \theta) \end{cases} \tag{A.12}$$

We can now, in Equation A.13 inverse the parameter transformation to express ARMA(1,1), coefficients in terms of AR(∞) parameters.

$$\implies \begin{cases} \theta = -\frac{\pi_k}{\pi_{k-1}} = -\frac{\pi_2}{\pi_1} \\ \phi = \pi_1 - \theta = \pi_1 + \frac{\pi_2}{\pi_1} \end{cases} \tag{A.13}$$

If $|\theta| < 1$; using Equation A.12, and the ARMA to GARCH parameter transformations, the GARCH(1,1) parameters can be derived in terms of the AR(∞) ones. This is provided in Equation A.14.

$$\implies \begin{cases} \gamma = -\theta = \frac{\pi_2}{\pi_1} \\ \alpha = \phi + \theta = \pi_1 \end{cases} \tag{A.14}$$

Finally the derivation for the constant term. Let ω denote the intercept of the GARCH(1,1) specification in Equation A.4. From the derivation of the model equivalence in Equation A.14 it follows that the intercept of the ARMA(1,1) for ϵ_t^2 is also ω. However for the ARMA(1,1)-AR(∞) equivalence the intercept has to be converted. Let ω_∞ denotes the intercept in the AR(∞) specification. Then introducing the lag polynomial, if the MA component is invertible, Equation A.8 can be rewritten as in the process followed in Equation A.15, Equation A.16, and Equation A.17.

$$\epsilon_t^2(1 - \phi L) = \omega + \nu_t(1 + \theta L) \tag{A.15}$$

$$\implies \epsilon_t^2 \Phi_\infty(L) = \frac{\omega}{1 + \theta L} + \nu_t \tag{A.16}$$

$$\implies \omega_\infty = \frac{\omega}{1 + \theta L} = \left(\frac{1 + \theta L}{\omega}\right)^{-1} = \left(\frac{1}{\omega} + \theta L\left(\frac{1}{\omega}\right)\right)^{-1} = \left(\frac{1 + \theta}{\omega}\right)^{-1} = \frac{\omega}{1 + \theta} \tag{A.17}$$

Therefore the GARCH(1,1), and ARMA(1,1) can be computed by inverting this identity as: $\omega = (1 + \theta)\omega_\infty = (1 - \frac{\pi_2}{\pi_1})\omega_\infty$. In the derivation: $L\frac{1}{\omega} = \frac{1}{\omega}$, as the lag operator on a constant returns the constant itself.

B Data Tables

Table B.1: Summary Statistics for Group 1: USA Output and Income

Transformation	Fred	Description	Mean	SD	JB	ADF
$\Delta \log(x_t)$	RPI	Real Personal Income	0.002	0.017	>0.000	0.01
$\Delta \log(x_t)$	W875RX1	Real personal income ex transfer receipts	0.002	0.007	>0.000	0.01
$\Delta \log(x_t)$	INDPRO	IP Index	0.001	0.011	>0.000	0.01
$\Delta \log(x_t)$	IPFPNSS	IP: Final Products and Nonindustrial Supplies	0.001	0.012	>0.000	0.01
$\Delta \log(x_t)$	IPFINAL	IP: Final Products (Market Group)	0.001	0.013	>0.000	0.01
$\Delta \log(x_t)$	IPCONGD	IP: Consumer Goods	0.000	0.012	>0.000	0.01
$\Delta \log(x_t)$	IPDCONGD	IP: Durable Consumer Goods	0.001	0.043	>0.000	0.01
$\Delta \log(x_t)$	IPNCONGD	IP: Nondurable Consumer Goods	0.000	0.008	>0.000	0.01
$\Delta \log(x_t)$	IPBUSEQ	IP: Business Equipment	0.002	0.021	>0.000	0.01
$\Delta \log(x_t)$	IPMAT	IP: Materials	0.002	0.012	>0.000	0.01
$\Delta \log(x_t)$	IPDMAT	IP: Durable Materials	0.002	0.017	>0.000	0.01
$\Delta \log(x_t)$	IPNMAT	IP: Nondurable Materials	0.000	0.014	>0.000	0.01
$\Delta \log(x_t)$	IPMANSICS	IP: Manufacturing (SIC)	0.001	0.013	>0.000	0.01
$\Delta \log(x_t)$	IPB51222S	IP: Residential Utilities	0.001	0.022	>0.000	0.01
	IPFUELS	IP: Fuels	0.001	0.023	>0.000	0.01
Δx_t	CUMFNS	Capacity Utilization: Manufacturing	-0.018	0.921	>0.000	0.01

Table B.2: Summary Statistics for Group 2: Stock Market

Transformation	Fred	Description	Mean	SD	JB	ADF
$\Delta \log(x_t)$	S&P 500	S&P's Common Stock Price Index: Composite	0.005	0.039	>0.000	0.01
$\Delta \log(x_t)$	S&P: indust	S&P's Common Stock Price Index: Industrials	0.005	0.038	>0.000	0.01
Δx_t	S&P div yield	S&P's Composite Common Stock: Dividend Yield	0	0.082	>0.000	0.01
$\Delta \log(x_t)$	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio	0	0.057	>0.000	0.01

Table B.3: Summary Statistics for Group 3: USA Labor Market

Transformation	Fred	Description	Mean	SD	JB	ADF
Δx_t	HWI	Help-Wanted Index for United States	15.531	284.317	>0.000	0.01
Δx_t	HWIURATIO	Ratio of Help Wanted/No. Unemployed	0.003	0.068	>0.000	0.01
$\Delta \log(x_t)$	CLF16OV	Civilian Labor Force	0.001	0.003	>0.000	0.01
$\Delta \log(x_t)$	CE16OV	Civilian Employment	0.001	0.01	>0.000	0.01
Δx_t	UNRATE	Civilian Unemployment Rate	-0.003	0.654	>0.000	0.01
Δx_t	UEMPMEAN	Average Duration of Unemployment (Weeks)	0.019	1.175	>0.000	0.01
$\Delta \log(x_t)$	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	-0.001	0.137	>0.000	0.01
$\Delta \log(x_t)$	UEMP5TO14	Civilians Unemployed for 5-14 Weeks	0	0.12	>0.000	0.01
$\Delta \log(x_t)$	UEMP15OV	Civilians Unemployed - 15 Weeks & Over	0	0.078	>0.000	0.01
$\Delta \log(x_t)$	UEMP15T26	Civilians Unemployed for 15-26 Weeks	0	0.122	>0.000	0.01
$\Delta \log(x_t)$	UEMP27OV	Civilians Unemployed for 27 Weeks and Over	0.001	0.071	>0.000	0.01
$\Delta \log(x_t)$	CLAIMSx	Initial Claims	-0.001	0.161	>0.000	0.01
$\Delta \log(x_t)$	PAYEMS	All Employees: Total nonfarm	0.001	0.009	>0.000	0.01
$\Delta \log(x_t)$	USGOOD	All Employees: Goods-Producing Industries	0	0.008	>0.000	0.01
$\Delta \log(x_t)$	CES1021000001	All Employees: Mining and Logging: Mining	0	0.011	>0.000	0.01
$\Delta \log(x_t)$	USCONS	All Employees: Construction	0.001	0.011	>0.000	0.01
$\Delta \log(x_t)$	MANEMP	All Employees: Manufacturing	-0.001	0.007	>0.000	0.01
$\Delta \log(x_t)$	DMANEMP	All Employees: Durable goods	-0.001	0.009	>0.000	0.01
$\Delta \log(x_t)$	NDMANEMP	All Employees: Nondurable goods	-0.001	0.006	>0.000	0.01
$\Delta \log(x_t)$	SRVPRD	All Employees: Service-Providing Industries M	0.001	0.009	>0.000	0.01
$\Delta \log(x_t)$	USTPU	All Employees: Trade, Transportation & Utilities M	0	0.008	>0.000	0.01
$\Delta \log(x_t)$	USWTRADE	All Employees: Wholesale Trade M	0	0.005	>0.000	0.01
$\Delta \log(x_t)$	USTRADE	All Employees: Retail Trade M	0	0.01	>0.000	0.01
$\Delta \log(x_t)$	USFIRE	All Employees: Financial Activities M	0.001	0.003	>0.000	0.01
$\Delta \log(x_t)$	USGOVT	All Employees: Government M	0	0.004	>0.000	0.01
	CES0600000007	Avg Weekly Hours : Goods-Producing M	-0.002	0.26	>0.000	0.01
Δx_t	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing M	-0.005	0.127	>0.000	0.01
	AWHMAN	Avg Weekly Hours : Manufacturing M	-0.003	0.245	>0.000	0.01
$\Delta^2 \log(x_t)$	CES0600000008	Avg Hourly Earnings : Goods-Producing M	0	0.003	>0.000	0.01
$\Delta^2 \log(x_t)$	CES2000000008	Avg Hourly Earnings : Construction M	0	0.006	>0.000	0.01
$\Delta^2 \log(x_t)$	CES3000000008	Avg Hourly Earnings : Manufacturing M	0	0.003	>0.000	0.01

