Modeling the 1-year U.S. Treasury Yield Curve

Marco Cavallari, Maria Sole Chiaramonti

June 3, 2024

1 Problem Statement

The Market Yield on U.S. Treasury Securities at 1-Year Constant Maturity, often referred to as the "1-Year Treasury Yield," represents the return an investor would expect to earn from holding a U.S. Treasury security with a one-year maturity.

Forecasting the future behavior of the 1-Year Treasury Yield is crucial for financial institutions (e.g., banks, insurance companies) which are exposed to interest rate risk, as it allows them to better manage this risk, to maintain appropriate capital reserves, and to adjust their asset-liability strategies. Moreover, the yield curve is often considered an indicator of broader economic conditions, and forecasting its conditional volatility can provide insights into market sentiment and risk [7].

To this scope, we will conduct a time series analysis of the 1-Year Treasury Yield curve, modeling it with ARIMA models and GARCH models. We will follow a robust procedure to ensure that our results are unbiased: we will fit the models on a train set, evaluate them on the test set and analyze the residuals of our forecast to ensure that the assumptions of the models hold [8].

2 Dataset

We obtained the time series of the 1-Year Constant Maturity Treasury Yield curve, quoted on an investment basis, from the Federal Reserve Economic Data repository, containing daily data from 1978 to 2024 [6]. Such an extensive range allows for an in-depth analysis of yield fluctuations over nearly five decades, offering valuable insights into long-term trends and patterns in U.S. Treasury yields.

Upon importing the datasets, missing values are imputed with the last available observation. This technique known as "last observation carried forward" (LOCF) is one of the simplest methods to account for missing data in a time series which succeeds at mantaining data continuity without introducing undue bias [10].

Aside from our time series of interest, we retrieved other time series to use as exogenous regressors for an ARIMAX model from the same repository. Specifically, we obtained the time series of Treasury Yield curves at different maturities (e.g., 2, 3, 5, 7, 10, 20, 30 years).

3 Methods

In our analysis, the initial step involves assessing the stationarity of the time series, which may include differencing the time series until it resembles a realization of some stationary process $\{Y_t\}$ [1]. We will then use the theory of stationary processes for the modeling, analysis and prediction of $\{Y_t\}$, and hence for the original process [2]. When speaking about stationarity, we mean weakly stationary, which is defined as

- The unconditional mean $E[Y_t]$ is not a function of t.
- The unconditional variance $Var(Y_t) = \sigma^2$ is not a function of t.
- $Cov(Y_t, T_{t+h})$ is independent of t for each h.

To test for stationarity of a time series, we plot the time series alongside its autocorrelation function (ACF), which displays the correlation ρ_k between Y_t and Y_{t-k}, which is defined as

$$\rho_k = \frac{Cov(Y_t, Y_{t-k})}{Var(Y_t)} [9],$$

and perform a Dickey–Fuller (DF), which examines the null hypothesis that a unit root is present in the time series, implying non-stationarity. Introduced by Dickey and Fuller in 1981 [3], the DF test involves estimating the following regression equation:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$$

where $\Delta y_t = y_t - y_{t-1}$ is the first difference of the time series, y_{t-1} is the lagged value of the series, α is a constant (optional, can be included to account for a non-zero mean), and ϵ_t is the error term, assumed to be a white noise process with zero mean and variance σ^2. We are testing for $H_0 : \gamma = 0$, which implies that a unit root is present. Conversely, the alternative hypothesis $H_1 : \gamma < 0$ suggests that the series does not have unit root. We use a significance level $\alpha = 0.01$ to test the null hypothesis.

Once we have established that the differentiated time series is stationary, we try to understand the underlying DGP by fitting autoregressive models, such as ARIMA and GARCH, to estimate the conditional mean and variance of the time series. We estimate the parameters on the training set (70% of obs) using maximum likelihood estimation (MLE) and compare the goodness of the models on the test set using information criteria such as Akaike's Information Criterion (AIC), AICC, the Bayesian Information Criterion

(BIC), and the Log Likelihood (LL), in order to select the best-fitting model from the remaining out-of-sample observations. After selecting the best performing model, we will carry out diagnostic checks on the residuals to check whether they follow a white noise process, thus ensuring that the model captures the underlying DGP [2].

