Predicting iShares 20+ Year Treasury Bond ETF Price with Regression

Allie Ordentlich, Swaraj Dash, Joshua Petrikat, Elif Turgut

Department of Statistics, University of California Riverside

STAT 170: Regression Analysis

July 25, 2025

Introduction (swaraj):

Research Question Formation: Bond Exchange-Traded Funds (ETFs), such as the iShares 20+ Year Treasury Bond ETF (TLT), are widely used by institutional and retail investors to gain exposure to U.S. Treasury bonds without directly purchasing individual securities. TLT's price is influenced by macroeconomic conditions such as inflation, unemployment, and other economic output. Through this Project, we aim to find out whether macroeconomic indicators can be used in a Linear Regression model to predict near-term movements in TLT's price.

Motivation: Understanding the relationship between macroeconomic factors and TLT prices can be valuable for investors seeking short-term bond ETF positioning strategies. Through this project we also got hands-on experience collecting financial and economic data, cleaning, transforming, and merging datasets, and applying Regression modeling for financial forecasting.

Plan of Execution: Data Collection & Preparation -> Creation of lagged macro variables -> Exploratory Data Analysis (EDA) -> Scatter plot matrix to check relationships between TLT and predictors -> Modeling & Validation -> Stepwise selection to refine predictors -> Final model selected based Diagnostics and Results -> Prediction & Evaluation. Introduction of Variables & Data: Dependent Variable (Y) -> TLT_Avg_1_10, Average TLT ETF price over the 1st–10th trading days each month. Independent Variables (X) are the following. CPI_lag -> Previous month Consumer Price Index (inflation proxy), GDP_lag -> Previous month GDP estimate (economic growth proxy), $Unemp^2$ -> Quadratic transformation of last month's unemployment rate to capture non-linear effects, FedFunds_lag1 -> Lagged Federal Funds Rate (monetary policy indicator). UMCSENT_lag -> Lagged Consumer Sentiment Index (measure of economic optimism)

EDA (Ali):

After combining the various .csv files, we created a monthly dataset with 34 observations and 8 variables, covering from 1/01/2022 through 12/01/2024

We visualized **histograms** for each variable, in order to view shape and spread **(Fig. 1)**: CPI: Left skewed multimodal, with a range of about 280 to 320. Fed Funds: Heavily left skewed, with a range of about 0 to 5%. GDP: Mostly symmetrical with a potential outlier to the left (~ -2), with a range of -2.5 to 6. TLT Avg Price: Right skewed with a range of about 80 to 140, and two potential outliers towards the right of the range. Consumer Sentiment: Fairly symmetrical and multimodal, with some gaps at around 55 and 75, and a range of around 50 to 80 Unemployment Range: Right skewed with many gaps and a very small range (3.4 to 4.2). This likely reflects recent stability in the labor market, but could result in the untransformed variable not being beneficial to a model without interaction.

We also visualized **boxplots (Fig. 2),** which served to confirm some skews and outliers that we hypothesized about earlier. TLT has two outliers towards the higher end (~130 & ~140), Fed Funds has an outlier towards the low end (~0), and GDP has an outlier towards the low end (~ -2).

From the initial EDA, we are able to see that the variables all have vastly different ranges. We researched this topic, and concluded that the different ranges may result in large-scale variables dominating the regression model, and potentially artificial inflation of interaction effects. We decided that the best path would be to standardize the variables, (Fig. 3 & Fig. 4) so each reflects deviation from its mean in standard units. This also makes effect size comparison more meaningful in later modeling.

The next step was to visualize the time progression of each variable (Fig. 5) to try and observe any patterns or trends over time. We can see that CPI, Fed Funds, Consumer Sentiment, and Unemployment slowly rise over time, while TLT appears to slope downwards. This suggests a potential inverse relationship, where the increase of these variables (one or a combination) could result in a decrease in TLT.

We attempted to list key economic and political events that may have affected our variables in order to try and identify some causes of rises and dips. (Fig. 6) Unfortunately, we weren't able to identify any significant rises or dips from this list of events, but given more time we would add more events and dive deeper into the causes of the trends.