Table B.4: Summary Statistics for Group 4: USA Consumption and Orders

Transformation	Fred	Description	Mean	SD	JB	ADF
$\Delta \log(x_t)$	HOUST	Housing Starts: Total New Privately Owned	0	0.082	0.132	0.01
$\log(x_t)$	HOUSTNE	Housing Starts, Northeast	-0.002	0.268	0.003	0.01
$\log(x_t)$	HOUSTMW	Housing Starts, Midwest	-0.001	0.208	>0.000	0.01
$\Delta \log(x_t)$	HOUSTS	Housing Starts, South	0.001	0.105	0.063	0.01
$\Delta \log(x_t)$	HOUSTW	Housing Starts, West	0	0.148	>0.000	0.01
$\Delta \log(x_t)$	PERMIT	New Private Housing Permits (SAAR)	0	0.054	>0.000	0.01
$\log(x_t)$	PERMITNE	New Private Housing Permits, Northeast (SAAR)	-0.001	0.192	>0.000	0.01
$\Delta \log(x_t)$	PERMITMW	New Private Housing Permits, Midwest (SAAR)	-0.002	0.093	0.184	0.01
$\Delta \log(x_t)$	PERMITS	New Private Housing Permits, South (SAAR)	0.001	0.063	0.871	0.01
$\Delta \log(x_t)$	PERMITW	New Private Housing Permits, West (SAAR)	0	0.095	>0.000	0.01

Table B.5: Summary Statistics for Group 5: USA Orders and Inventories

Transformation	Fred	Description	Mean	SD	JB	ADF
$\Delta \log(x_t)$	DPCERA3M086SBEA	Real personal consumption expenditures	0.002	0.011	>0.000	0.01
$\Delta \log(x_t)$	CMRMTSPLx	Real Manu. and Trade Industries Sales	0.001	0.013	>0.000	0.01
$\Delta \log(x_t)$	RETAILx	Retail and Food Services Sales	0.004	0.019	>0.000	0.01
$\Delta \log(x_t)$	ACOGNO	New Orders for Consumer Goods	0.002	0.025	>0.000	0.01
$\Delta \log(x_t)$	AMDMNOx	New Orders for Durable Goods	0.002	0.046	>0.000	0.01
$\Delta \log(x_t)$	ANDENOx	New Orders for Nondefense Capital Goods	0.001	0.092	>0.000	0.01
$\Delta \log(x_t)$	AMDMUOx	Unfilled Orders for Durable Goods	0	0.01	>0.000	0.01
$\Delta \log(x_t)$	BUSINVx	Total Business Inventories	0.003	0.006	>0.000	0.01
Δx_t	ISRATIOx	Total Business: Inventories to Sales Ratio	0	0.025	>0.000	0.01

Table B.6: Summary Statistics for Group 6: USA Money and Credit

Transformation	Fred	Description	Mean	SD	JB	ADF
$\Delta^2 \log(x_t)$	M1SL	M1 Money Stock	0	0.095	>0.000	0.01
$\Delta^2 \log(x_t)$	M2SL	M2 Money Stock	0	0.005	>0.000	0.01
$\Delta \log(x_t)$	M2REAL	Real M2 Money Stock	0.003	0.007	>0.000	0.01
$\Delta^2 \log(x_t)$	BOGMBASE	Monetary Base	0	0.031	>0.000	0.01
$\Delta^2 \log(x_t)$	TOTRESNS	Total Reserves of Depository Institutions	0	0.093	>0.000	0.01
$\Delta \left(\frac{x_t}{x_{t-1}} - 1.0\right)$	NONBORRES	Reserves Of Depository Institutions	0	1.568	>0.000	0.01
$\Delta^2 \log(x_t)$	BUSLOANS	Commercial and Industrial Loans	0	0.011	>0.000	0.01
$\Delta^2 \log(x_t)$	REALLN	Real Estate Loans at All Commercial Banks	0	0.007	>0.000	0.01
$\Delta^2 \log(x_t)$	NONREVSL	Total Nonrevolving Credit	0	0.008	>0.000	0.01
Δx_t	CONSPI	Nonrevolving consumer credit to Personal Income	0	0.003	>0.000	0.01
$\Delta^2 \log(x_t)$	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding	0	0.032	>0.000	0.01
$\Delta^2 \log(x_t)$	DTCTHFNM	Total Consumer Loans and Leases Outstanding	0	0.033	>0.000	0.01
$\Delta^2 \log(x_t)$	INVEST	Securities in Bank Credit at All Commercial Banks	0	0.011	>0.000	0.01

Table B.7: Summary Statistics for Group 7: Interest rate and Exchange Rates

Transformation	Fred	Description	Mean	SD	JB	ADF
Δx_t	FEDFUNDS	Effective Federal Funds Rate	-0.001	0.186	>0.000	0.01
Δx_t	CP3Mx	3-Month AA Financial Commercial Paper Rate	0.002	0.199	>0.000	0.01
Δx_t	TB3MS	3-Month Treasury Bill	0	0.198	>0.000	0.01
Δx_t	TB6MS	6-Month Treasury Bill	0	0.190	>0.000	0.01
Δx_t	GS1	1-Year Treasury Rate	-0.001	0.196	>0.000	0.01
Δx_t	GS5	5-Year Treasury Rate	-0.006	0.227	0.003	0.01
Δx_t	GS10	10-Year Treasury Rate	-0.007	0.216	>0.000	0.01
Δx_t	AAA	Moody's Seasoned Aaa Corporate Bond Yield	-0.007	0.183	>0.000	0.01
Δx_t	BAA	Moody's Seasoned Baa Corporate Bond Yield	-0.005	0.210	>0.000	0.01
	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS	0.151	0.248	>0.000	0.01
	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS	-0.141	0.275	>0.000	0.01
	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS	-0.044	0.327	>0.000	0.01
	T1YFFM	1-Year Treasury C Minus FEDFUNDS	0.114	0.398	>0.000	0.039
Δx_t	T5YFFM	5-Year Treasury C Minus FEDFUNDS	-0.005	0.256	>0.000	0.01
Δx_t	T10YFFM	10-Year Treasury C Minus FEDFUNDS	-0.005	0.263	>0.000	0.01
Δx_t	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS	-0.005	0.259	>0.000	0.01
Δx_t	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS	-0.003	0.302	>0.000	0.01
$\Delta \log(x_t)$	TWEXAFEGSMTHx	Trade Weighted U.S. Dollar Index: Major Currencies	>0.000	0.016	0.016	0.01
$\Delta \log(x_t)$	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	-0.002	0.023	>0.000	0.01
$\Delta \log(x_t)$	EXJPUSx	Japan / U.S. Foreign Exchange Rate	0	0.024	>0.000	0.01
$\Delta \log(x_t)$	EXUSUKx	U.S. / U.K. Foreign Exchange Rate	-0.001	0.033	>0.000	0.01
$\Delta \log(x_t)$	EXCAUSx	Canada / U.S. Foreign Exchange Rate	0	0.018	>0.000	0.01

