CFA INSTITUTE RESEARCH FOUNDATION / **BRIEF**

MACROECONOMIC DRIVERS OF STOCKS AND BONDS

FRIEDRICH BAUMANN
ABDOLREZA NAZEMI
FRANK J. FABOZZI, CFA

CFA Institute Research Foundation

MACROECONOMIC DRIVERS OF STOCKS AND BONDS

FRIEDRICH BAUMANN
ABDOLREZA NAZEMI
FRANK J. FABOZZI, CFA

CFA Institute
Research
Foundation

Statement of Purpose

The CFA Institute Research Foundation is a not-for-profit organization established to promote the development and dissemination of relevant research for investment practitioners worldwide.

Photo credit: Pete Rowbottom / Moment / Getty Images

Print ISBN: 978-1-952927-55-3

Ebook ISBN: 978-1-952927-56-0

CONTENTS

CFA Institute

PROFESSIONAL LEARNING QUALIFIED ACTIVITY
This publication qualifies for 0.5 PL credits under the guidelines of the CFA Institute Professional Learning Program.

MACROECONOMIC DRIVERS OF STOCKS AND BONDS

Friedrich Baumann
Karlsruhe Institute of Technology, School of Economics and Management, Karlsruhe, Germany

Abdolreza Nazemi
Karlsruhe Institute of Technology, School of Economics and Management, Karlsruhe, Germany

Frank J. Fabozzi, CFA
Carey School of Business, Johns Hopkins University, Baltimore

Introduction

Investors may be at an inflection point that defines a new era. Monetary policy has long served as a predominant stabilizing factor for asset prices, and many of today's younger practitioners have built their entire professional careers on the back of asset appreciation boosted by liquidity and low interest rates provided by loosened central bank policies. While inflation expectations gradually increased in the aftermath of the markets' COVID-19 recovery, inflation in the United States spiked to levels unseen for more than 40 years in 2022. Inflation uncertainty seems to have gotten ahead of the Federal Reserve Board (Fed), which increased interest rates seven times in 2022, having previously raised rates even against expectations and its own guidance. At the same time, a growing number of analysts are voicing concerns that the Fed's recently started unwinding of its near USD9 trillion balance sheet could trigger an economy-wide recession.

Today's tightening monetary policy may even be the dawn of a long-term paradigm shift away from decades of increasing liquidity. Will there be a soft landing, or is the United States headed toward stagflation? For investors, answering this question is of paramount importance. Having benefited from rising valuation levels, low yields, and a longstanding negative correlation of equity and bond returns that allowed portfolio hedging, investors may need to rethink what they have learned about portfolio composition. As macroeconomic uncertainties mount, equity and bond returns have recently converged. This pattern is not new, yet it has not been seen for more than 20 years.

Practitioners must now assess how to deal with this changing paradigm. Hence, accurate estimation of the stock–bond return correlation helps decision makers navigate the uncertainty around this correlation and helps them improve the allocation of resources and assets.

In this report, we use a machine learning approach to identify the most meaningful macroeconomic drivers of the stock–bond return correlation, relying on Stocks, Bonds, Bills, and Inflation (SBBI®) data available from the CFA Institute Research Foundation Investment Data Alliance and the Federal Reserve Bank of St. Louis's Federal Reserve Economic Data (FRED). Our methodology selects the most relevant variables and ranks them by importance for predicting the correlation, providing a useful tool for reducing the uncertainty of the correlation between stock and bond returns.

Some of the technical terminology in this report relating to machine learning and statistical methods is nicely described in Simonian (2024).

Literature and Theoretical Background

Stock–Bond Return Correlation

According to Ibbotson and Harrington (2021), the correlation of annual returns of long-term US corporate bonds and US stocks was positive on average during the period 1926–2020, with wide variations over time. To improve our understanding of the stock–bond return correlation, theoretical and empirical researchers have devoted various studies to it. Campbell, Pflueger, and Viceira (2020), for example, used a consumption-based model developed by Campbell and Cochrane (1999) to model macroeconomic dynamics together with the stock–bond return correlation and explain the correlation's switch from positive to negative in 2001 with the inflation–output gap.