Initial Model (Elif):

- lm(formula = TLT_Avg_1_10 ~ CPI_lag + Unemp_lag + GDP_lag + UMCSENT_lag + FedFunds_lag1, data = df_lagged_scaled)

The purpose of this initial regression model is to predict the 10-day average of TLT bond prices using lagged macroeconomic indicators. The predictors used in the initial model are lagged consumer price index (CPI_lag), lagged unemployment rate (Unemp_lag), lagged gross domestic product (GDP_lag), lagged consumer sentiment index (UMCSENT_lag), and lagged federal funds interest rate (FedFunds_lag1). The reason we use these lagged variables is because macroeconomic data is released with a delay, and markets typically react after the information becomes available, making lagged values more realistic and predictive for modeling bond prices like TLT. These variables are chosen based on economic theory, as factors like inflation, interest rates, and GDP growth are known to influence bond market behavior.

We used a scatterplot matrix to visualize the relationship between the 10-day average TLT bond prices and each of the lagged macroeconomic predictors. (Fig. 7) The top triangle displays correlation coefficients, with significance levels denoted by asterisks. We observe strong negative correlations between TLT and several variables, particularly CPI_lag (-0.863) and

FedFunds_lag1 (-0.876), suggesting that higher inflation and interest rates are associated with lower bond prices. Additionally, there is high correlation between some predictors, such as CPI_lag and FedFunds_lag1 (0.888) and CPI_lag and UMCSENT_lag (0.689), which may lead to multicollinearity issues in the regression model — a concern that is can be investigated by the elevated Variance Inflation Factors (VIF) found during diagnostics. We can also see in the lower triangle illustrate linear and nonlinear patterns among the variables. This allowed us to have room to improve the model as there are quadratic trends present. Overall, this plot provides initial insight into which predictors are most related to TLT and highlights potential collinearity between economic indicators.

In the initial model summary (Fig. 8), the model explains 84.3% of the variance in TLT ($R^2 = 0.8429$, Adjusted $R^2 = 0.8149$) and is statistically significant overall ($p < 0.001$). Since this is our initial model, we only have one successful significant predictor, GDP_lag ($p = 0.0287$). This predictor indicates that stronger economic growth tends to lower bond prices, likely due to rising interest rate expectations.

The Variance Inflation Factors (vif) reveal high multicollinearity, especially for CPI_lag (18.29) and FedFunds_lag1 (10.71), suggesting that some predictors are highly correlated and may be redundant. This multicollinearity can inflate standard errors and reduce the reliability of coefficient estimates, highlighting the need to refine the model. Additionally in blue, the Akaike Information Criterion (AIC) provides a baseline for model comparison. Our score for this model is 46.53, which allows us to compare and achieve the lowest AIC value we can.

Model Building, Final Model, and Analysis (Josh):

To first build our full model, we used each unique pairwise combination of the predictors from the initial model for interaction terms. This was done based on prior economic knowledge and different economic theories that alluded to each interaction in the full model being potentially significant. Then we used forward and backward selection to simplify our model, with the null model having no predictors in the model. Our final model ended up giving us (Fig. 9):

$$\begin{aligned}\hat{y} = \; & -0.30356 - 0.80313 \cdot \mathrm{CPI}_{\mathrm{lag}} - 0.26312 \cdot \mathrm{GDP}_{\mathrm{lag}} \\ & + 0.23328 \cdot \mathrm{Unemp2} + 0.13111 \cdot (\mathrm{GDP}_{\mathrm{lag}} \times \mathrm{Unemp2}) + 0.12954 \cdot (\mathrm{CPI}_{\mathrm{lag}} \times \mathrm{GDP}_{\mathrm{lag}})\end{aligned}$$

$$\beta_0 = -0.30356, \quad \beta_1 = -0.80313, \quad \beta_2 = -0.26312, \quad \beta_3 = 0.23328, \quad \beta_4 = 0.13111, \quad \beta_5 = 0.12954$$

$$\mathrm{CPI}_{\mathrm{lag}} = x_1 \qquad \mathrm{GDP}_{\mathrm{lag}} = x_2 \qquad \mathrm{Unemp2} = x_3$$

The AIC of our final model was 30.8590 which was less than the AIC of the initial model which was 46.537; the adjusted R^2 for the final model was 0.8833 which was higher than the adjusted R^2 for the initial model of 0.8149). (Both shown in Fig. 9 and 8 respectfully highlighted in blue).