The ARIMA(p,d,q) model assumes that the stationary, mean-centered process $\Delta^d\{Y_t\}$ is a linear combination of three terms: autoregressive (p), differencing (d), and moving average (q). It models the relationship between a variable, its lagged values and the lagged values of its error term. The ARIMA model estimates the parameters (ϕ, θ, σ^2) that define the process

$$\text{ARIMA(p,d,q)} : \Delta^d y_t = \phi_1 \Delta^d y_{t-1} + \ldots + \phi_p \Delta^d y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \ldots - \theta_q \epsilon_{t-q}, \quad \epsilon_t \sim WN(0, \sigma^2)$$

$$\Delta^d y_t = \sum_{i=1}^{p} \phi_i \Delta^d y_{t-i} + \epsilon_t - \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

We can extend the concept of ARIMA by introducing in the model a $n \times r$ matrix, containing r stationary processes, in order to improve forecasting accuracy. The ARMAX model can be expressed as:

$$\text{ARIMAX(p,d,q)} : \Delta^d y_t = \sum_{i=1}^{p} \phi_i \Delta^d y_{t-i} + \epsilon_t - \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{r} \beta_k X_{t-k}$$

If the residuals of the ARIMA model ϵ_t resemble a white noise process with zero mean and finite variance σ^2, but they display significant autocorrelation in the squares, it means that we haven't captured all the dinamics of the DGP.

The GARCH(p,q) model assumes that the process is a martingale-difference sequence, i.e. $E[y_t|\mathcal{F}_{t-1}] = 0$, so that $y_t = \epsilon_t$ and the error term is equal to some random variable z_t, which follows a normal distribution with zero mean and finite variance (usually equal to 1), scaled by an $\mathcal{F}_{t-1}$-measurable quantity $\sigma_t : \epsilon_t = \sigma_t z_t, z_t \sim IID(0,1)$, implying that $\text{Var}[y_t|\mathcal{F}_{t-1}] = \sigma^2_{t|t-1}$. The notations $\sigma^2_{t|t-1}$ and σ^2_t are used interchangeably to refer to the conditional variance $\text{Var}[y_t|\mathcal{F}_{t-1}]$. GARCH(p,q) models the conditional variance as:

$$\text{GARCH(p,q)} : \sigma^2_{t|t-1} = \omega + \alpha_1 y^2_{t-1} + \ldots + \alpha_r y^2_{t-p} + \beta_1 \sigma^2_{t-1} + \ldots + \beta_q \sigma^2_{p-q},$$

which can be written as:

$$\sigma^2_{t|t-1} = \omega + \sum_{i=1}^{q} \alpha_i y^2_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j}, \quad t \in Z$$

with

$$\alpha \geq 0, \beta \geq 0, \omega \geq 0, \alpha + \beta < 1$$

.

Considering that all the variability of the ARIMA model comes from the error term ϵ_t, if it satisfied the assumption of a GARCH model, we can predict the conditional variance of ϵ_t to build a robust confidence interval around the point estimates $\hat{y}_t$. Eventually, we will evaluate the goodness of fit of the GARCH model by studying the behavior its residuals z_t, which should be IID(0,1) [5].

4 Analysis

Figure 1 suggests that the 1-Year Treasury Yield, which we will now denote as Y_t, has unit root, as it displays significant, slowly decreasing autocorrelation up to high lags and partial autocorrelation close to one at lag 1. The results of the DF test performed on Y_t (p-value $= 0.7973$) confirm that the time series has a unit root, and thus it is non-stationary.

