Multiple Regression Analysis of Bond Interest Rates

Posted on March 24, 2019

Executive Interpretation

This multiple regression estimates how four economic factors are associated with bond interest rates. The estimated model is:

$$\hat{Y} = -4.022 + 0.8791X_1 + 0.4653X_2 - 0.2932X_3 + 0.4917X_4$$

where Y is the bond interest rate, X_1 is economic growth, X_2 is the unemployment rate, X_3 is the political climate index, and X_4 is the inflation rate. The model explains approximately 97.1% of the sample variation in bond interest rates, and the overall F-test is statistically significant. At the conventional 5% significance level, economic growth, unemployment, and inflation are statistically significant within the fitted model, while the political climate coefficient is not. These findings describe conditional associations in this dataset. They do not establish that changes in the predictors cause changes in bond rates, nor do they guarantee the same relationship in another period, market, or model specification.

Data and Research Question

The analysis asks whether economic growth, unemployment, political climate, and inflation help explain variation in bond interest rates. Multiple regression is appropriate when the objective is to estimate the association between one dependent variable and several predictors while holding the other included predictors constant. The central equation can be written:

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \varepsilon$$

The intercept β_0 represents the fitted value when all predictors equal zero. Each slope coefficient represents the expected change in Y associated with a one-unit change in that predictor, conditional on the others remaining fixed. The error term ε represents influences not included explicitly in the model.

Estimated Regression Equation

The estimated coefficients from the supplied output are:

Intercept = −4.022

Economic growth = 0.8791

Unemployment rate = 0.4653

Political climate = −0.2932

Inflation rate = 0.4917

The fitted equation is therefore:

Predicted bond rate = −4.022 + 0.8791(Economic Growth) + 0.4653(Unemployment) − 0.2932(Political Climate) + 0.4917(Inflation).

The equation should be reported with units. A “one-unit” change means one percentage point if a variable is recorded in percentage points, but it could mean a one-unit change in an index for political climate. Without definitions and units, the economic magnitude cannot be interpreted fully.

Economic Growth Coefficient

The economic growth coefficient is 0.8791. Holding unemployment, political climate, and inflation constant, a one-unit increase in the economic growth variable is associated with an estimated 0.8791-unit increase in the bond interest rate. If both variables are measured in percentage points, this would be interpreted as approximately 0.879 percentage points.

The supplied t-statistic is 13.936 and the p-value is extremely small. This provides strong evidence against the null hypothesis that the population coefficient on economic growth equals zero within this model. Statistical significance does not demonstrate economic causality. Growth may be correlated with monetary policy, expectations, credit conditions, or omitted factors that also influence bond yields.

Unemployment Coefficient

The unemployment coefficient is 0.4653. Conditional on the other included variables, a one-unit increase in unemployment is associated with an estimated 0.4653-unit increase in the bond interest rate. The t-statistic of approximately 6.602 and the reported p-value indicate statistical significance at conventional levels.

The positive sign may appear surprising because higher unemployment is often associated with weaker activity and lower policy rates. Regression coefficients are conditional, however. The sign reflects the relationship after holding the other included variables constant in this particular sample. It

may be influenced by the period studied, inflation dynamics, credit risk, government borrowing, or correlation among predictors. A surprising sign should prompt investigation rather than automatic rejection.

Political Climate Coefficient

The political climate coefficient is −0.2932. Holding the other variables constant, a one-unit increase in the political climate index is associated with a 0.2932-unit reduction in the predicted bond interest rate. The sign cannot be interpreted substantively until the index is defined. If larger values represent greater political stability, the negative sign might indicate lower borrowing costs under greater stability. If larger values represent greater political risk, the substantive interpretation would be reversed.

The reported t-statistic is approximately −1.468 and the p-value is about 0.147. At the 5% significance level, the coefficient is not statistically distinguishable from zero. The appropriate conclusion is not that political climate has no effect in reality. The sample provides insufficient evidence of an independent linear association after controlling for the other predictors.

Inflation Coefficient

The inflation coefficient is 0.4917. Holding economic growth, unemployment, and political climate constant, a one-unit increase in inflation is associated with an estimated 0.4917-unit increase in the bond interest rate. The t-statistic is approximately 3.678 and the p-value is less than 0.001, indicating statistical significance at the 5% level.

A positive relationship between inflation and nominal bond yields is economically plausible because investors may require compensation for expected loss of purchasing power. The regression uses the supplied inflation measure, but bond yields are forward-looking. Expected inflation may be more relevant than realized current inflation. A richer model could distinguish inflation expectations from observed inflation.

Interpretation of the Intercept

The estimated intercept is −4.022. It is the predicted bond interest rate when all four predictors equal zero. The intercept is mathematically necessary for the fitted model unless theory justifies constraining it. Its economic interpretation may be weak if zero growth, zero unemployment, zero inflation, and a zero political-climate index are outside the observed data or are not meaningful simultaneously.