Other studies have used macroeconomic factors to explain the variation in stock–bond correlations. Li (2002) argued that the stock–bond correlation is mainly driven by uncertainty about expected inflation. Ilmanen (2003) examined the impact of growth, inflation, volatility, and monetary policy on the stock–bond return correlation. The focus of Connolly, Stivers, and Sun (2005) was on stock market uncertainty as a factor, while Baele, Bekaert, and Inghelbrecht (2010) used a dynamic factor model to analyze the co-movement between the returns on stocks and bonds. Novel studies, such as Wu, DiCiurcio, Yeo, and Wang (2022), also employ standard machine learning models to predict stock–bond correlations.

Predicting bond returns has recently gained increased attention from academics, leading to several innovations. Ludvigson and Ng (2009) demonstrated that bond returns vary with macroeconomic fundamentals. Kelly, Palhares, and Pruitt (2023) used a conditional and time-varying factor model to describe corporate bond returns. Huang, Jiang, and Tong (2022) used real-time macro variables for nonlinear bond return prediction.

Macroeconomic Drivers of the Stock–Bond Correlation

Various other models attempt to explain the correlation of stock and bond returns using macroeconomic data. Ilmanen, Maloney, and Ross (2014) studied the sensitivity of stocks and bonds to various macroeconomic regimes: growth, inflation, real yields, volatility, and illiquidity. They showed that stock and bond returns are positively related under certain regimes and negatively related under others. Ermolov (2022) used macroeconomic shocks to estimate stock and bond returns, and Campbell et al. (2020) analyzed the macroeconomic drivers of equity and bond risks.

Cieslak and Pang (2021) identified economic shocks with distinct effects on stocks and bonds, whereas Tao, Wang, Wang, and Wu (2022) used the concept of economic policy uncertainty for corporate bond return prediction. The implications of the findings of these studies are that to construct stock–bond portfolios today, investors need to navigate uncertainty about macroeconomic and market conditions, as well as their influence on both stock and bond returns.

The existing research provides explanations of stock and bond returns under varying macroeconomic regimes. In this report, we offer practitioners a framework to use a large amount of available macroeconomic data to identify the most important macroeconomic factors related to the stock–bond return correlation. Although Ilmanen et al. (2014) considered only five macroeconomic variables that are difficult to predict, we use a large amount of data from which we select to derive a small set of explanatory variables. Hence, we provide a valuable extension that is ready for practitioners to use. We identify the most meaningful macroeconomic variables for predicting the stock–bond correlation using machine learning methods to derive practical implications for balancing portfolios under given macroeconomic conditions.