Table B.8: Summary Statistics for Group 8: USA Prices

Transformation	Fred	Description	Mean	SD	JB	ADF
$\Delta^2 \log(x_t)$	OILPRICEx	Crude Oil, spliced WTI and Cushing	0	0.127	>0.000	0.01
$\Delta^2 \log(x_t)$	PPICMM	PPI: Metals and metal products	0	0.042	>0.000	0.01
$\Delta^2 \log(x_t)$	CPIAUCSL	CPI : All Items	0	0.003	>0.000	0.01
$\Delta^2 \log(x_t)$	CPIAPPSL	CPI : Apparel	0	0.007	>0.000	0.01
$\Delta^2 \log(x_t)$	CPITRNSL	CPI : Transportation	0	0.017	>0.000	0.01
$\Delta^2 \log(x_t)$	CPIMEDSL	CPI : Medical Care	0	0.002	>0.000	0.01
$\Delta^2 \log(x_t)$	CUSR0000SAC	CPI : Commodities	0	0.007	>0.000	0.01
$\Delta^2 \log(x_t)$	CUSR0000SAD	CPI : Durables	0	0.004	>0.000	0.01
$\Delta^2 \log(x_t)$	CUSR0000SAS	CPI : Services	0	0.001	>0.000	0.01
$\Delta^2 \log(x_t)$	CPIULFSL	CPI : All Items Less Food	0	0.004	>0.000	0.01
$\Delta^2 \log(x_t)$	CUSR0000SA0L2	CPI : All items less shelter	0	0.005	>0.000	0.01
$\Delta^2 \log(x_t)$	CUSR0000SA0L5	CPI : All items less medical care	0	0.003	>0.000	0.01
$\Delta^2 \log(x_t)$	PCEPI	Personal Cons. Expend.: Chain Index	0	0.002	>0.000	0.01
$\Delta^2 \log(x_t)$	DDURRG3M086SBEA	Personal Cons. Exp: Durable goods	0	0.003	>0.000	0.01
$\Delta^2 \log(x_t)$	DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods	0	0.008	>0.000	0.01
$\Delta^2 \log(x_t)$	DSERRG3M086SBEA	Personal Cons. Exp: Services	0	0.002	>0.000	0.01

Table B.9: Summary Statistics for Group 9: UK Labour Market

Transformation	Name	Mean	SD	JB	ADF
$\Delta \log(x_t)$	EMP	0,001	0,002	>0.000	0,01
$\Delta \log(x_t)$	EMP_PART	0,001	0,005	>0.000	0,01
$\Delta \log(x_t)$	EMP_TEMP	0	0,013	0,722	0,01
Δx_t	UNEMP_RATE	-0,007	0,102	>0.000	0,01
$\Delta \log(x_t)$	UNEMP_DURA_6mth	0	0,024	>0.000	0,01
$\Delta \log(x_t)$	UNEMP_DURA_6.12mth	-0,001	0,05	>0.000	0,01
$\Delta \log(x_t)$	UNEMP_DURA_12mth.	-0,002	0,031	>0.000	0,01
$\Delta \log(x_t)$	UNEMP_DURA_24mth.	-0,002	0,042	>0.000	0,01
$\Delta \log(x_t)$	EMP_RATE	0,013	0,127	>0.000	0,01
$\Delta \log(x_t)$	EMP_ACT	0,001	0,001	>0.000	0,01
Δx_t	EMP_ACT_RATE	0,008	0,107	0,327	0,01
$\Delta \log(x_t)$	CLAIMS	0	0,039	>0.000	0,01
Δx_t	CLAIMS_RATE	-0,002	0,173	>0.000	0,01
$\Delta \log(x_t)$	TOT_WEEK_HRS	0,001	0,009	>0.000	0,01
$\Delta \log(x_t)$	AVG_WEEK_HRS	0	0,008	>0.000	0,01
$\Delta \log(x_t)$	AVG_WEEK_HRS_FULL	0	0,008	>0.000	0,01
$\Delta \log(x_t)$	AWE_ALL	0,003	0,009	>0.000	0,01
$\Delta \log(x_t)$	AWE_CONS	0,003	0,015	>0.000	0,01
$\Delta \log(x_t)$	AWE_MANU	0,002	0,005	>0.000	0,01
$\Delta \log(x_t)$	AWE_PRIV	0,003	0,004	>0.000	0,01
$\Delta \log(x_t)$	AWE_PUB	0,003	0,01	>0.000	0,01
$\Delta \log(x_t)$	AWE_SERV	0,003	0,01	>0.000	0,01
$\Delta \log(x_t)$	VAC_TOT	0,002	0,04	>0.000	0,01
$\Delta \log(x_t)$	VAC_CONS	0,001	0,087	>0.000	0,01
$\Delta \log(x_t)$	VAC_MANU	-0,001	0,049	>0.000	0,01

Table B.10: Summary Statistics for Group 10: UK Production

Transformation	Name	Mean	SD	JB	ADF
$\Delta \log(x_t)$	IOP_PROD	0	0,017	>0.000	0,01
$\Delta \log(x_t)$	IOP_CAP_GOOD	0,002	0,035	>0.000	0,01
$\Delta \log(x_t)$	IOP_DUR	0,005	0,038	>0.000	0,01
$\Delta \log(x_t)$	IOP_ENER	-0,003	0,024	>0.000	0,01
$\Delta \log(x_t)$	IOP_GOOD	0,002	0,021	>0.000	0,01
$\Delta \log(x_t)$	IOP_INT_GOOD	0	0,026	>0.000	0,01
$\Delta \log(x_t)$	IOP_MACH	0	0,055	>0.000	0,01
$\Delta \log(x_t)$	IOP_MANU	0,001	0,022	>0.000	0,01
$\Delta \log(x_t)$	IOP_MINE	-0,006	0,047	>0.000	0,01
$\Delta \log(x_t)$	IOP_NON_DUR	0,001	0,023	>0.000	0,01
$\Delta \log(x_t)$	IOP_PETRO	-0,002	0,068	0,001	0,01
$\Delta \log(x_t)$	IOP_OIL_EXTRACT	-0,007	0,058	>0.000	0,01

Table B.11: Summary Statistics for Group 11: UK Retail and Services

Transformation	Name	Mean	SD	JB	ADF
$\Delta \log(x_t)$	IOS	0,002	0,016	>0.000	0,01
$\Delta \log(x_t)$	IOS_45	0	0,177	>0.000	0,01
$\Delta \log(x_t)$	IOS_46	0,002	0,037	>0.000	0,01
$\Delta \log(x_t)$	IOS_47	0	0,025	>0.000	0,01
$\Delta \log(x_t)$	IOS_G	0,001	0,032	>0.000	0,01
$\Delta \log(x_t)$	IOS_EDUC	0,001	0,021	>0.000	0,01
$\Delta \log(x_t)$	IOS_PNDS	0,002	0,013	>0.000	0,01
$\Delta \log(x_t)$	RSI	0,002	0,02	>0.000	0,01
$\Delta \log(x_t)$	CAR_REGIS	-0,001	0,335	>0.000	0,01
$\Delta \log(x_t)$	RETAIL_TRADE_INDEX	0,002	0,018	>0.000	0,01
$\Delta \log(x_t)$	AVG_WEEK_RETAIL_SALE	0,003	0,02	>0.000	0,01
$\Delta \log(x_t)$	AVG_WEEK_RETAIL_SALE_NON_FOOD	0,002	0,051	>0.000	0,01