The first beta coefficient of i = 0 denotes the intercept and its interpretation is not meaningful since this is a standardized model. The beta coefficients from i = 1 to 5 mean that for every single additional increase in the standard deviation for either the associated singular predictor variable or the interaction term, given that all other variables are constant, the ETF prices (TLT_Avg where ETF denotes exchange traded fund) will change by that beta coefficient's number of standard deviations; increasing if positive and decreasing if negative. The predictor variables that were found statistically significant at an alpha = 0.05 significance level were CPI_lag (inflation rate), GDP_lag (growth domestic product growth rate), Unemp2 (unemployment rate squared), and CPI_lag:GDP_lag (interaction between CPI_lag and GDP_lag). (Also see Fig. 9)

Discussing the significant variables, a single increase in the standard deviation of CPI_lag (inflation rate) will lower TLT_Avg (ETF prices) by 0.8031 standard deviations on average. This makes sense as rising inflation often means consumers can afford less which lowers demands for goods and services, in turn lowering ETF prices. A single increase in the standard deviation of GDP_lag (growth domestic product growth rate) will lower TLT_Avg (ETF prices) by 0.2631 standard deviations on average. This corresponds to how an increase in GDP growth rate can sometimes create overheating in the economy or if the increases in GDP growth rate were below expectations even though the GDP growth rate did actually increase. A single increase in the standard deviation of Unemp2 (unemployment rate squared) will increase TLT_Avg (ETF prices) by 0.2333 standard deviations on average. While it might seem strange that with an increase in the squared unemployment rate there will be an increase in ETF prices, this is actually feasible since when unemployment rates are very high, investors will expect a turning point in government economic policy that will help alleviate the bad economy and stimulate economic growth. This possible increase in investor optimism can lead to higher ETF prices, where the squared term describes the non linear shift in expectations. With a single increase in the standard deviation of the interaction of CPI_lag (inflation rate) and GDP_lag (GDP growth rate), the ETF prices will go up by 0.1295 standard deviations on average. This makes sense as when the inflation rate is high and the GDP growth rate is low or negative, this product then captures stagflation, which is where high inflation, slow economic growth, and high unemployment occur all together. Stagflation is a particularly dangerous combination in a nation's economy, so it makes sense why this interaction was deemed statistically significant.

With low VIF scores for each predictor variable (Fig. 10), we determined that multicollinearity is not a large problem for our model. For a significance level of alpha = 0.05, our model passes the Durbin-Watson test for independence with a p value of 0.1475; the model also passes the Shapiro-Wilk Normality test with a p value of 0.6017. (Fig. 10) This result also corresponds with the Q-Q Residuals plot where the points of the residuals make a positive sloped 45 degree straight line. Our Residuals vs Fitted plot shows that the assumption that the expected value of the errors being 0 is met. Our Residuals vs Leverage plot shows that we do not have many problematic leverage points with a high Cook's Distance. While our Scale-Location plot line shows slight curvature, no major funnel shape is found so we can say that the constant variance assumption for the errors is met. (Fig. 11)

Conclusion (Swaraj):

In terms of further improvement, We considered using the previous period's ETF prices (TLT_avg_lagged) as part of the unique pairwise combinations of the predictors from the initial model for interaction terms in the full model since based on economic theories and knowledge it seemed plausible. However, we ultimately declined on that idea since we only wanted to use economic measurements and metrics to predict ETF prices, and figured that including the previous period's ETF prices in the model could be redundant. We also tried to remove outliers and leverage points with our final model, but it resulted in an extremely high R^2 adjusted of 0.95. We figured out that the model without those points has high overfitting and which means it introduced more bias in exchange for necessary variance. In order to account for a good Bias vs Variance trade-off, and predictions for uncertain scenarios, we decided to not remove the high leverage points, not to mention that we already had a low amount of observations to begin with. The Scale-Location plot of the final model without the leverage points and outliers had a higher line suggesting a slightly worse fit, which was an additional reason to forgo removing the outliers and leverage points in the final model. (Fig. 12).

Impact of Results and Answering Research Question: The results from the final model confirms that macroeconomic conditions, particularly CPI, are significant short-term influencers of Treasury bond ETF prices during the beginning of the month. For investors, this means macroeconomic signals can be leveraged for tactical positioning in bond ETFs such as TLT, complementing existing interest-rate focused investment strategies. For aspiring analysts, it demonstrates how economic theory can be integrated with Regression modeling to use broad indicators for financial forecasting.