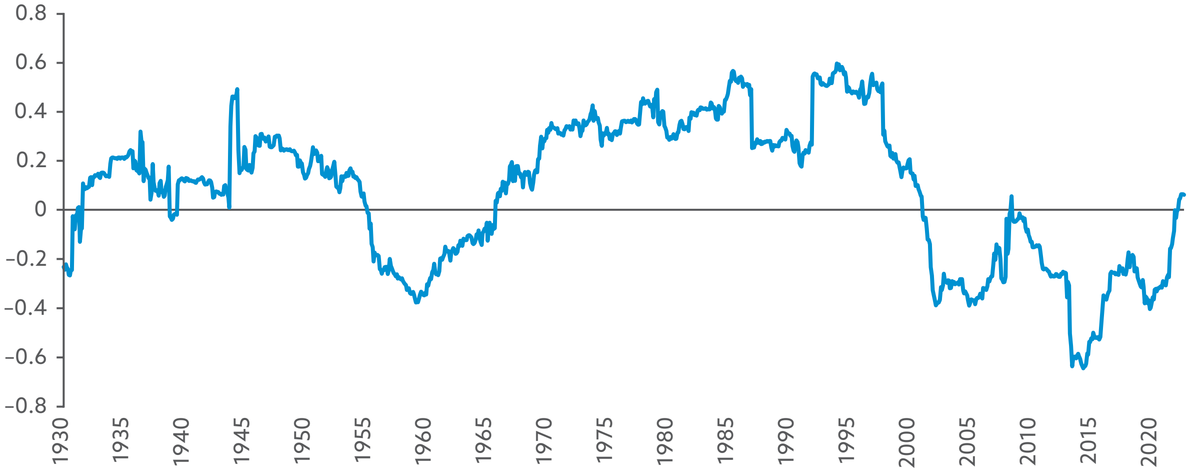
Because we use data over a long time horizon, we can produce predictions over a long time period, making it possible to derive generalizations about expected future stock–bond return correlations under given macroeconomic conditions. By doing so, our analysis shows the ability of certain macroeconomic factors to serve as early indicators of the stock–bond correlation, thereby giving practitioners valuable help in guiding investment decisions about portfolios of stocks and bonds.

Data

We compiled a dataset of stock–bond return correlations relying on stock and bond return data obtained from Morningstar Direct through the Investment Data Alliance of CFA Institute Research Foundation. The return data that we use are from the SBBI database. In particular, we collected monthly total return observations for the US large-cap stock market (Ibbotson SBBI US Large-Cap Stocks [Total Return]) and the monthly total returns of long-term government bonds (Ibbotson SBBI US Long-term [20-Year] Government Bonds [Total Return]). With the monthly return data, we created monthly rolling five-year stock–bond return correlations.

Exhibit 1 presents the five-year rolling stock–bond correlation from 1930 to 2023. The figure shows that the correlation remained either largely positive or largely negative over

Exhibit 1. Stock–Bond Return Correlations, December 1930–December 2023



Source: Drawn by the authors using Ibbotson SBBI® data.

persistent time frames. That is, from the early 1930s until the mid-1950s, the stock–bond return correlation was positive, followed by about a decade of negative correlation. From the mid-1960s to the early 2000s, the correlation was positive again, after which it switched to a predominantly negative correlation that persisted for the most recent 20 years. In late 2022, the return correlation rose again to a slightly positive level.

We used macroeconomic data to predict the stock–bond correlation. We retrieved 134 macroeconomic indicators from FRED, provided by McCracken and Ng (2016). FRED, which provides monthly updates for a large range of US macroeconomic indicators from 1959 to 2024, is available from the website of the Federal Reserve Bank of St. Louis.[1] We matched the stock–bond return correlation data with the 127 available macroeconomic indicators. Although the return data start in 1926, we needed to limit the time frame to 1959–2023 because of the availability of the macroeconomic data. Overall, we created a comprehensive dataset ranging from 1959 to 2023, which we used for our empirical analysis.

Methodology

In Exhibit 1, we presented the trailing five-year rolling windows, but for out-of-sample prediction of time-series variation, we used a forward-looking approach. This method uses distinct historical data windows as inputs, separate from the forward-looking time window. Additionally, we used various statistical and machine learning methodologies, described in detail in the following subsections.

Stability Selection

To select the most meaningful macroeconomic variables for predicting the stock–bond correlation, we use a stability selection technique that is based on the least absolute shrinkage and selection operator (LASSO), similar to the methodology used by Nazemi and Fabozzi (2018). LASSO shrinks coefficients to zero for the least significant variables via regularization defined by $\sum_{j=1}^{N}\left|\beta_j\right| = \|\beta\|_1$, thereby effectively performing selection. Compared with ridge regression, which uses L2 regularization to shrink coefficients to nonzero values, LASSO removes variables entirely. The following expression estimates the LASSO coefficient:

$$\sum_{i=1}^{N}(y_i - \alpha - \sum_{j=1}^{k}\beta_j x_{ij})^2 + \lambda_1 \sum_{j=1}^{k}\left|\beta_j\right|.$$

However, small perturbations in the data can alter the LASSO variable selection, however, resulting in instability in the selection. Meinshausen and Bühlmann (2010) suggested a stability selection technique to overcome these shortcomings of using LASSO. The stability selection technique performs variable selection via LASSO repeatedly and considers those variables that are most frequently selected as stable. Thus, only variables that are selected more often than a predefined threshold will be included in the final set of selected variables. We implemented the stability selection technique to select the top 10 macroeconomic variables from the large dataset.