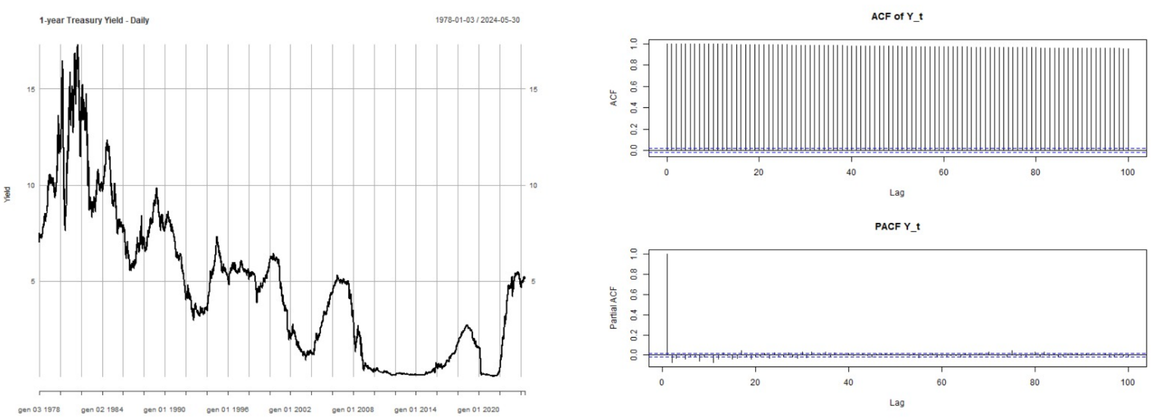

Figure 1: Historical plot of Y_t (left), ACF of Y_t (top-right) and PACF of Y_t (bottom-right)

As suggested by Box and Jenkins [1], we transform the time series Y_t to achieve stationarity by taking its first difference, denoted as ΔY_t. Figure 2 shows that the differenced time series has mean close to zero and exhibits no apparent unit root. The PACF plot reveals significant correlations at various lags, indicating the presence of autoregressive structure. The DF test on ΔY_t produces a p-value smaller then our significance level $\alpha = 0.01$, confirming that the differenced series does not have unit root. This suggests that we should use a differencing order $d \geq 1$ in the ARIMA models fitted to Y_t.

We fitted multiple ARIMA models, exploring 128 combinations of p (0 to 7), d (1 to 2), and q (0 to 7), to estimate the model parameters $\hat{\Phi}$, $\hat{\Theta}$, and $\hat{\sigma^2}$ on the training set. These parameters were then used to make one-step-ahead predictions on the test set. The model specification (p,d,q) that minimized the BIC on the test set was selected. In Table 1 we report the specifications and coefficients of the preferred ARIMA model, while its information criteria and $\hat{\sigma^2}$ are displayed in Table 2.

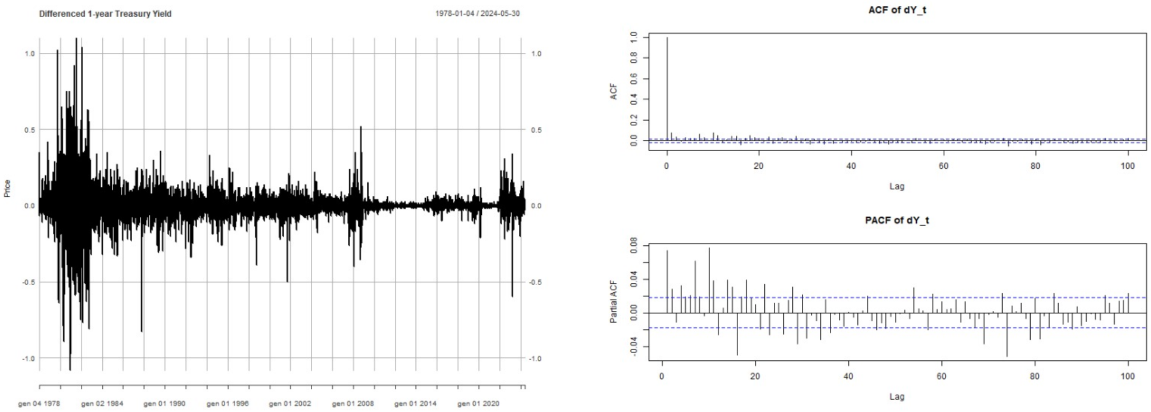

Figure 2: Plot of ΔY_t (left), ACF of ΔY_t (top-right) and PACF of ΔY_t (bottom-right)