Table B.12: Summary Statistics for Group 12: UK Consumer and Retail Price Indices

Transformation	Name	Mean	SD	JB	ADF
$\Delta \log(x_t)$	CPIH_ALL	0	0,003	>0.000	0,01
$\Delta \log(x_t)$	CPI_ALL	0	0,002	>0.000	0,01
$\Delta \log(x_t)$	CPI_EX_ENER	0	0,002	>0.000	0,01
$\Delta \log(x_t)$	CPI_GOOD	0,001	0,004	>0.000	0,01
$\Delta \log(x_t)$	CPI_DUR	0	0,004	>0.000	0,01
$\Delta \log(x_t)$	CPI_NON_DUR	0	0,004	>0.000	0,01
$\Delta \log(x_t)$	CPI_SERV	0	0,003	>0.000	0,01
$\Delta \log(x_t)$	CPI_CLOTH	-0,002	0,005	>0.000	0,01
$\Delta \log(x_t)$	CPI_TRANS	0,002	0,007	0,209	0,01
$\Delta \log(x_t)$	RPI_ALL	0,003	0,003	>0.000	0,01
$\Delta \log(x_t)$	RPI_GOOD	0	0,003	>0.000	0,01
$\Delta \log(x_t)$	RPI_SERV	0	0,008	>0.000	0,01
$\Delta \log(x_t)$	RPI_HOUSE	0,003	0,005	>0.000	0,01

Table B.13: Summary Statistics for Group 13: UK International Trade

Transformation	Name	Mean	SD	JB	ADF
$\Delta \log(x_t)$	EXP_TOT	0,002	0,029	>0.000	0,01
$\Delta \log(x_t)$	EXP_GOOD	0,001	0,049	>0.000	0,01
$\Delta \log(x_t)$	IMP_ALL	0,003	0,035	>0.000	0,01
$\Delta \log(x_t)$	IMP_GOOD	0,003	0,046	>0.000	0,01
$\Delta \log(x_t)$	EXP_FUEL	-0,002	0,119	>0.000	0,01
$\Delta \log(x_t)$	IMP_FUEL	0,002	0,134	0,181	0,01
$\Delta \log(x_t)$	EXP_OIL	-0,002	0,197	>0.000	0,01
$\Delta \log(x_t)$	IMP_OIL	0,001	0,238	0,678	0,01
$\Delta \log(x_t)$	EXP_MACH	0,002	0,064	>0.000	0,01
$\Delta \log(x_t)$	IMP_MACH	0,004	0,066	>0.000	0,01
$\Delta \log(x_t)$	EXP_METAL	0,005	0,146	>0.000	0,01
$\Delta \log(x_t)$	IMP_METAL	0,001	0,152	0,009	0,01
$\Delta \log(x_t)$	EXP_CRUDE_MAT	0,002	0,091	>0.000	0,01
$\Delta \log(x_t)$	IMP_CRUDE_MAT	0,001	0,068	0,002	0,01
$\Delta \log(x_t)$	GBP_BROAD	-0,001	0,015	>0.000	0,01
$\Delta \log(x_t)$	GBP_CAN	-0,001	0,02	0,093	0,01
$\Delta \log(x_t)$	GBP_EUR	-0,001	0,017	>0.000	0,01
$\Delta \log(x_t)$	GBP_JAP	-0,001	0,029	>0.000	0,01
$\Delta \log(x_t)$	GBP_US	-0,001	0,021	>0.000	0,01
$\Delta \log(x_t)$	OIL_PRICE	0,006	0,107	>0.000	0,01

Table B.14: Summary Statistics for Group 14: UK Money and Credit

Transformation	Name	Mean	SD	JB	ADF
Δx_t	BANK_RATE	-0,009	0,165	>0.000	0,01
$\Delta \log(x_t)$	CONS_CREDIT_ex_student_loan	0,003	0,009	>0.000	0,01
$\Delta \log(x_t)$	TOT_LENDING_APP	-0,001	0,08	>0.000	0,01
$\Delta \log(x_t)$	TOT_HOUSE_APP	-0,002	0,134	>0.000	0,01
Δx_t	MORT_FIXED_RATE_5YRS	-0,009	0,173	>0.000	0,01
Δx_t	MORT_FIXED_RATE_2YRS	-0,008	0,193	>0.000	0,01
$\Delta \log(x_t)$	M1	0,006	0,011	>0.000	0,01
$\Delta \log(x_t)$	M2	0,005	0,01	>0.000	0,01
$\Delta \log(x_t)$	M3	0,005	0,01	>0.000	0,01
$\Delta \log(x_t)$	M4	0,005	0,008	>0.000	0,01
Δx_t	LIBOR_3mth	-0,009	0,193	>0.000	0,01
Δx_t	BGS_5yrs_yld	-0,008	0,201	>0.000	0,01
Δx_t	BGS_10yrs_yld	-0,007	0,192	>0.000	0,01
Δx_t	BGS_20yrs_yld	-0,004	0,162	>0.000	0,01

Table B.15: Summary Statistics for Group 15: UK Semtiment and Leading Indicators

Transformation	Name	Mean	SD	JB	ADF
$\Delta \log(x_t)$	FTSE_ALL	0,001	0,04	>0.000	0,01
$\Delta \log(x_t)$	FTSE250	0,004	0,05	>0.000	0,01
	VIX	20,673	8,057	>0.000	0,01
$\Delta \log(x_t)$	SP500	0,005	0,046	>0.000	0,01
$\Delta \log(x_t)$	UK_focused_equity	-0,001	0,05	>0.000	0,01
Δx_t	EUR_UNC_INDEX	0,639	42,559	>0.000	0,01
Δx_t	BCI	-0,004	0,443	>0.000	0,01
Δx_t	CCI	-0,018	0,363	>0.000	0,01
Δx_t	CLI	-0,005	0,742	>0.000	0,01
$\Delta \log(x_t)$	PPI_MANU	0,002	0,006	>0.000	0,01
$\Delta \log(x_t)$	PPI_MACH	0,002	0,004	>0.000	0,01
$\Delta \log(x_t)$	PPI_OIL	0,005	0,054	>0.000	0,01
$\Delta \log(x_t)$	PPI_METAL	0,003	0,018	>0.000	0,01
$\Delta \log(x_t)$	PPI_MOTOR	0,001	0,004	>0.000	0,01

C ARCH(p) h-step ahead forecast: formula derivation

A definition of conditional volatility in volatility models is given by Equation A.1, Equation A.2, and Equation A.3. This definition is used, where z_t is a *i.i.d.* standard normal series, and I_{t-1} is the information set available at one period before, to define a specification of the ARCH(p) in Equation C.1.

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 \tag{C.1}$$

Next, by recursively substituting the error term in Equation 4.3 a h-step ahead forecast formula for a fixed horizon h and varying order p can be obtained. The value of σ at time $t+1$ is only dependent on the information on the previous period and therefore it's known at time t. Therefore, the following is used when recursively substituting: $E(\sigma_{t+1}|I_t) = \sigma_{t+1}$. Equation C.2 defines the one step ahead forecast for $p \geq 1$.

$$\hat{\sigma}_{t+1} = w + \sum_{i=1}^{p} \alpha_i \varepsilon_{t+1-i}^2 \tag{C.2}$$

Equation C.3 defines the two step ahead forecast for $p \geq 2$.

$$\hat{\sigma}_{t+2} = w + \alpha_1 \sigma_{t+1}^2 + \sum_{i=2}^{p} \alpha_i \varepsilon_{t+2-i}^2 \tag{C.3}$$

Equation C.4 defines the three step ahead forecast for $p \geq 3$.