[1]Federal Reserve Bank of St. Louis, "FRED-MD and FRED-QD: Monthly and Quarterly Databases for Macroeconomic Research," https://research.stlouisfed.org/econ/mccracken/fred-databases/.

Permutation Importance Ranking

Altmann, Toloşi, Sander, and Lengauer (2010) demonstrated that variable importance derived from random forests is biased, given that categorical variables with many categories are considered more important than others. To rank the macroeconomic variables by importance, we conducted permutation importance ranking, a methodology that identifies the importance of individual variables in explaining the dependent variable. Our approach is closely related to that of Nazemi, Baumann, and Fabozzi (2022), who used stability selection to identify the most meaningful variables from a large set of macroeconomic variables to determine corporate bond recovery rates as key drivers of credit risk.

Initially, a model is trained involving all variables, and model fit is determined. Then, the permutation importance ranking methodology randomly permutes all values of a feature numerous times, keeping all other features constant. By doing so, the methodology breaks the relationship between the explanatory variable and the dependent variable, offering insights into how much the model performance deteriorates on average. The procedure is repeated for each variable separately, thus allowing us to quantify the influence of each variable on model performance. Finally, the explanatory variables are ranked by how much the model fit deteriorates, assigning higher importance to variables that induce a greater performance drop when permuted and ultimately providing insights into the relative importance ranking among the explanatory variables in the model.

Random Forests

We used machine learning via random forests for out-of-sample prediction. Random forests, introduced by Breiman (2001), improve upon the bagging approach by combining many regression trees that rely on random selection of features. The predictions from these trees are then aggregated, typically by averaging, to obtain more precise and stable results. In traditional bagging, individual regression trees are trained solely on random subsets of the data. This approach aims to increase accuracy and reduce prediction variance.

Principal Component Analysis

Given the large number of macroeconomic variables, it is difficult to differentiate among some of these variables that are highly correlated with each other. For example, the dataset contains data series on different aggregation levels, such as levels of industrial production of consumer goods, but also for the subcategories durable and nondurable consumer goods. Although each of these variables captures different dimensions of similar underlying economic information, they remain highly correlated and add to the complexity of modeling and interpretation.

We use the statistical technique of principal component analysis (PCA) to reduce dimensions within our dataset. We first identify thematic groups of variables based on their underlying economic relationships and then derive and isolate the first principal component for each of these groups. Doing so allows us to shrink the size of the dataset, condensing the data to thematic groups and decreasing the overall number of variables, keeping underlying information and linkages. Applying the stability selection technique on the thematic variables derived via PCA allows for simplification while also improving the interpretability and efficiency of the modeling efforts.

Empirical Implementation and Results

We first estimate the stock–bond return correlation in an in-sample setting, selecting the most meaningful macroeconomic variables as explanatory variables using a stability selection technique. We further consider the principal components of thematic groups of these variables. Then, we create out-of-sample prediction models with random forests, using the selected variables. Finally, we rank these variables by importance to gain insights into which variables are the most important indicators for explaining the stock–bond return correlation.

At each point in time, we consider the correlation of stock–bond returns over the next five years as the dependent variable. Therefore, our data span the period from February 1959 to June 2018, reflecting the return correlation up to June 2023. For our analysis, we define in- and out-of-sample time periods. The in-sample time period spans 70% of the data, ranging from 1959 to 2000, and the test set is the remainder of the data.