A negative intercept does not mean that the model predicts negative rates under normal observed

conditions. Interpretation should focus on fitted values within the range of the data. Extrapolating far outside the sample can produce unrealistic predictions.

Model Fit

The R-squared value is approximately 0.9713. This means that about 97.13% of the sample variation in bond interest rates is explained by the fitted linear relationship with the four included predictors. The adjusted R-squared is approximately 0.9695, which remains extremely high after accounting for the number of predictors.

A high R-squared is not proof that the model is correct. It does not demonstrate causality, absence of omitted variables, accurate future prediction, or valid standard errors. Time-series variables can move together because of common trends and produce a high R-squared even when the relationship is not economically stable. Diagnostics remain essential.

Overall F-Test

The overall F-statistic is approximately 546.2, with a very small significance probability. The joint null hypothesis is:

H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$.

The alternative is that at least one slope coefficient differs from zero. The very small p-value leads to rejection of the joint null at the 5% level. The predictors, considered together, provide statistically significant explanatory power relative to an intercept-only model. (National Institute of Standards and Technology, 2012)

This test does not mean every predictor is individually significant. The political climate coefficient illustrates the distinction: the model can be highly significant overall even when one individual slope is not.

Individual Hypothesis Tests

For each predictor, the conventional two-sided hypothesis is:

H_0: $\beta_i = 0$

H_1: $\beta_i \neq 0$.

At $\alpha = 0.05$, economic growth, unemployment, and inflation reject the null based on the supplied p-values. Political climate does not. Decisions should ideally be reported with coefficient estimates,

standard errors, confidence intervals, and p-values rather than only labels of significant and insignificant.

Confidence Intervals

Approximate 95% confidence intervals can be constructed as:

coefficient ± critical t × standard error.

For economic growth, the estimate is 0.8791 with a standard error of about 0.0631, yielding a narrow interval clearly above zero. For unemployment, 0.4653 with a standard error of about 0.0705 also yields an interval above zero. Political climate, −0.2932 with a standard error near 0.1997, produces an interval that crosses zero. Inflation, 0.4917 with a standard error near 0.1337, produces an interval above zero.

Confidence intervals communicate magnitude and precision more clearly than p-values alone. An insignificant estimate may still be compatible with economically meaningful positive or negative values when its interval is wide.

Standardized Versus Unstandardized Effects

The coefficients shown are unstandardized. They cannot be compared directly to decide which predictor is “most important” when variables use different units. Economic growth’s coefficient is numerically larger than inflation’s, but a typical one-unit change may represent a very different movement.

Standardized coefficients, partial R-squared, incremental explanatory power, and domain knowledge can help compare influence. Even standardized measures should not be interpreted causally without an appropriate research design.

Multicollinearity

Economic growth, unemployment, and inflation may be correlated. Multicollinearity occurs when predictors contain overlapping information. It does not bias ordinary least-squares coefficients automatically, but it can increase standard errors, make coefficients unstable, and produce unexpected signs.

Variance inflation factors should be reported. Large VIF values would suggest that individual coefficient interpretations are sensitive to the predictor set. Pairwise correlations are useful but insufficient because multicollinearity can involve combinations of several variables.

Residual Diagnostics

Regression assumptions concern residuals, not simply the raw variables. Analysts should plot residuals against fitted values and each predictor to look for curvature, changing variance, and influential observations. A random cloud supports the linear and constant-variance assumptions, while systematic patterns suggest misspecification.

A normal Q-Q plot can assess whether residuals are approximately normal for small-sample inference. Normality is less central for unbiased coefficient estimation than for exact finite-sample tests. Severe outliers can distort both coefficients and tests.

Heteroskedasticity

Heteroskedasticity means that the variance of errors changes with predictors or fitted values. This is common in economic data. Ordinary least-squares coefficient estimates can remain unbiased under exogeneity, but conventional standard errors become unreliable.

Robust standard errors should be considered. Tests such as Breusch-Pagan or White can supplement graphical analysis. The original output should not be treated as final evidence until the standard-error assumptions have been checked.

Autocorrelation

If observations are ordered over time, residuals may be serially correlated. Interest rates and macroeconomic variables often show persistence. Positive autocorrelation can make conventional significance tests overly optimistic.

The Durbin-Watson statistic, residual autocorrelation function, or other time-series tests can assess dependence. Corrective approaches may include lagged variables, generalized least squares, or heteroskedasticity-and-autocorrelation-consistent standard errors. The appropriate method depends on the data-generating process.

Stationarity and Spurious Regression

Time-series regressions require particular caution. If bond rates, inflation, or other variables contain stochastic trends, unrelated series can produce high R-squared and statistically significant coefficients. Analysts should inspect plots and apply stationarity tests such as augmented Dickey-Fuller where appropriate.

If nonstationary variables share a long-run equilibrium, cointegration methods may be suitable. If not,

differencing or alternative specifications may be required. The exceptionally high R-squared of 0.971 makes these checks especially important if the data are a time series.