Model Specification: ARIMA(5,1,5)										
	ar1	**ar2**	**ar3**	**ar4**	**ar5**	**ma1**	**ma2**	**ma3**	**ma4**	**ma5**
Value	0.1863	0.7739	0.2550	-0.6343	0.3074	-0.1154	-0.7738	-0.3151	0.6156	-0.2303
SE	0.1402	0.0721	0.1536	0.0874	0.1233	0.1410	0.0751	0.1508	0.0838	0.1205

Table 1: Best ARIMA model summary

	σ^2	**LL**	**AIC**	**AICc**	**BIC**
In-sample	0.009354	7775.73	-15529.46	-15529.43	-15451.97
Out-of-sample	0.009354	7299.72	-14597.45	-14597.45	-14591.25

Table 2: Summary statistics of ARIMA (5,1,5) model

To improve the accuracy of our estimates, we applied the same approach to identify the best specification for an ARIMAX model. This model incorporates the differentiated Treasury Yields at various maturities (e.g., 2, 3, 5, 7, 10, 20, 30 years) as external exogenous regressors to predict Y_t. As illustrated in Figure 3, the cross-correlation function (CCF) plots between ΔY_t and its potential regressors display significant cross-correlation at the first lag, suggesting the usefulness of these regressors in predicting Y_t. The model that yields the lowest BIC on the test set is an ARIMAX(0,1,0), which essentially corresponds to a linear regression of ΔY_t on the differentiated regressors $\Delta Y_{2t}, \ldots, \Delta Y_{30t}$. The coefficients and information criteria of this model are displayed in Tables 3 and 4, respectively.

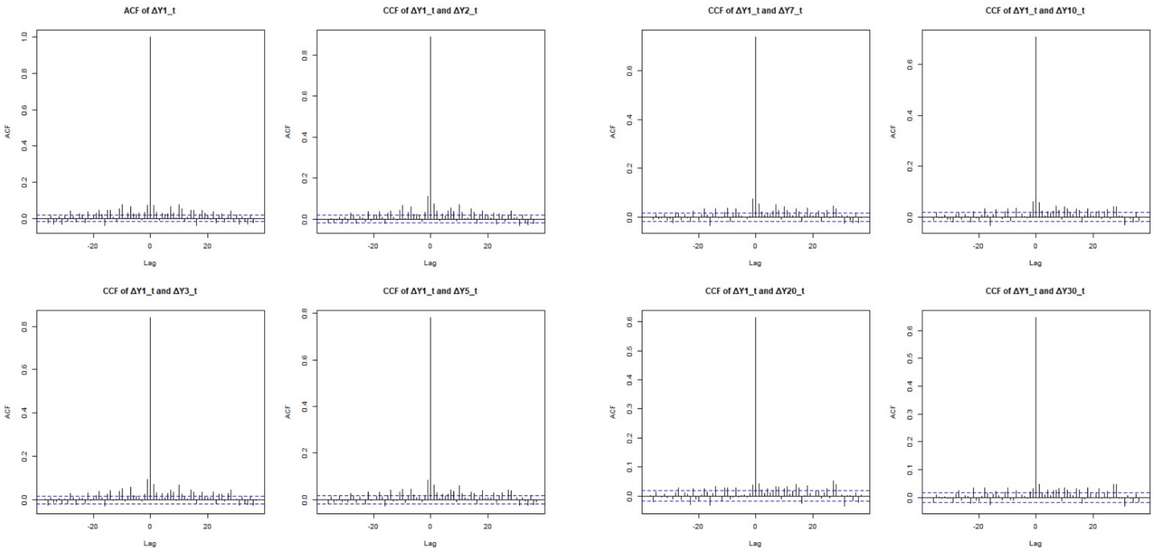

Figure 3: CCF between ΔY_t and the differentiated potential regressors ($\Delta Y_{2t}, \ldots, \Delta Y_{30t}$)