In-Sample Linear Regressions

We define the baseline model as a linear model using all observations from 1959 to 2000 in an in-sample setting. The model includes macroeconomic variables that are often associated with the stock–bond return correlation: the inflation rate, industrial production, the Federal funds rate, and the slope of the yield curve (i.e., 10-year Treasury bill yield minus the 1-year Treasury bill yield). We adjust standard errors for heteroskedasticity and use Newey–West standard errors that are robust to autocorrelation using 59 lags.

The in-sample results presented in **Exhibit 2** show that about 46% of the variation in the stock–bond return correlation is accounted for by four variables: inflation rate, industrial production output, federal funds rate, and yield curve slope. However, only the federal funds rate and the slope of the yield curve are found to be significant. Only if we do not adjust standard errors as described previously do we find that all four variables are highly significant.

Exhibit 2. In-Sample Linear Regression of the Stock–Bond Correlation on Commonly Used Stock–Bond Correlation Drivers

Variable	Coefficient
Inflation rate	–0.0295
Industrial production	–0.0786
Federal funds rate	0.1832***
Yield curve slope	0.1293***
Adjusted R^2	0.460

Notes: Standard errors are adjusted for heteroskedasticity and autocorrelation using 59 lags. Explanatory variable significance is as follows: *** denotes 1%.

Exhibit 3. In-Sample Linear Regression of the Stock–Bond Correlation on Selected Macroeconomic Variables

Variable	Coefficient (4 variables)	Coefficient (10 variables)
S&P 500	–0.1219***	–0.0983***
Employees in ND goods mfg.	0.1086***	0.0908***
Unemployment < 5 weeks	0.0816***	0.0592**
New orders consumer goods mfg.	–0.0074	–0.0219
Housing starts West		–0.0318*
Nonborrowed reserves		0.0879***
Canada/United States FX rate		–0.0555*
New private housing starts		–0.0224
Housing starts South		–0.0034
New private housing authorized Northeast		0.0358***
Adjusted R^2	0.779	0.865

Notes: Standard errors are adjusted for heteroskedasticity and autocorrelation using 59 lags. Explanatory variable significance is as follows: *** denotes 1%, ** denotes 5%, and * denotes 10%.

Next, we consider the stability selection technique to select macroeconomic variables, rather than relying on economic intuition. This approach allows us to extract the most important explanatory variables from the large set of macroeconomic factors. The stability selection iterates LASSO 200 times, and we selected the top 4 and the top 10 variables by selection frequency, respectively. In **Exhibit 3** we present the in-sample regression results. By determining the model specification through stability selection, our data-driven approach improves the model fit from the adjusted R^2 of the baseline model of 0.46 to an adjusted R^2 of 0.779.

The selection technique considers the S&P 500 Index, the number of employees in nondurable (ND) goods manufacturing, the number of unemployed people less than five weeks, and the number of new orders for consumer goods manufacturing as the most meaningful in explaining the stock–bond return correlation. In our linear model, we find that these variables, except for the number of new orders for consumer goods manufacturing, are highly significant in explaining the correlation.

In a separate specification, we let the stability selection technique select the top 10 variables, which further improves the adjusted R^2 to 0.865. Now, we find that housing starts in the Western regions, new private housing authorizations in the Northeast, nonborrowed reserves, and the foreign exchange (FX) rate of Canadian dollars to US dollars are also significant.

In the next step, we use PCA for dimensionality reduction. Because the individual relationships uncovered in the linear regressions of the selected variables are difficult to rationalize, we group the 127 macroeconomic variables within 18 thematic categories and derive for each group the first principal component of all variables subsumed within the group. We then let the stability selection technique select the 4 and 10 most meaningful variables, respectively, for explaining the stock–bond correlation.