$$\hat{\sigma}_{t+3} = w\,(1+\alpha_1) + \sigma_{t+1}^2 \left(\alpha_1^2 + \alpha_2\right) + \sum_{i=1}^{p-2} \varepsilon_{t+1-i}^2 \left(\alpha_1 \alpha_{i+1} + \alpha_{i+2}\right) + \varepsilon_{t+p-2} \left(\alpha_1 \alpha_p\right) \tag{C.4}$$

D Results Tables

Table D.1: Hyperparameter Tuning: LARS, RL, and RWR

	MAF-Tuning					
	RL = 0.1, RWR = 0	RL = 0.1, RWR = 0.5	RL = 0.1, RWR = 0.95	RL = 0.5, RWR = 0	RL = 0.5, RWR = 0.5	RL = 0.5, RWR = 0.95
Horizon 1						
Lars = 30	**0.0357**	0.0789	0.1498	0.0431	0.0795	0.1560
Lars = 40	0.0387	0.0788	0.1484	0.0403	0.0751	0.1511
Lars = 50	0.0627	0.1042	0.1554	0.0653	0.0908	0.1529
Horizon 2						
Lars = 30	0.2743	0.1713	0.1971	0.1964	0.1494	0.1910
Lars = 40	**0.1647**	0.1553	0.2883	**0.1637**	0.1553	0.2137
Lars = 50	0.4099	0.2099	0.3056	0.3256	0.3196	0.2962
Horizon 3						
Lars = 30	0.9130	0.5334	0.4141	0.7796	0.4740	**0.4140**
Lars = 40	0.7606	0.6293	0.6506	0.6137	0.4859	0.5998
Lars = 50	1.3415	0.9658	1.0041	1.0638	1.0848	0.8587
Horizon 4						
Lars = 30	0.8254	0.7331	0.6459	0.7004	0.6503	**0.5881**
Lars = 40	1.7213	1.5908	1.5784	1.3746	1.3349	1.3833
Lars = 50	3.1087	1.8400	1.9135	2.1715	1.6687	1.7038
Horizon 6						
Lars = 30	1.6978	1.2881	1.3300	1.3295	1.2184	**1.1454**
Lars = 40	2.3776	2.2631	2.3029	2.2652	2.1286	2.3313
Lars = 50	4.7073	5.7453	4.9348	2.6356	3.8790	4.8383
Horizon 12						
Lars = 30	4.8362	3.6990	2.5870	4.0128	3.0476	**2.4269**
Lars = 40	3.5503	3.5917	3.7745	3.1225	2.8367	3.0000
Lars = 50	7.2472	7.7522	6.8151	4.5301	6.8062	5.2762

Notes: This table reports the Mean Squared Prediction Error (MSPE) of the MRF relative to the MSPE of the Random Walk. The best forecast configuration for horizon h is in bold.

Table D.2: Hyperparameter Tuning: MAF/NO-MAF

MAF

Horizon 1	RL		
lags	0.35	0.5	0.65
3	0.2476	0.2493	0.2464
5	0.1603	0.1559	0.1555
8	0.0970	**0.0925**	0.0972

Horizon 2	RL		
lags	0.35	0.5	0.65
3	0.3274	0.3784	0.3688
5	0.1778	0.1910	0.1567
8	**0.1541**	0.1643	0.1631

Horizon 4	RL		
lags	0.35	0.5	0.65
3	0.8182	0.7228	0.7067
5	0.7005	0.5881	**0.6491**
8	0.9467	1.1379	0.9877

NO-MAF

Horizon 1	RL		
lags	0.35	0.5	0.65
3	0.2595	0.2569	0.2646
5	0.1707	0.1741	0.1722
8	**0.1114**	0.1144	0.1128

Horizon 2	RL		
lags	0.35	0.5	0.65
3	0.3476	0.3428	0.3363
5	0.1926	**0.1736**	0.1903
8	0.1833	0.1847	0.1831

Horizon 4	RL		
lags	0.35	0.5	0.65
3	0.6401	0.6317	**0.5798**
5	0.7217	0.7065	0.6504
8	1.0206	0.8814	0.9917

Notes: This table reports the Mean Squared Prediction Error (MSPE) of the MRF relative to the MSPE of the Random Walk. The best forecast configuration for horizon h is in bold.

Table D.3: Hyperparameter tunning: PCA and MAF

Horizon 1	PCA		
MAF	5	10	15
5	**0.0988**	0.1000	-
10	0.1095	0.1074	0.1012
15	-	0.1083	0.1086

Horizon 2	PCA		
MAF	5	10	15
5	0.1646	0.1440	-
10	0.1349	**0.1333**	0.1392
15	-	0.1385	0.1377

Horizon 3	PCA		
MAF	5	10	15
5	**0.1266**	0.4733	-
10	0.4091	0.4520	0.4398
15	-	0.4268	0.4373

Notes: This table reports the Mean Squared Prediction Error (MSPE) of the MRF relative to the MSPE of the Random Walk. The best forecast configuration for horizon h is in bold.

Table D.4: MSPE of AR(p) for h = 1, 2, 4, 6, and 12.

	Horizons				
p	**1**	**2**	**4**	**6**	**12**
1	0.59684	0.53648	**0.45984**	0.48408	0.50137
2	0.59809	0.53762	0.45985	**0.48406**	**0.50135**
3	**0.59555**	**0.53574**	0.46525	0.48419	0.50150
4	0.60221	0.54126	0.46933	0.48454	0.50156
5	0.60469	0.54162	0.47040	0.48511	0.50148
6	0.61364	0.55143	0.47780	0.49260	0.50165
7	0.62223	0.56396	0.48688	0.49963	0.50678
8	0.62410	0.56577	0.48884	0.50063	0.50700
9	0.62615	0.56788	0.48991	0.50136	0.50681
10	0.62691	0.56696	0.48958	0.50132	0.50628

Notes: Lowest MSPE for each horizon h is in bold.

Table D.5: VaR results of MRF-ARCH for h = 1

	Quantile level = 0.05			Quantile level = 0.1			Quantile level = 0.15		
p	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**
1	0.003995	0.004163	0.003729	0.008070	0.008750	0.008405	NP	NP	NP
2	NP	NP	0.003859	0.008302	0.008993	0.008642	NP	NP	NP
3	NP	NP	NP	NP	NP	NP	NP	NP	NP
4	0.004074	0.004244	0.003807	0.008163	0.008846	0.008499	NP	NP	NP
5	**0.003952**	**0.004118**	**0.003691**	**0.007989**	**0.008670**	**0.008324**	NP	NP	NP
6	0.003998	0.004166	0.003734	0.008071	0.008756	0.008408	NP	NP	NP
7	NP	NP	NP	NP	NP	NP	NP	NP	NP
8	NP	NP	NP	NP	NP	NP	NP	NP	NP
9	NP	NP	NP	NP	NP	NP	NP	NP	NP
10	NP	NP	NP	NP	NP	NP	NP	NP	NP
11	NP	NP	NP	NP	NP	NP	NP	NP	NP
12	NP	NP	NP	NP	NP	NP	NP	NP	NP

Notes: i) This table reports the regulatory loss function value for different VaR models, where the volatility is forecasted by MRF-ARCH(p) for different autoregressive orders p and different quantile distributions at different quantile levels. ii) 'NP' signifies that either the conditional coverage, unconditional coverage, or independence of the Var model is violated.