Model Specification: ARIMAX(0,1,0)							
	ΔY_{2t}	ΔY_{3t}	ΔY_{5t}	ΔY_{7t}	ΔY_{10t}	ΔY_{20t}	ΔY_{30t}
Value	0.4398	0.0592	0.0442	-0.0028	-0.1102	0.0163	0.0371
SE	0.0192	0.0253	0.0302	0.0324	0.0328	0.0183	0.0235

Table 3: Best ARIMAX model summary

	σ^2	**LL**	**AIC**	**AICc**	**BIC**
In-sample	0.006012	9647.17	-19278.33	-19278.32	-19221.98
Out-of-sample	0.006012	7747.75	-15493.49	-15493.49	-15487.30

Table 4: Summary statistics of ARIMAX (0,1,0) model

We continue by analyzing the residuals of the best model ARIMAX(0,1,0) to assess whether they exhibit characteristics of a white noise process, i.e. no significant autocorrelation, zero mean and constant variance. By inspecting Figure 4, we can see that the error term of the ARIMAX(0,1,0) model can't be white noise since it displays significant autocorrelation at various lags. Consequently, we go back to considering the ARIMA(5,1,5) model to assess whether its residuals are more similar to a white noise process. Figure 5 suggests that the error term ϵ_t is a student-t distributed, zero-mean process with no significant autocorrelation. However, there is significant autocorrelation in the squared residuals, suggesting that there is likely dependence in the conditional variance of ϵ_t. This indicates that the ARIMA model has not fully captured all underlying structures of the

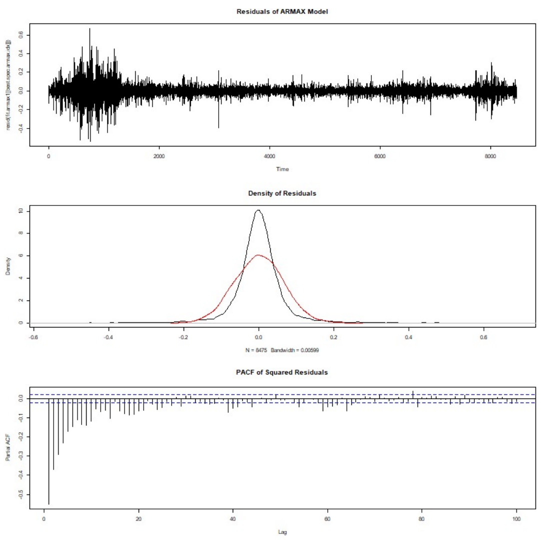

Figure 4: Residual plots of ARIMAX(0,1,0): residuals (top), density of residuals with normal distribution (center), PACF of squared residuals (bottom)

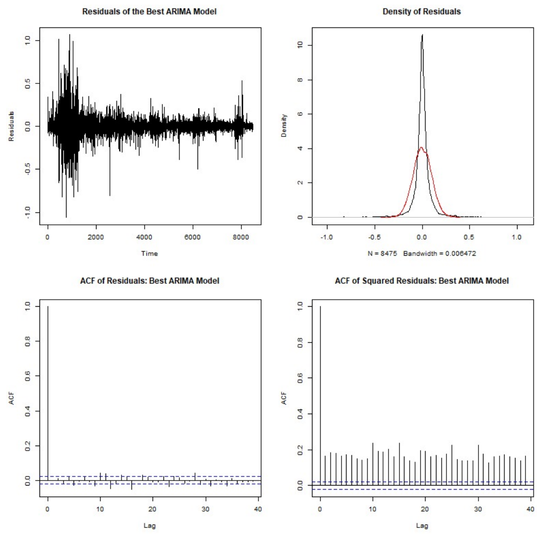

Figure 5: Residual plots of ARIMA(5,1,5): residuals (top-left), density of residuals with normal distribution (top-right), ACF of residuals (bottom-left), ACF of squared residuals (bottom-right)

DGP, pointing to the presence of time-varying conditional variance or volatility clustering, which is a common characteristic in financial time series [5].