As shown in **Exhibit 4**, the stability selection technique selects the four themes of commodity prices, financial markets, foreign exchange, and manufacturing orders. Only financial markets and manufacturing orders, however, are significant for estimating the stock–bond correlation in the linear in-sample setting. When we select the 10 most important themes, we further find that economic and consumer sentiments are also significant.

Interestingly, the selection of themes partially corresponds to the selection of individual macroeconomic variables in Exhibit 3, given that financial markets and leading indicators for manufacturing are selected. However, the two sentiment themes were not uncovered in the individual macroeconomic variable selection in Exhibit 3. Furthermore, it is important to note that the coefficients and signs of the thematic variables cannot be reliably interpreted

Exhibit 4. In-Sample Linear Regression of the Stock–Bond Correlation on Selected First Principal Components of Thematic Groups of Macroeconomic Variables

Variable	Coefficient (4 Variables)	Coefficient (10 Variables)
Commodity prices	0.0403	0.0518
Financial markets	–0.4457***	–0.5143***
Foreign exchange	–0.0087	–0.0564
Manufacturing orders	0.3379**	0.4744***
Interest rates		–0.0194
Economic sentiment		–0.0822**
Employment		–0.0569
Work hours		0.0195
Consumer sentiment		0.0250**
Housing market		–0.0115
Adjusted R^2	0.705	0.779

Notes: Standard errors are adjusted for heteroskedasticity and autocorrelation using 59 lags. Explanatory variable significance is as follows: *** denotes 1% and ** denotes 5%.

because of the transformation of underlying macroeconomic data by applying principal component analysis.

Out-of-Sample Random Forest Regressions

We now report the out-of-sample predictions using linear and random forest regressions. We create the regression models on the training data and test their performance for predicting the stock–bond return correlation using the test data. Therefore, this correlation is regressed on the four factors commonly used for explaining stock–bond return correlation (similar to Exhibit 2); the 4 and 10, respectively, selected macroeconomic variables (similar to Exhibit 3); and the 4 and 10, respectively, selected themes derived from the macroeconomic variables (similar to Exhibit 4).

As a performance measure, we compute the out-of-sample root mean square error (RMSE). For creating random forest models, we follow a suggestion by Breiman (2001) and use one-third of the explanatory variables for each tree. We further define two more essential settings: the minimum size of leaves and the number of trees in the forest. To determine the ideal settings, we run a grid search in fivefold cross-validation on the training dataset.

In **Exhibit 5**, we present the out-of-sample prediction results. As this exhibit demonstrates, the linear regression performs worse than random forests for all model specifications, producing RMSEs in the range of about 0.59 to 0.73, whereas the random forest regressions produce RMSEs of 0.17–0.18. Thus, the random forest regressions are likely to better capture nonlinearities in the data. Further, note that the best out-of-sample linear regression model is the model that relies on the four factors commonly used for stock–bond return correlation estimation, whereas the 10 selected macroeconomic factors perform only slightly worse. The selected principal components derived from macroeconomic themes perform the worst, with RMSEs of 0.65 using 4 themes and 0.73 using 10 themes.

Exhibit 5. RMSE Results of Out-of-Sample Linear Regression and Out-of-Sample Random Forest Regression

	RMSE	
	Linear Regression	**Random Forests**
Four common factors	0.5872	0.1694
Four selected factors	0.6271	0.1695
Ten selected factors	0.5932	0.1836
Four selected themes (PCA)	0.6533	0.1687
Ten selected themes (PCA)	0.7270	0.1667

Note: The stock–bond correlation is regressed on the four factors commonly used for stock–bond correlation (similar to Exhibit 2); the 4 and 10, respectively, selected macroeconomic variables (similar to Exhibit 3); and the 4 and 10, respectively, selected themes derived from the macroeconomic variables (similar to Exhibit 4).

Nevertheless, when we use random forests, the selected principal components derived from macroeconomic themes perform best. With random forests, however, there are no large deviations in the RMSEs produced by the different models, except for the model using 10 selected macroeconomic variables, which has the worst RMSE among the models using random forest techniques.