Table D.6: VaR results of MRF-ARCH for h = 2

	Quantile level = 0.05			Quantile level = 0.1			Quantile level = 0.15		
p	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**
1	0.003321	0.003476	0.003078	0.007149	0.007798	0.007467	0.011083	**0.012059**	**0.011925**
2	0.003461	0.003621	0.003210	0.007382	0.008047	0.007709	0.011379	NP	NP
3	NP	NP	NP	0.008492	NP	0.008832	NP	NP	NP
4	0.003315	0.003470	0.003072	0.007141	0.007789	0.007459	NP	NP	NP
5	**0.003221**	**0.003373**	**0.002985**	**0.006996**	**0.007641**	**0.007313**	**0.010938**	NP	NP
6	0.003274	0.003428	0.003034	0.007087	0.007736	0.007406	0.011050	NP	NP
7	0.003600	0.003770	0.003333	0.007713	0.008404	0.008053	NP	NP	NP
8	NP	NP	NP	NP	NP	NP	NP	NP	NP
9	NP	NP	NP	NP	NP	NP	NP	NP	NP
10	NP	NP	NP	NP	NP	NP	NP	NP	NP
11	NP	NP	NP	NP	NP	NP	NP	NP	NP
12	NP	NP	NP	NP	NP	NP	NP	NP	NP

Notes: i) This table reports the regulatory loss function value for different VaR models, where the volatility is forecasted by MRF-ARCH(p) for different autoregressive orders p and different quantile distributions at different quantile levels. ii) 'NP' signifies that either the conditional coverage, unconditional coverage, or independence of the Var model is violated.

Table D.7: VaR results of MRF-ARCH for h = 3

	Quantile level = 0.05			Quantile level = 0.1			Quantile level = 0.15		
p	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**
1	0.003195	0.003347	0.002958	0.006974	0.007617	0.007290	0.010885	**0.011860**	**0.011725**
2	0.003254	0.003408	0.003012	0.007083	0.007734	0.007403	0.011031	0.012008	0.011873
3	0.003834	NP	NP	0.008215	NP	0.008549	**0.010644**	NP	NP
4	0.003196	0.003345	0.002964	0.006922	0.007551	0.007230	NP	NP	NP
5	**0.003108**	**0.003253**	**0.002883**	**0.006778**	**0.007404**	**0.007085**	NP	NP	NP
6	0.003139	0.003286	0.002910	0.006842	0.007472	0.007151	NP	NP	NP
7	0.003353	0.003517	0.003097	0.007349	0.008024	0.007681	NP	NP	NP
8	0.002306	NP	0.002083	0.006119	NP	NP	NP	NP	NP
9	0.002546	0.002691	0.002321	0.006372	NP	NP	NP	NP	NP
10	NP	NP	NP	NP	NP	NP	NP	NP	NP
11	NP	NP	NP	NP	NP	NP	NP	NP	NP
12	NP	NP	NP	NP	NP	NP	NP	NP	NP

Notes: i) This table reports the regulatory loss function value for different VaR models, where the volatility is forecasted by MRF-ARCH(p) for different autoregressive orders p and different quantile distributions at different quantile levels. ii) 'NP' signifies that either the conditional coverage, unconditional coverage, or independence of the Var model is violated.

Table D.8: VaR results of ARCH for h = 1

	Quantile level = 0.05			Quantile level = 0.1			Quantile level = 0.15		
p	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**
1	0.003399	NP	0.003411	0.006385	NP	NP	0.009730	NP	NP
2	0.010256	NP	0.010265	0.012743	0.014071	0.013686	0.015629	NP	0.017453
3	0.002969	0.003488	0.002983	0.006378	0.008107	0.007612	0.010063	NP	0.012303
4	**0.002536**	**0.003020**	**0.002548**	**0.005709**	**0.007324**	**0.006856**	0.009208	NP	0.011411
5	0.002595	0.003080	0.002608	0.005823	0.007502	0.007019	0.009432	NP	NP
6	0.002580	0.003057	0.002592	0.005790	0.007478	0.006992	0.009414	NP	0.011660
7	0.003109	0.003660	0.003123	0.006679	0.008481	0.007969	**0.000039**	NP	NP
8	0.003333	0.003901	0.003348	0.006957	0.008729	0.008225	0.010708	NP	0.012971
9	0.003353	0.003925	0.003368	0.006996	0.008774	0.008269	0.000045	NP	0.013023
10	0.003548	0.003563	NP	0.007284	0.009088	0.008577	0.011085	NP	**0.000043**
11	0.003642	NP	NP	0.007399	0.009207	0.008695	0.011204	NP	NP
12	0.003610	NP	0.003625	0.007361	0.009169	0.008656	0.011166	NP	0.000053

Notes: 'NP' signifies that either the conditional coverage, unconditional coverage, or independence of the Var model is violated. Lower values for each quantile level and quantile distribution is marked in bold.

Table D.9: VaR results of ARCH for h = 2

	Quantile level = 0.05			Quantile level = 0.1			Quantile level = 0.15		
p	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**
1	0.004908	0.005367	0.004920	0.007944	0.009472	0.009038	0.011207	0.013440	0.013225
2	0.006648	0.007001	0.006657	0.009170	0.010559	0.010159	0.012192	0.014378	0.014165
3	0.001891	0.002327	0.001902	0.004936	0.006562	0.006095	0.008449	0.010934	0.010693
4	**0.001869**	**0.002301**	**0.001880**	**0.004895**	**0.006517**	**0.006052**	**0.008399**	**0.010881**	**0.010639**
5	0.001903	0.002335	0.001914	0.004929	0.006551	0.006085	0.008434	0.010915	0.010674
6	0.001912	0.002343	0.001923	0.004942	0.006576	0.006106	0.008472	0.010967	0.010724
7	0.002600	0.003140	0.002614	0.006071	0.007800	0.007305	0.009777	0.012321	0.012075
8	0.002754	0.003309	0.002769	0.006295	0.008023	0.007532	0.009980	0.012507	0.012263
9	0.002780	0.003339	0.002794	0.006339	0.008072	0.007580	0.010033	0.012564	0.012319
10	0.002978	0.003561	0.002993	0.006641	0.008399	0.007902	0.010367	0.012893	0.012649
11	0.003058	0.003646	0.003073	0.006744	0.008510	0.008010	0.010486	0.013016	0.012772
12	0.003023	0.003609	0.003038	0.006700	0.008464	0.007965	0.010440	0.012970	0.012725

Notes: 'NP' signifies that either the conditional coverage, unconditional coverage, or independence of the Var model is violated. Lower values for each quantile level and quantile distribution is marked in bold.

Table D.10: VaR results of ARCH for h = 3

	Quantile level = 0.05			Quantile level = 0.1			Quantile level = 0.15		
p	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**	**Norm**	**Fitted t-distr.**	**Fitted skewed-t**
1	NP	NP	NP	0.007573	0.009144	0.008693	0.010953	0.013292	0.013066
2	0.001603	0.001978	0.001612	0.004374	0.005949	0.005494	0.007797	0.010270	0.010028
3	0.003429	0.003838	0.003440	0.006286	0.007822	0.007381	0.009600	0.011950	0.011721
4	**0.001155**	**0.001536**	**0.001165**	**0.003982**	**0.005577**	**0.005116**	**0.007441**	**0.009922**	**0.009680**
5	0.001184	0.001567	0.001194	0.004017	0.005612	0.005152	0.007480	0.009961	0.009719
6	0.001302	0.001693	0.001312	0.004190	0.005804	0.005338	0.007687	0.010181	0.009937
7	0.002039	0.002558	0.002052	0.005440	0.007168	0.006672	0.009146	0.011700	0.011453
8	0.002081	0.002612	0.002095	0.005539	0.007267	0.006776	0.009224	0.011763	0.011517
9	0.002105	0.002641	0.002119	0.005581	0.007314	0.006822	0.009276	0.011818	0.011572
10	0.002283	0.002843	0.002297	0.005866	0.007625	0.007127	0.009596	0.012136	0.011890
11	0.002400	0.002968	0.002414	0.006020	0.007788	0.007288	0.009767	0.012310	0.012064
12	0.002368	0.002934	0.002383	0.005978	0.007744	0.007244	0.009722	0.012265	0.012019

Notes: 'NP' signifies that either the conditional coverage, unconditional coverage, or independence of the Var model is violated. Lower values for each quantile level and quantile distribution is marked in bold.

Table D.11: DM test results for h = 1, 2, 4.

