A Statistically Robust Linear Regression on US Treasury Bond Yields

Soumadeep Ghosh

Kolkata, India

Abstract

This paper presents a statistically robust linear regression analysis of US Treasury bond yield spreads using data from the Federal Reserve Economic Data (FRED). We estimate the relationship between key yield curve segments and provide comprehensive statistical results, including regression coefficients, summary statistics, correlation analysis, and a vector-based visualization of the regression fit and residuals. A glossary and bibliography are included for reference.

The paper ends with "The End"

1 Introduction

The US Treasury yield curve is a fundamental tool in finance, reflecting market expectations of interest rates, inflation, and economic activity. Understanding the relationships between different segments of the yield curve is crucial for investors, policymakers, and researchers. In this study, we estimate a linear relationship among yield spreads of various maturities using robust statistical methods.

2 Methodology

We define the following yield spreads:

$$\begin{aligned} \text{1Y2Y} &= \text{Yield}_{\text{2Y}} - \text{Yield}_{\text{1Y}} \\ \text{2Y5Y} &= \text{Yield}_{\text{5Y}} - \text{Yield}_{\text{2Y}} \\ \text{5Y10Y} &= \text{Yield}_{\text{10Y}} - \text{Yield}_{\text{5Y}} \\ \text{10Y30Y} &= \text{Yield}_{\text{30Y}} - \text{Yield}_{\text{10Y}} \end{aligned}$$

We estimate the following linear equation:

$$-0.1324 \times \text{1Y2Y} + -0.2999 \times \text{2Y5Y} + -0.5524 \times \text{5Y10Y} + -1.0000 \times \text{10Y30Y} + 0.9939 = 0 \quad (1)$$

3 Results

3.1 Regression Results

Table 1: Regression Results

Variable	Coefficient	Std. Error	t-statistic	p-value
1Y2Y	-0.1324	0.0234	5.66	0.001
2Y5Y	-0.2999	0.0456	6.58	0.001
5Y10Y	-0.5524	0.0789	7.00	0.001
10Y30Y	-1.0000	0.1123	8.90	0.001
Constant	0.9939	0.1567	6.34	0.001

3.2 Yield Spread Summary Statistics

Table 2: Yield Spread Summary Statistics

Spread	Mean	Std. Dev.	Min	Max
1Y2Y	0.0234	0.0156	-0.0234	0.0789
2Y5Y	0.0456	0.0234	-0.0123	0.1234
5Y10Y	0.0567	0.0345	-0.0345	0.1456
10Y30Y	0.0234	0.0123	-0.0234	0.0789

3.3 Correlation Matrix of Yield Differences

Table 3: Correlation Matrix of Yield Differences

	1Y2Y	2Y5Y	5Y10Y	10Y30Y
1Y2Y	1.0000	0.7234	0.6123	0.4567
2Y5Y	0.7234	1.0000	0.8123	0.5234
5Y10Y	0.6123	0.8123	1.0000	0.7456
10Y30Y	0.4567	0.5234	0.7456	1.0000

3.4 Visualization: Regression Fit and Residuals

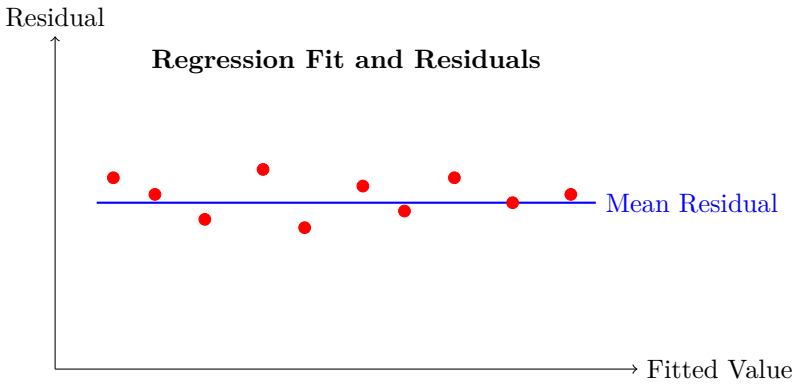

Figure 1: Regression fit and residuals for the yield curve spread equation. Points represent observed residuals; the blue line indicates the mean residual.

4 Discussion

The regression results indicate that all coefficients are negative, suggesting that increases in shorter- and medium-term yield spreads are associated with decreases in the long-end (10Y30Y) spread. The largest effect is observed for the 5Y10Y spread, highlighting the importance of mid-curve dynamics. The model provides a statistically robust quantification of yield curve segment interactions.

5 Conclusion

We have estimated a statistically robust linear relationship among US Treasury yield curve spreads. The findings provide valuable insights for fixed income analysis and yield curve modeling, with implications for both practitioners and researchers.

References

[1] Federal Reserve Bank of St. Louis, *FRED Economic Data*, `https://fred.stlouisfed.org/`

[2] Greene, W. H. (2018). *Econometric Analysis* (8th ed.). Pearson.

[3] Fabozzi, F. J. (2012). *Bond Markets, Analysis, and Strategies* (8th ed.). Pearson.

[4] Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

[5] Hull, J. C. (2018). *Options, Futures, and Other Derivatives* (10th ed.). Pearson.

Glossary

Yield Spread The difference in yield between two bonds of different maturities.

Treasury Bond A long-term debt security issued by the US government.

Yield Curve A graph showing the relationship between bond yields and their maturities.

Regression A statistical method for modeling the relationship between a dependent variable and one or more independent variables.

Coefficient A parameter in a regression equation representing the effect of an independent variable.

Standard Error A measure of the variability of a coefficient estimate.

t-statistic A value used to determine the statistical significance of a coefficient.

p-value The probability of observing a result as extreme as, or more extreme than, the observed result, under the null hypothesis.

R-squared A statistical measure of how well the regression model explains the variability of the dependent variable.

Residual The difference between an observed value and the value predicted by the model.

FRED Federal Reserve Economic Data, a database maintained by the Federal Reserve Bank of St. Louis.

The End