Predictions:

	Month <dbl>	Observed <dbl>	Predicted <dbl>	Lwr_95 <dbl>	Upr_95 <dbl>
1	1	86.52833	84.96644	74.27327	95.6596
2	2	89.10000	81.39760	67.29479	95.5004

In terms of predictions, we could only predict the TLT price average from the 1st to the 10th of January and February only because of the lack of GDP data to predict March and April. The predicted price for January is $84.966 which is close to the observed price of $86.528. However, February prediction was off by around $8. However, the observed value was well within the 95% confidence interval.

Future Improvements: If we had more time, we could have done some further adjustments. We could have included more historical observations that would have allowed for more stable coefficient estimates, reduced variance, and improved the generalization of the model. Furthermore, while we used an unemployment squared term to capture nonlinearity, we could have also explored additional polynomial terms or interaction effects manually. This could improve the model's ability to detect structural changes in macroeconomic conditions. Also, major events like Federal Reserve policy shifts, geopolitical shocks and categorical market conditions such as recession and expansion could be integrated as dummy variables or one-hot encoded features. Including these qualitative variables in the regression can help capture sudden shifts not explained by continuous macroeconomic indicators alone.

Appendix:

Fig. 1

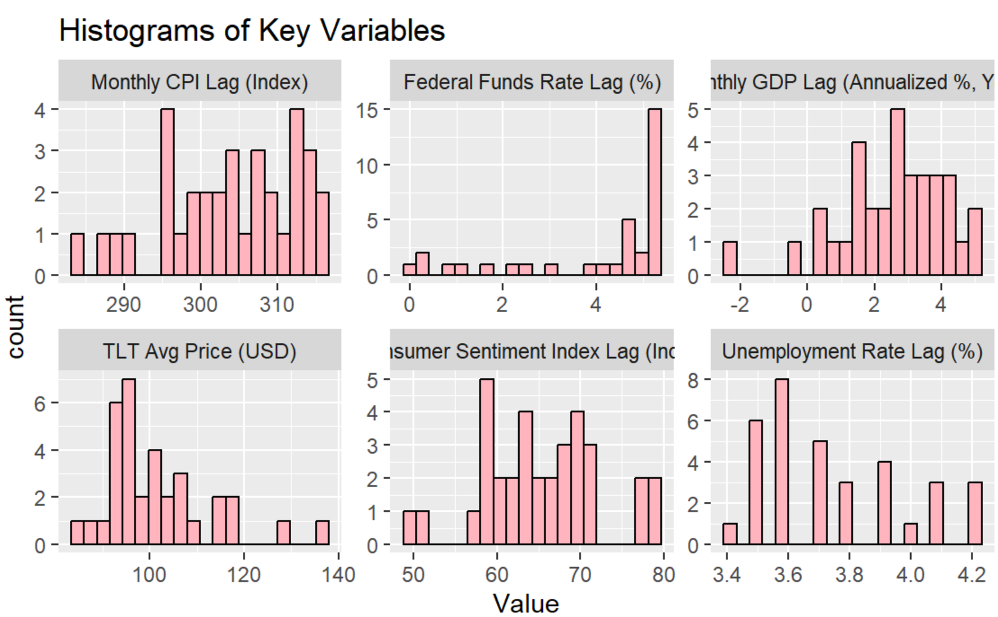


Fig. 2

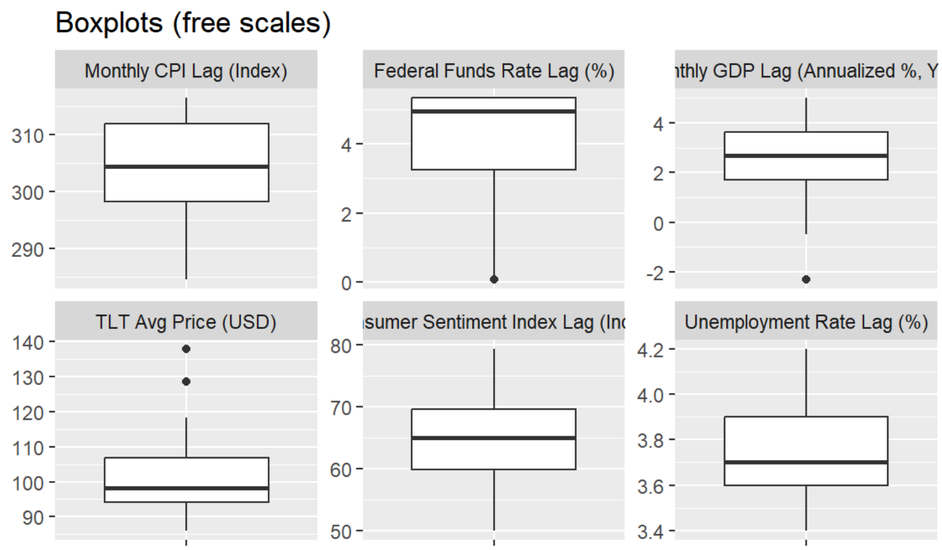

Boxplots (free scales)
Monthly CPI Lag (Index)
Federal Funds Rate Lag (%)
nthly GDP Lag (Annualized %, Y
TLT Avg Price (USD)
nsumer Sentiment Index Lag (Inc
Unemployment Rate Lag (%)
310
300
290
4
2
0
4
2
0
-2
140
130
120
110
100
90
80
70
60
50
4.2
4.0
3.8
3.6
3.4

Fig. 3

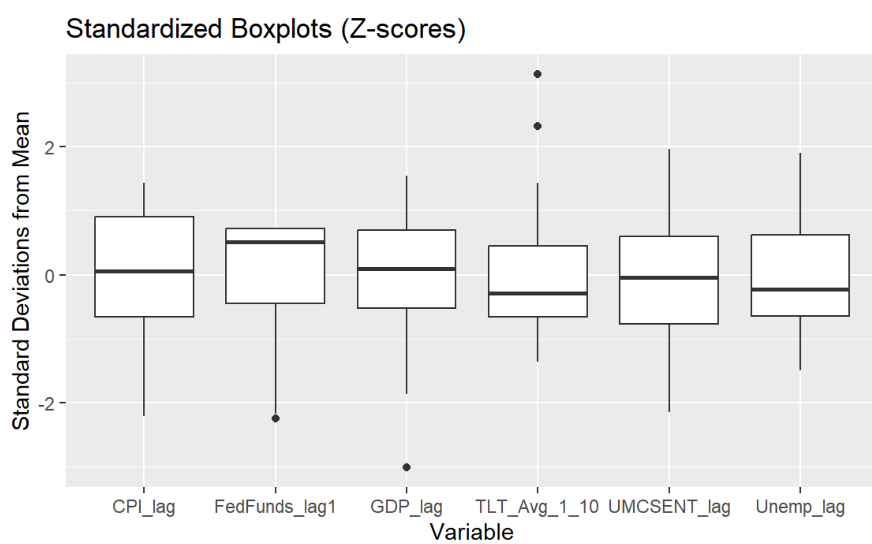

Standardized Boxplots (Z-scores)
Standard Deviations from Mean
2
0
-2
CPI_lag
FedFunds_lag1
GDP_lag
TLT_Avg_1_10
UMCSENT_lag
Unemp_lag
Variable

Fig. 4

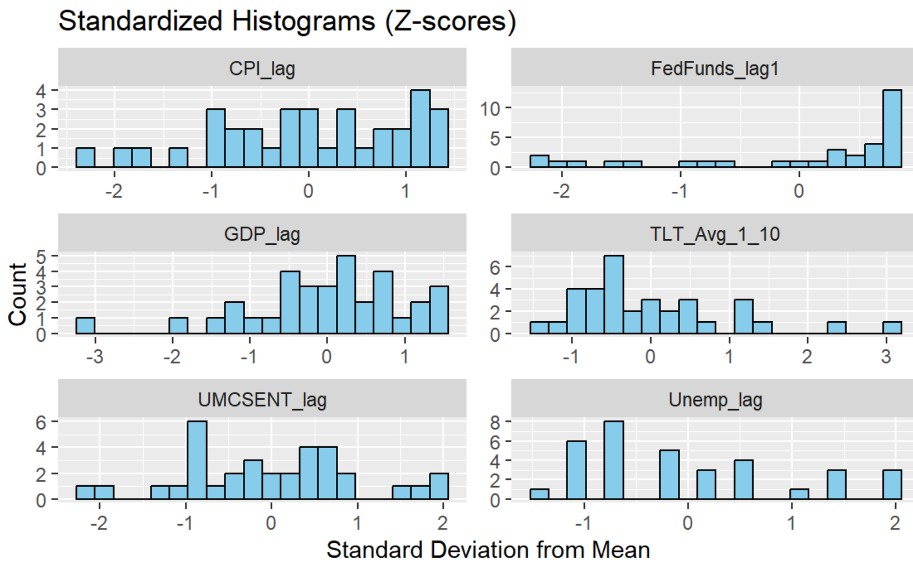

Standardized Histograms (Z-scores)
CPI_lag
FedFunds_lag1
GDP_lag
TLT_Avg_1_10
UMCSENT_lag
Unemp_lag
Count
Standard Deviation from Mean

