

Macroeconomic Drivers of S&P 500 Returns: A Boruta-Selected Regression Analysis

Emil Blaignan, Robert Sellman, Sean Eizadi

2025-04-30

Contents

1. Variable Selection	1
2. Descriptive Analysis (Preview)	4

1. Variable Selection

1a. Quantitative Predictors via Boruta

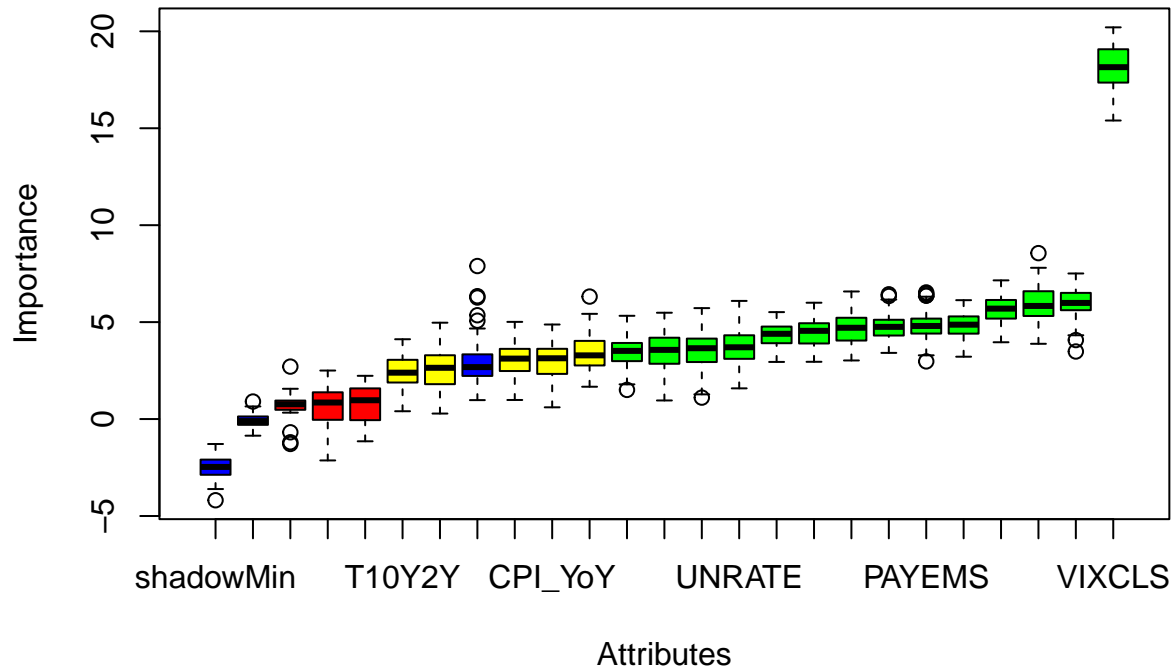
We applied the Boruta algorithm to identify macroeconomic variables most strongly associated with monthly S&P 500 returns. Boruta uses a random forest classifier to assess feature importance, comparing real variables against randomly permuted “shadow” features.

```
# Load data
cleaned_data <- read_csv("data/processed_data.csv")

# Prepare Boruta data
boruta_data <- cleaned_data %>%
  select(-date, -USREC, -tightening) %>%
  select(where(is.numeric))

# Run Boruta
set.seed(123)
boruta_results <- Boruta(monthly_return ~ ., data = boruta_data, doTrace = 1, maxRuns = 100)

# Plot Boruta results
plot(boruta_results)
```



```
# Extract selected attributes
```

```
important_vars <- getSelectedAttributes(boruta_results, withTentative = FALSE)
important_vars
```

```
## [1] "GS10"      "TB3MS"      "BAA"        "AAA"        "FEDFUNDS"
## [6] "CPIAUCSL"  "PPIACO"     "INDPRO"     "UNRATE"     "PAYEMS"
## [11] "M2SL"      "VIXCLS"     "close_price" "INDPRO_YoY"
```

From the confirmed important variables, we retained the **top five** based on importance scores:

1. **VIXCLS** – CBOE Volatility Index (investor risk sentiment)
2. **AAA** – Yield on AAA-rated corporate bonds (long-term credit conditions)
3. **PAYEMS** – Nonfarm payroll employment
4. **M2SL** – M2 money stock
5. **UNRATE** – Unemployment rate

These variables reflect key macroeconomic themes: sentiment, credit, labor, and monetary conditions. All are grounded in theory as plausible drivers of equity returns.

1b. Selection of Factor Variables

We also evaluated two binary indicators:

- **USREC**: NBER recession indicator (1 = recession)
- **tightening**: Dummy for Fed tightening (1 = FEDFUNDS > lagged value)

To assess predictive power, we ran two-sample t-tests:

```
# Recession t-test
recession_test <- t.test(
  cleaned_data$monthly_return[cleaned_data$USREC == 1],
  cleaned_data$monthly_return[cleaned_data$USREC == 0]
)

# Tightening t-test
tightening_test <- t.test(
  cleaned_data$monthly_return[cleaned_data$tightening == 1],
  cleaned_data$monthly_return[cleaned_data$tightening == 0]
)

recession_test

##
## Welch Two Sample t-test
##
## data: cleaned_data$monthly_return[cleaned_data$USREC == 1] and cleaned_data$monthly_return[cleaned_
## t = -1.2482, df = 28.149, p-value = 0.2222
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.05308317 0.01287935
## sample estimates:
## mean of x mean of y
## -0.014345378 0.005756531

tightening_test

##
## Welch Two Sample t-test
##
## data: cleaned_data$monthly_return[cleaned_data$tightening == 1] and cleaned_data$monthly_return[cle
## t = -1.3109, df = 300.94, p-value = 0.1909
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.016512500 0.003308699
## sample estimates:
## mean of x mean of y
## 0.0001948923 0.0067967930
```

Neither test produced statistically significant results at conventional levels ($p = 0.22$ for USREC; $p = 0.19$ for tightening). Nevertheless, we **retained both factors** based on theoretical relevance. Recession periods and monetary policy shifts are widely accepted as important structural influences on asset returns. Their effects may emerge in multivariate settings or through interaction terms.

2. Descriptive Analysis (Preview)

Before model estimation, we conduct exploratory analysis on the **five selected quantitative predictors** and **two factor variables**. This includes:

- Distributional plots (histograms, density curves, boxplots)
- Correlation diagnostics (correlation matrix, scatterplots)
- Outlier checks and transformation considerations (e.g., log or squared terms)

This step ensures variables are well-behaved and suitable for regression modeling. Code and figures will follow in Section 2.