Omitted Variable Bias

The model includes four predictors, but bond yields may also depend on central-bank policy, expected future short rates, maturity, credit risk, sovereign risk, liquidity, fiscal deficits, exchange rates, global interest rates, and risk appetite. If an omitted factor is correlated with an included predictor, the estimated coefficient may capture part of the omitted effect.

Adding every available variable is not the solution because overfitting and multicollinearity can increase. Variables should be selected using theory and the purpose of the model.

Endogeneity and Causality

Macroeconomic variables influence one another. Interest rates can affect unemployment, growth, and inflation, while those variables also influence policy rates and bond yields. This simultaneity creates endogeneity.

Ordinary multiple regression estimates conditional association unless strong exogeneity assumptions hold. A causal study may require instrumental variables, natural experiments, structural models, or time-series identification strategies. The word “effect” should therefore be used cautiously in the current descriptive analysis.

Functional Form

The model assumes linear additive relationships. The impact of inflation may differ when inflation is low versus extremely high, and political risk may matter only after a threshold. Interactions may also exist; for example, inflation’s association with bond rates could differ in periods of weak growth.

Scatterplots, partial residual plots, polynomial terms, splines, or theory-driven interactions can evaluate nonlinearity. Alternative models should be validated rather than chosen solely because they improve in-sample R-squared.

Political Climate Measurement

The political variable needs detailed documentation. A categorical rating, expert index, or numerical stability scale carries different interpretation. Researchers should state the source, range, direction, and frequency of measurement.

Measurement error can attenuate or distort the coefficient. Political conditions may also affect yields through nonlinear risk premiums or interact with fiscal conditions. The insignificant coefficient may reflect poor measurement rather than absence of political influence.

Prediction

The fitted equation can produce a predicted bond interest rate when values for all four predictors are supplied. A useful prediction should include uncertainty. A confidence interval estimates the mean response for a set of predictor values, while a prediction interval for an individual future observation is wider because it includes residual variability.

Predictions should stay within or near the observed range. Using the model for an economy with inflation or unemployment far outside the sample would be extrapolation and may generate unreliable values.

Out-of-Sample Validation

In-sample fit can be misleading. A model may explain historical observations extremely well but forecast poorly. Data can be divided into training and test periods, or rolling-origin validation can be used for time series.

Forecast accuracy measures such as root mean squared error and mean absolute error should be reported. Candidate models can be compared on future-like data rather than R-squared alone.

Economic Significance

Statistical significance asks whether an estimate is distinguishable from zero under model assumptions. Economic significance asks whether the magnitude matters. A coefficient of 0.49 on inflation may be highly meaningful or relatively modest depending on typical inflation variation and the bond market.

Analysts should calculate predicted changes for realistic scenarios. For example, the predicted effect of a two-percentage-point inflation change is twice the one-unit coefficient if linearity holds. These scenarios make the model more useful to decision-makers.

Model Stability

Relationships among macroeconomic variables can change after financial crises, policy regime shifts, wars, or structural reforms. A coefficient estimated over one period may not hold over another.

Rolling regressions, structural-break tests, and interaction terms with regime indicators can assess stability. The model should be re-estimated when new data become available rather than treated as permanent.

Reporting Standards

A complete regression report should include the sample size, observation period, variable definitions and units, data sources, coefficient estimates, standard errors, t-statistics, p-values, confidence intervals, R-squared, adjusted R-squared, F-statistic, and diagnostic tests. It should explain missing-data treatment and any transformations.

The original essay provides coefficient and ANOVA information but needs these contextual details. Reproducible analysis also requires the underlying dataset or a clear description of how it can be obtained.

Conclusion

The estimated regression suggests that economic growth, unemployment, and inflation are positively associated with bond interest rates after controlling for the other included variables, while political climate has a negative but statistically insignificant coefficient. The model explains approximately 97.1% of sample variation, and the joint F-test strongly rejects the hypothesis that all slopes are zero. These are strong in-sample statistical results. (Wooldridge, 2020)

The results should not yet be interpreted as causal or permanently predictive. Macroeconomic regression requires checks for multicollinearity, heteroskedasticity, serial correlation, stationarity, omitted variables, endogeneity, nonlinear relationships, influential observations, and out-of-sample performance. The political-climate index also needs a clear definition. A defensible final analysis combines numerical results with economic theory, diagnostic evidence, uncertainty, and transparent limitations.

References

1. National Institute of Standards and Technology. *NIST/SEMATECH e-Handbook of Statistical Methods: Multiple Linear Regression*. 2012.
2. Wooldridge, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 7th ed., Cengage, 2020.
3. Gujarati, Damodar N., and Dawn C. Porter. *Basic Econometrics*. 5th ed., McGraw-Hill, 2009.
4. Federal Reserve Bank of St. Louis. *FRED Economic Data*.
5. Statsmodels. *Regression Diagnostics and Specification Tests*.