Fig. 5

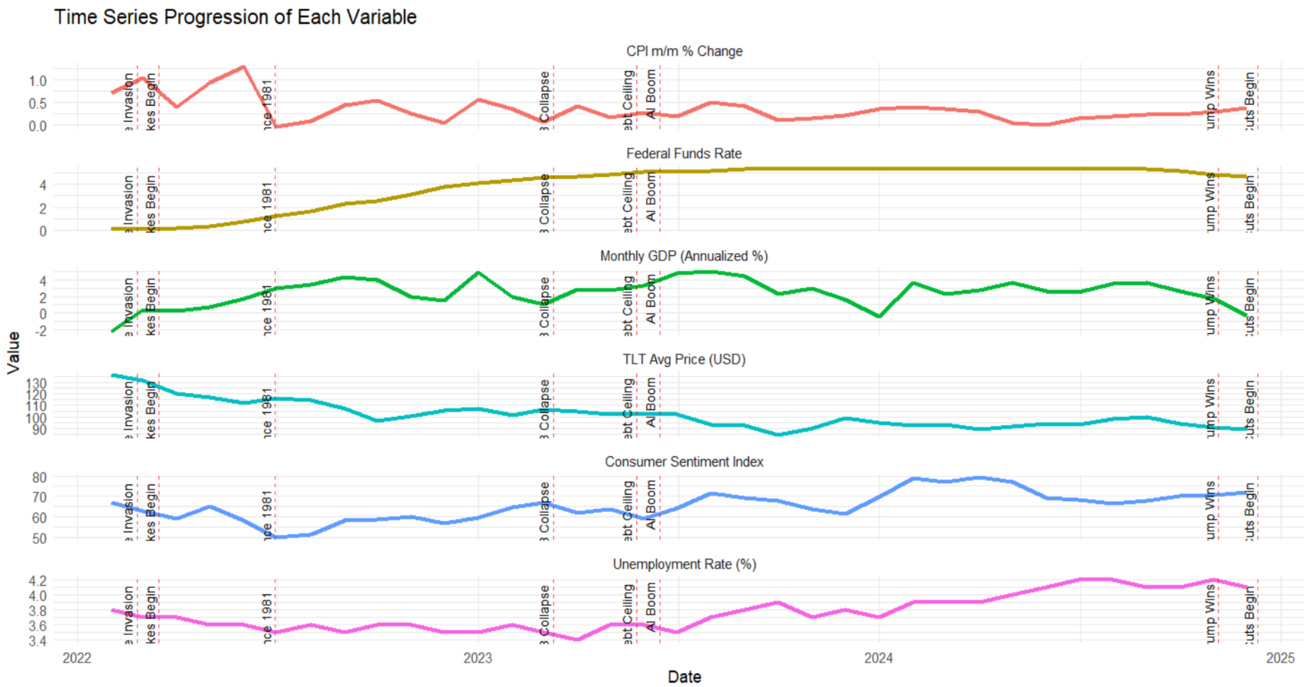


Fig. 6 Significant Events

```r
# Potential Event Dates
event_dates <- as.Date(c(
  "2022-03-16",  # Fed Rate Hike begins (first hike)
  "2022-02-24",  # Russia invades Ukraine
  "2022-06-30",  # Inflation peaks
  "2023-03-10",  # SVB collapse
  "2023-05-25",  # Debt ceiling crisis peaks
  "2023-06-15",  # AI boom takes off (Nvidia surge post-earnings)
  "2024-11-05",  # Trump wins election
  "2024-12-11"   # First rate cut
))
```