We address this issue by fitting a GARCH(1,1) model on the residuals of the ARIMA(5,1,5) model. From Figure 6 we can see that the residuals of the GARCH(1,1) model are much more similar to white noise than the residuals of the original ARIMA(5,1,5) model, implying that the GARCH model is indeed capturing the autocorrelation structure of the squared residuals of the ARIMA model, i.e. the conditional variance of Y_t. Figure 7 shows that there is no left out autocorrelation in the residuals of the GARCH model, and their squares confirming that the GARCH(1,1) is a good fit for the data.

Model Specification: GARCH(1,1)			
	a0	**a1**	**b1**
Value	1.353261e-05	6.677057e-02	9.358033e-01

Table 5: GARCH(1,1) model parameters

Since all the variability in Y_t is due to ϵ_t such that $Var(Y_t) = Var(\epsilon_t)$, we can leverage the GARCH(1,1) model to predict the conditional variance of ϵ_t [4]. This prediction enables us to construct a prediction interval around the point forecast provided by the ARIMA(5,1,5) model: $CI = \hat{Y}_t \pm 1.96\hat{\sigma}_t$, assuming that $z_t \sim NID(0,1)$. To determine $\hat{\sigma}_t$, we need to use the coefficients of the GARCH(1,1) model and estimates $\hat{\sigma^2}_t$ as follows

$$\sigma^2_{t|t-1} = \omega + \alpha_1 y^2_{t-1} + \beta_1 \sigma^2_{t-1}$$

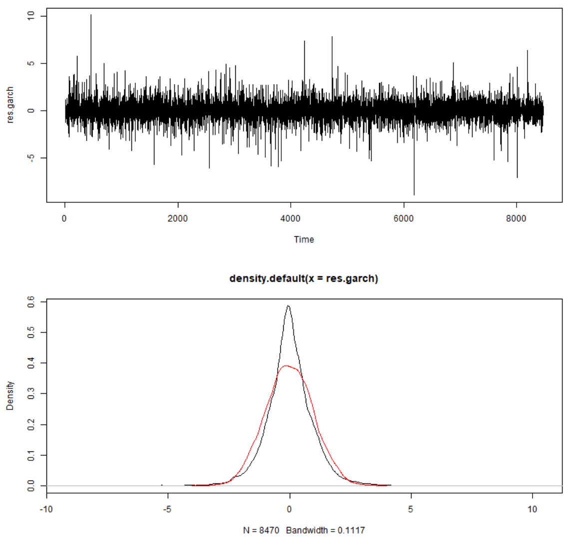

Figure 6: Plot (top) and density plot with normal distribution (bottom) of the residuals from GARCH(1,1)

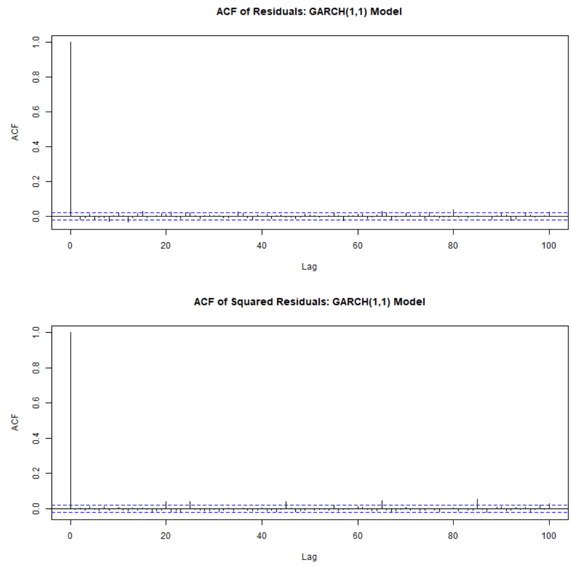

Figure 7: ACF of residuals (top) and of squared residuals (bottom) from GARCH(1,1)