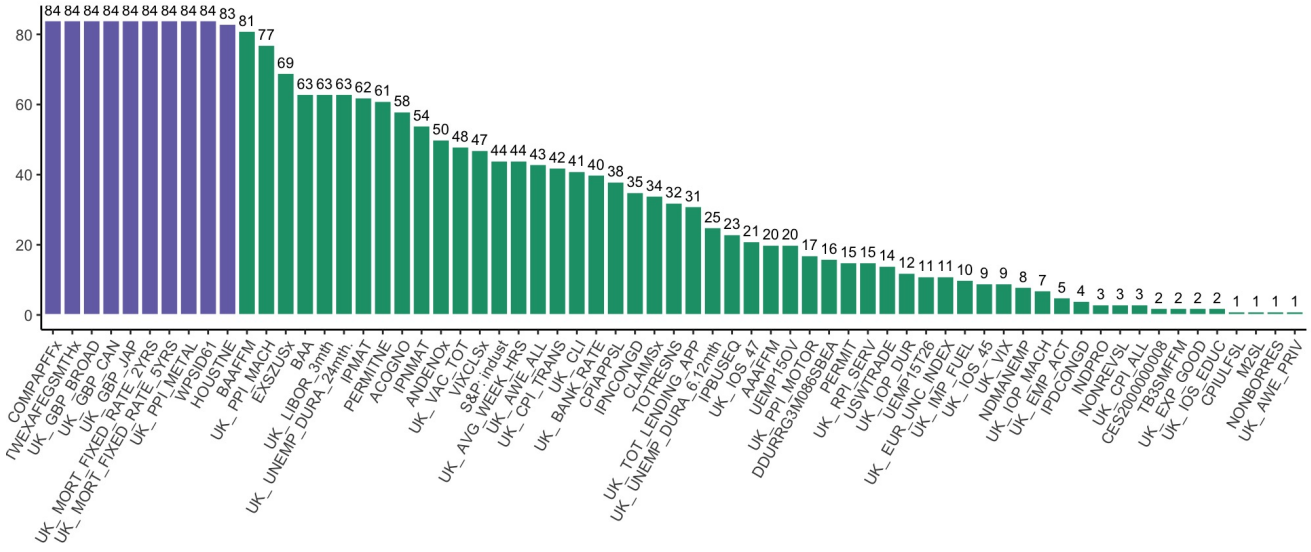
Horizon 1	Real-Time (5)	Real-Time (8)	Ex-Post (5)	Ex-Post (8)
Real-Time (5)	-	1.0000	0.9377	0.9994
Real-Time (8)	0.0000	-	0.5127	0.9721
Ex-Post (5)	0.0623	0.4873	-	1.0000
Ex-Post (8)	0.0006	0.0279	0.0000	-

Horizon 2	Real-Time (5)	Real-Time (8)	Ex-Post (5)	Ex-Post (8)
Real-Time (5)	-	**0.0090**	0.8627	0.3121
Real-Time (8)	0.9910	-	0.9916	0.9033
Ex-Post (5)	0.1373	0.0084	-	**0.0127**
Ex-Post (8)	0.6879	0.0967	0.9873	-

Horizon 4	Real-Time (5)	Real-Time (8)	Ex-Post (5)	Ex-Post (8)
Real-Time (5)	-	**0.0089**	0.8080	**0.0300**
Real-Time (8)	0.9911	-	0.9937	0.9093
Ex-Post (5)	0.1920	0.0063	-	0.0169
Ex-Post (8)	0.9700	0.0907	0.9831	-

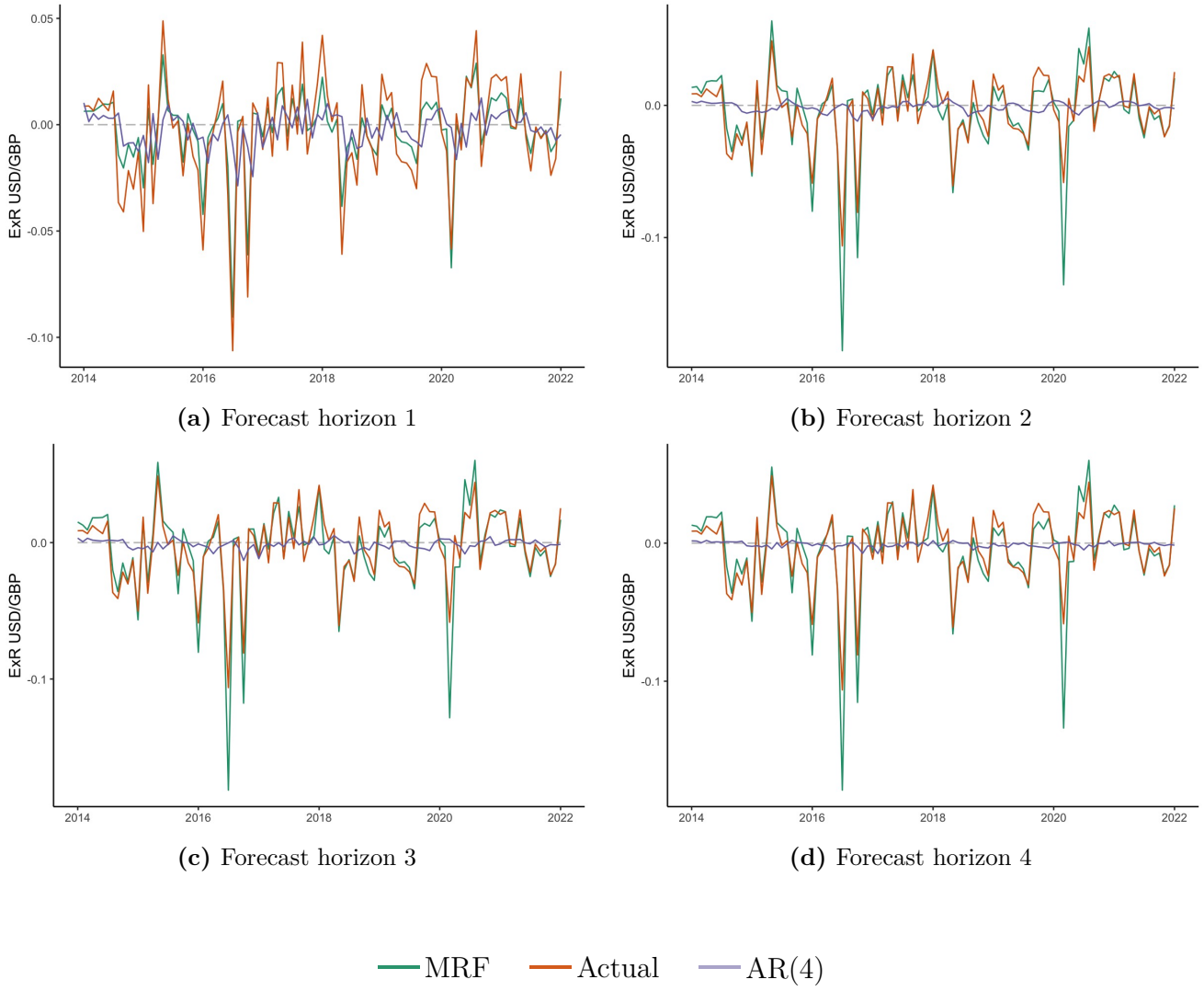
Note: Result P-values for one sided DM-test with small sample correction suggested in Harvey et al. (1997). P-values in bold indicate significance at a 5% level. Number of lags used for each dataset is indicated in parenthesis.

E Results Figures



Note: UK variables are prefixed with UK_. The 10 most chosen variables shown in purple.

Figure E.1: Number of times variables are selected by LARS algorithm.