Altogether, using random forests strongly improves the out-of-sample prediction performance of our models, and the performance of the models using selected macroeconomic factors or selected principal components derived from thematic groups of macroeconomic factors is overall comparable.

Variable Importance

Finally, we investigate the variable importance of the selected macroeconomic factors and the selected principal components derived from thematic groups of macroeconomic factors. Having previously shown that using these selected variables yields prediction performance similar to that of factors supported by the theoretical and empirical literature, we are now interested in identifying the importance of these variables for determining the stock–bond correlation. Therefore, we apply permutation importance ranking following Altmann et al. (2010). By doing so, we rank the importance of the selected variables for out-of-sample prediction via random forests.

We determine the permutation importance both for the selected macroeconomic variables and for the selected first principal components of thematic groups of macroeconomic variables. After obtaining the permutation importance scores, we scale them to the interval [0,1].

In **Exhibit 6**, we show the permutation importance for the selected macroeconomic variables. As the exhibit reveals, the four most important variables are leading indicators of housing

Exhibit 6. Permutation Importance Ranking of Selected Macroeconomic Variables

Rank	Variable	Permutation Importance
1	New private housing authorized Northeast	1.00
2	Employees ND goods mfg.	0.92
3	New orders consumer goods mfg.	0.92
4	Housing starts South	0.90
5	Nonborrowed reserves	0.83
6	Canada/United States FX rate	0.82
7	New private housing starts	0.56
8	Unemployment less than 5 weeks	0.56
9	Housing starts West	0.45
10	S&P 500	0.00

Exhibit 7. Permutation Importance Ranking of Selected First Principal Components of Thematic Groups of Macroeconomic Variables

Rank	Variable	Permutation Importance
1	Manufacturing orders	1.00
2	Consumer sentiment	0.36
3	Commodity prices	0.35
4	Interest rates	0.33
5	Financial markets	0.33
6	Economic sentiment	0.33
7	Work hours	0.32
8	Foreign exchange	0.32
9	Employment	0.05
10	Housing market	0.00

(new private housing authorized in the Northeast and new housing starts in the South) and manufacturing (nondurable goods manufacturing and the number of new orders for consumer goods manufacturing). The other variables drop in importance. Interestingly, the S&P 500 appears unimportant, although it is significant in the in-sample linear regression in Exhibit 3.

In the next step, we compute the permutation importance ranking for the selected first principal component of thematic groups of macroeconomic variables. We present the results in **Exhibit 7**. Here, we find that the most important group of variables is related to manufacturing orders and bears a considerably higher importance score than all other variables. Data related to consumer sentiment, commodity prices, interest rates, financial markets, economic sentiment, work hours, and foreign exchange follow in importance; however, they do not differ much from each other in importance. Interestingly, the factor related to the housing market appears to be least important, in contrast to the findings in the previous analysis.

Our analysis shows that when using either selected macroeconomic variables or variables derived from thematic groups of macroeconomic variables, macroeconomic data related to manufacturing orders are among the most important predictors of the stock–bond correlation. That is, we find that leading indicators of manufacturing activity serve as important predictors of the stock–bond correlation.

Conclusion

Determining the correlation of stocks and bonds is a key challenge for practitioners in today's markets, given the recent change from negative to positive correlation. In this report, we offer practitioners a framework to use a large amount of available macroeconomic data to identify the factors most impactful on the stock–bond correlation. Previous theoretical and empirical research has demonstrated the critical role of macroeconomic conditions in the formation of the stock–bond return correlation. Today, practitioners need reliable estimation methods to navigate economic volatility and policy adjustments.

By applying a stability selection technique, using selected variables in machine learning models to predict the stock–bond correlation out of sample, and ranking the macroeconomic variables by importance, we find that leading indicators of manufacturing activity are among the most important drivers of the stock–bond return correlation. These findings are underscored by an alternative approach in which we first combine the macroeconomic variables within thematic groups and then derive the first principal component for each group.