By substituting the estimates of the parameters of our model into the equation, we obtain

$$\sigma^2_{t|t-1} = 1.353261 \times 10^{-5} + 0.067 y^2_{t-1} + 0.936 \sigma^2_{t-1}$$

Figure 8 illustrates the effectiveness of the GARCH(1,1) model in capturing the time-varying variance of the residuals from the ARIMA(5,1,5) model. The estimated conditional variances, $\hat{\sigma}_t$, align closely with the observed volatility, demonstrating the model's accuracy. Additionally, the figure illustrates the confidence intervals that encompass approximately 95% of the observed volatility, providing a visual validation of the model's performance. In Figure 9, we present the one-step-ahead point forecast for the next unseen period. This final model is employed to predict both the point estimate and the corresponding confidence interval. This last figure highlights the predictive accuracy of our approach, underlining the robustness of our combined ARIMA-GARCH framework for forecasting future values and their associated uncertainty.

In conclusion, our analysis demonstrates that the ARIMA(5,1,5) model, augmented by a GARCH(1,1) model for the residuals, effectively captures the dynamics of the 1-Year U.S. Treasury Yield. While the initial ARIMAX(0,1,0) model using exogenous regressors did not produce white noise residuals, the ARIMA(5,1,5) model's residuals indicated the presence of conditional heteroskedasticity, addressed by the GARCH model. This combined approach successfully modeled both the conditional mean and variance, providing accurate point forecasts and prediction intervals. The results highlight the importance of considering both ARIMA and GARCH models for financial time series to account for both mean and volatility dynamics.

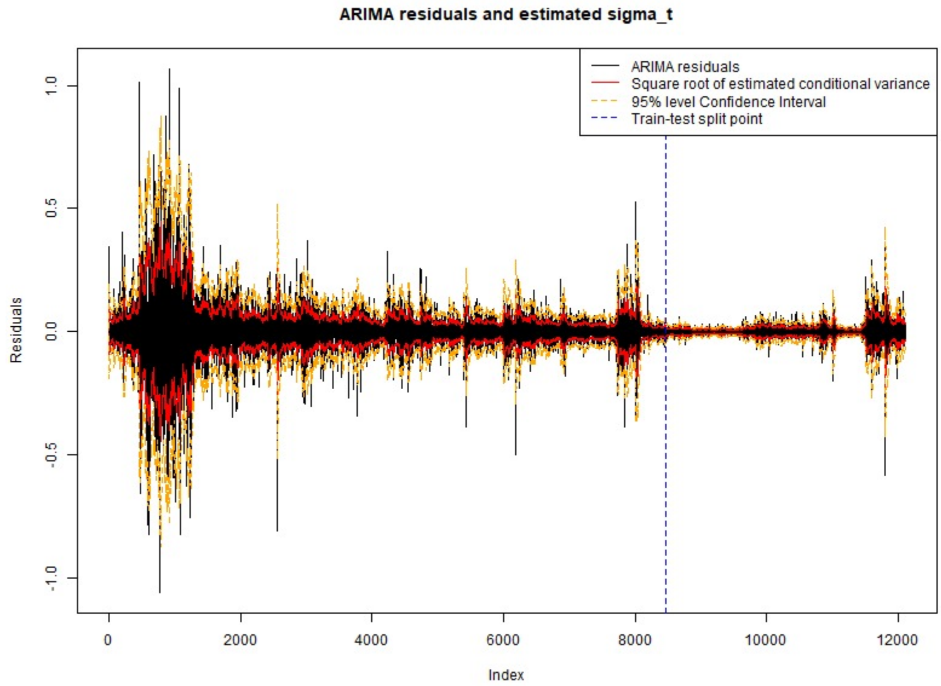

Figure 8: Residuals of the ARIMA(5,1,5) model (solid black), split between train and test set (dotted blue), square root of the estimated conditional variance (solid red) and the 95% confidence interval (dotted yellow)