Fig. 7 Scatterplot of Predictor Variables

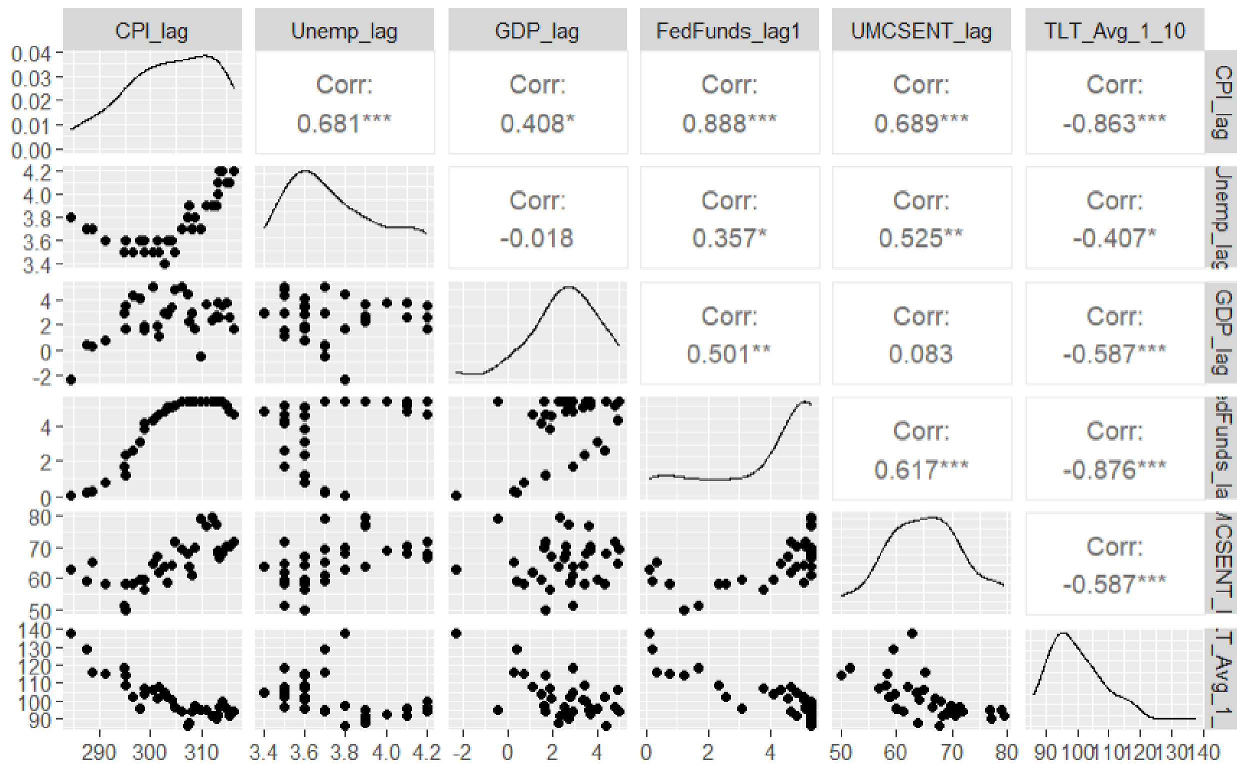


Fig. 8 Initial Model Summary with VIF scores

```
Call:
lm(formula = TLT_Avg_1_10 ~ CPI_lag + Unemp_lag + GDP_lag + UMCSENT_lag +
    FedFunds_lag1, data = df_lagged_scaled)

Residuals:
     Min       1Q   Median       3Q      Max
-0.93448 -0.18671 -0.02169  0.28750  0.62460

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.839e-15  7.379e-02   0.000   1.0000
CPI_lag       -5.940e-01  3.203e-01  -1.854   0.0742 .
Unemp_lag      1.176e-01  1.604e-01   0.733   0.4697
GDP_lag       -2.201e-01  9.542e-02  -2.306   0.0287 *
UMCSENT_lag   -7.839e-02  1.095e-01  -0.716   0.4800
FedFunds_lag1 -2.317e-01  2.451e-01  -0.945   0.3527
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4303 on 28 degrees of freedom
Multiple R-squared:  0.8429,	Adjusted R-squared:  0.8149
F-statistic: 30.05 on 5 and 28 DF,  p-value: 1.964e-10

[1] 46.537
      CPI_lag     Unemp_lag       GDP_lag   UMCSENT_lag FedFunds_lag1
    18.286656      4.588533      1.623160      2.137804     10.709170
```

Fig. 9 Full Model Summary

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.70095 -0.18507  0.04526  0.20116  0.70772

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     -0.30356    0.08769  -3.462  0.00174 **
CPI_lag         -0.80313    0.07928 -10.131 7.22e-11 ***
GDP_lag         -0.26312    0.09693  -2.714  0.01124 *
Unemp2           0.23328    0.06527   3.574  0.00130 **
GDP_lag:Unemp2   0.13111    0.10049   1.305  0.20265
CPI_lag:GDP_lag  0.12954    0.06144   2.108  0.04408 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3417 on 28 degrees of freedom
Multiple R-squared:  0.901,	Adjusted R-squared:  0.8833
F-statistic: 50.94 on 5 and 28 DF,  p-value: 3.386e-13