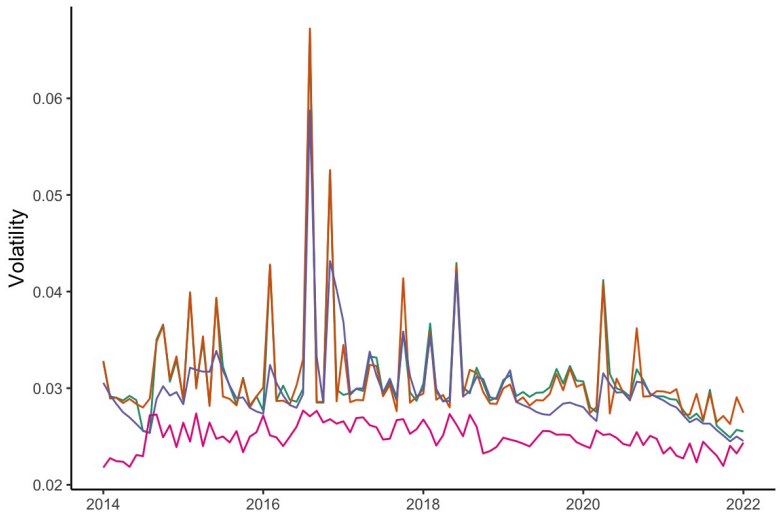
(a) Forecast horizon 1

(b) Forecast horizon 2

(c) Forecast horizon 3

(d) Forecast horizon 4

Figure E.2: USD/GBP forecast of MRF and AR for h = 1, 2, 3, and 4.



(a) 2-month ahead volatility forecasts

MRF-ARCH(5) ARCH(4) GARCH(1,1) ARMA(1,1)-GARCH(1,1)

Figure E.3: USD/GBP volatility forecasts of MRF-ARCH, ARCH, GARCH, and ARMA-GARCH for h = 2.

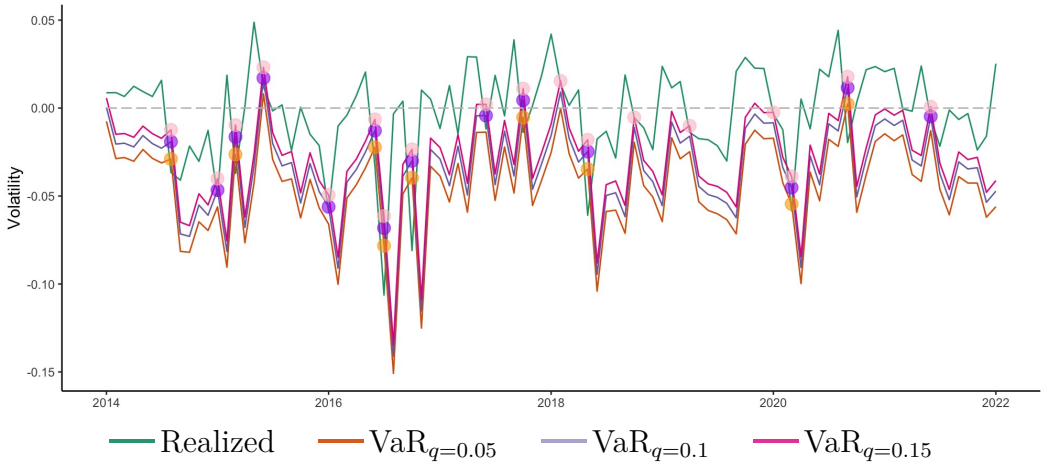
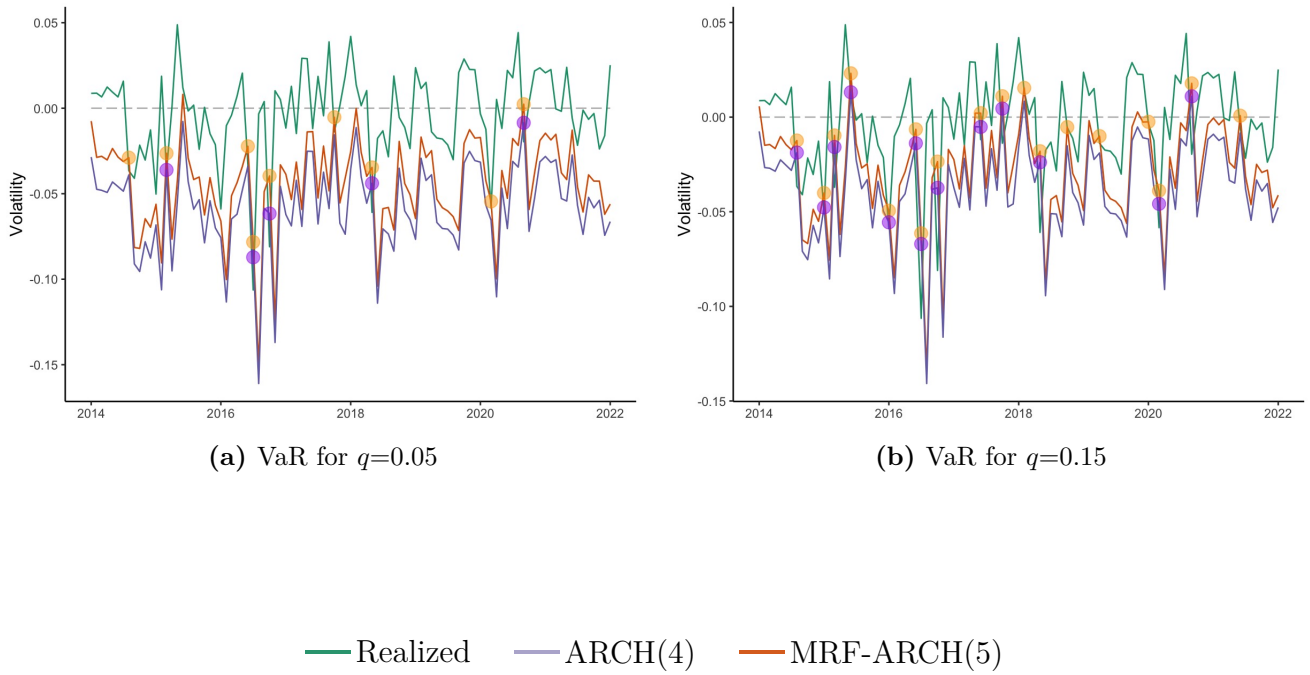


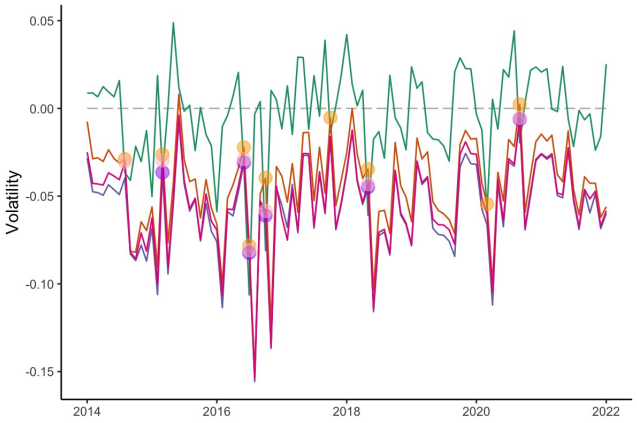
Figure E.4: MRF-ARCH VaR using normally distributed quantiles for q = 0.05, 0.1, and 0.15 and h = 1.



(a) VaR for q=0.05

(b) VaR for q=0.15

Figure E.5: MRF-ARCH and ARCH VaR using normally distributed quantiles for q = 0.05, 0.1, and 0.15 and h = 1.



(a) VaR for q=0.05

(b) VaR for q=0.1

(c) VaR for q=0.15

MRF GARCH(1,1) MRF-ARCH(5) ARMA(1,1)-GARCH(1,1)

Figure E.6: VaR of GARCH, MRF-ARCH, and ARMA-GARCH using normally distributed quantiles for q = 0.05, 0.1, and 0.15 and h = 1.