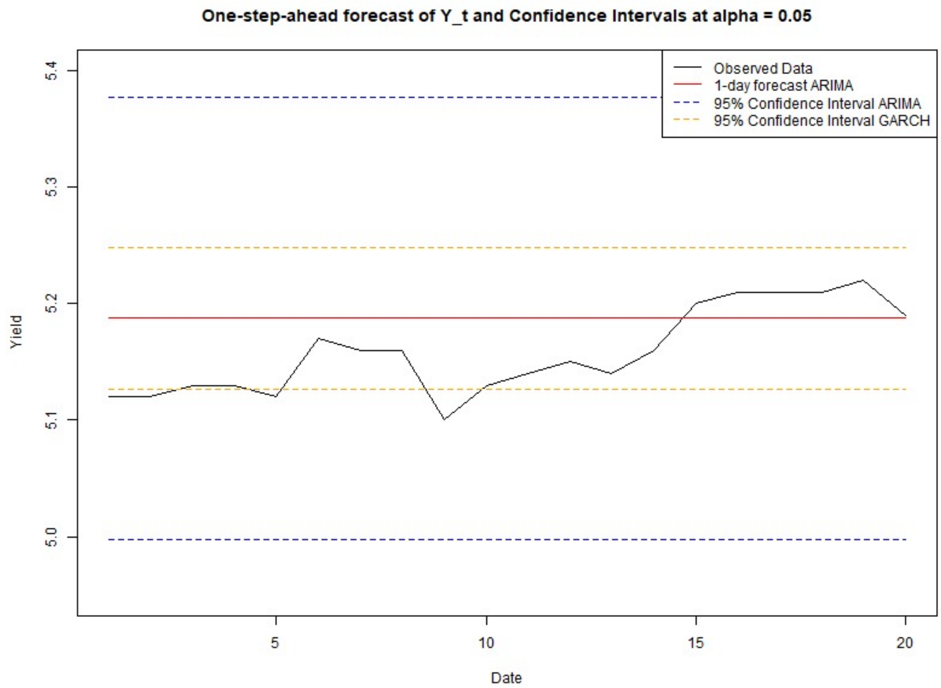

Figure 9: Observed 1-year treasury yield (solid black), one step ahead point estimate of the ARIMA(5,1,5) model (solid red), 95% CI of the ARIMA(5,1,5) model (dotted blue) and 95% CI of the GARCH(1,1) model (dotted yellow)

References

[1] George Box and Gwilym Jenkins. *Time series analysis: forecasting and control.* Holden-Day series in time series analysis and digital processing. Holden-Day, rev. ed edition, 1976.

[2] Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting.* Springer Texts in Statistics. Springer International Publishing.

[3] David A. Dickey and Wayne A. Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. 49(4):1057, 1981.

[4] Francis X Diebold. *Forecasting in Economics, Business, Finance and Beyond.* 2017.

[5] Christian Francq and Jean-Michel Zakoian. *GARCH models: structure, statistical inference and financial applications.* John Wiley & Sons, 2 edition edition.

[6] FRED. Market yield on u.s. treasury securities at 1-year constant maturity, quoted on an investment basis, 2024. Retrieved from Board of Governors of the Federal Reserve System (US).

[7] Refet S. Gürkaynak, Brian Sack, and Jonathan H. Wright. The u.s. treasury yield curve: 1961 to the present. 54(8):2291–2304, 2007.

[8] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer, 2013. OCLC: 1066177334.

[9] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning: with applications in R.* Springer texts in statistics. Springer, corrected at 8th printing edition, 2017.

[10] John M Lachin. Fallacies of last observation carried forward analyses. 13(2):161–168.