[1] 30.85904
```

Fig. 10 Vif Summaries of the Significant Variables and Normality and Independence Assumption Tests

```
        CPI_lag         GDP_lag          Unemp2  GDP_lag:Unemp2 CPI_lag:GDP_lag
       1.776772        2.656318        1.378998        1.775850        1.986421

	Durbin-Watson test

data:  final_model
DW = 1.849, p-value = 0.1475
alternative hypothesis: true autocorrelation is greater than 0


	Shapiro-Wilk normality test

data:  residuals(final_model)
W = 0.97469, p-value = 0.6017
```

Fig. 11 Residual Plots of the Final Model

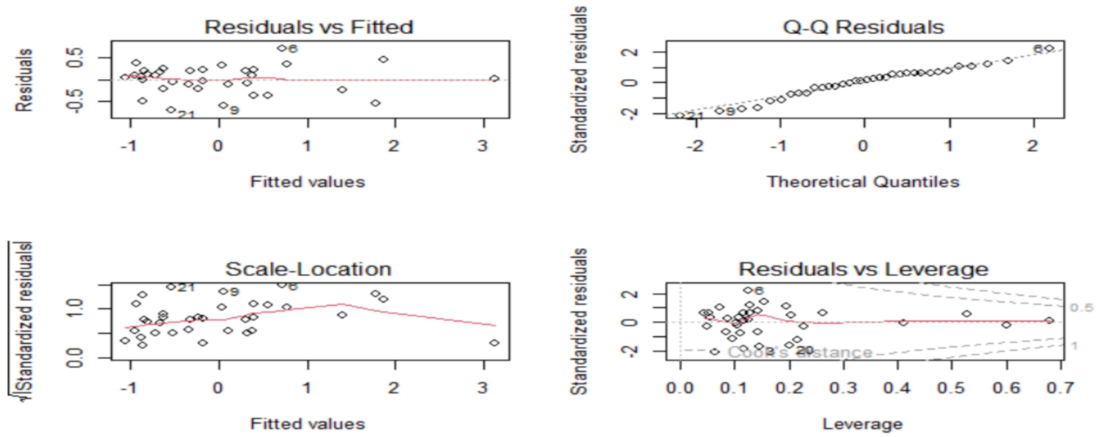
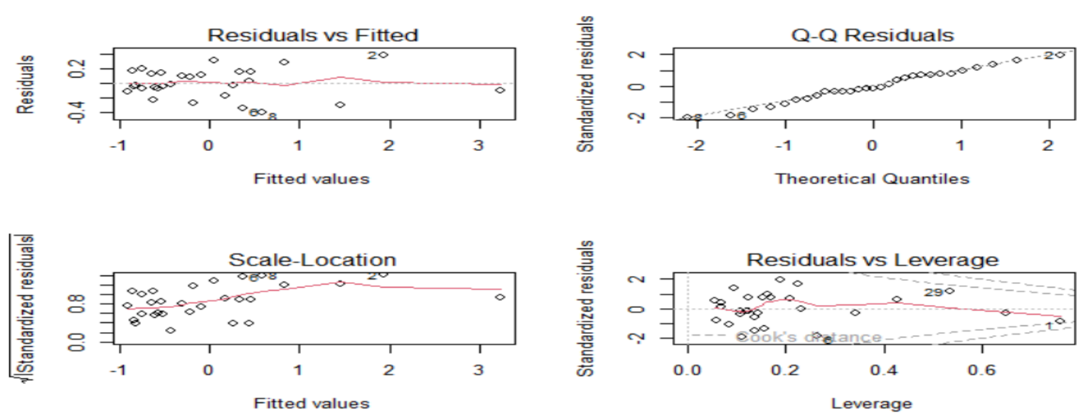


Fig. 12 Residual Plots of the Final Model with Outliers and Leverage Points Removed



Work Cited

Li, Keqing. “Predicting ETF Prices Using Linear Regression.” *BCP Business & Management CMAM*, vol. 36, 2023, pp. 25–31.

Hyndman, Rob J., and George Athanasopoulos. *Forecasting: Principles and Practice*. 2nd ed., OTexts, 2018, Chapter 5, *Time Series Regression Models*.