Using random forests for out-of-sample prediction improves modeling accuracy by a large fraction compared with linear regression. Nevertheless, the application of random forest models using commonly used stock–bond estimators, selected macroeconomic variables, or selected principal components of macroeconomic themes yields comparable prediction results.

Ultimately, our study demonstrates how machine learning techniques (specifically, stability selection and random forest models) distill the most meaningful information from a vast set of macroeconomic data to identify the most impactful predictors of the stock–bond correlation. Our approach equips practitioners with a toolkit that can be readily applied to gain insights into the stock–bond correlation, supporting decision making in asset allocation and risk management.

REFERENCES

Altmann, A., L. Toloşi, O. Sander, and T. Lengauer. 2010. "Permutation Importance: A Corrected Feature Importance Measure." *Bioinformatics* 26 (10): 1340–47.

Baele, L., G. Bekaert, and K. Inghelbrecht. 2010. "The Determinants of Stock and Bond Return Comovements." *Review of Financial Studies* 23 (6): 2374–2428.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32.

Campbell, J. Y., and J. H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy* 107 (2): 205–51.

Campbell, J. Y., C. Pflueger, and L. M. Viceira. 2020. "Macroeconomic Drivers of Bond and Equity Risks." *Journal of Political Economy* 128 (8): 3148–85.

Cieslak, A., and H. Pang. 2021. "Common Shocks in Stocks and Bonds." *Journal of Financial Economics* 142 (2): 880–904.

Connolly, R., C. Stivers, and L. Sun. 2005. "Stock Market Uncertainty and the Stock–Bond Return Relation." *Journal of Financial and Quantitative Analysis* 40 (1): 161–94.

Ermolov, A. 2022. "Time-Varying Risk of Nominal Bonds: How Important Are Macroeconomic Shocks?" *Journal of Financial Economics* 145 (1): 1–28.

Huang, D., F. Jiang, and G. Tong. 2022. "Real Time Macro Factors in Bond Risk Premium." *SSRN Electronic Journal*. https://ssrn.com/abstract=3107612.

Ibbotson, R. G., and J. P. Harrington. 2021. *Stocks, Bonds, Bills, and Inflation® (SBBI®): 2021 Summary Edition*. Charlottesville, VA: CFA Institute Research Foundation.

Ilmanen, A. 2003. "Stock–Bond Correlations." *Journal of Fixed Income* 13 (2): 55–66.

Ilmanen, A., T. Maloney, and A. Ross. 2014. "Exploring Macroeconomic Sensitivities: How Investments Respond to Different Economic Environments." *Journal of Portfolio Management* 40 (3): 87–99.

Kelly, B. T., D. Palhares, and S. Pruitt. 2023. "Modeling Corporate Bond Returns." *Journal of Finance* 78 (4): 1967–2008.

Li, L. 2002. "Macroeconomic Factors and the Correlation of Stock and Bond Returns." Yale ICF Working Paper No. 02-46.

Ludvigson, S., and S. Ng. 2009. "Macro Factors in Bond Risk Premia." *Review of Financial Studies* 22 (12): 5027–67.

McCracken, M. W., and S. Ng. 2016. "FRED-MD: A Monthly Database for Macroeconomic Research." *Journal of Business and Economic Statistics* 34 (4): 574–89.

Meinshausen, N., and P. Bühlmann. 2010. "Stability Selection." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 72 (4): 417–73.

Nazemi, A., F. Baumann, and F. J. Fabozzi. 2022. "Intertemporal Defaulted Bond Recoveries Prediction via Machine Learning." *European Journal of Operational Research* 297 (3): 1162–77.

Nazemi, A., and F. J. Fabozzi. 2018. "Macroeconomic Variable Selection for Creditor Recovery Rates." *Journal of Banking & Finance* 89 (April): 14–25.

Simonian, J. 2024. *Investment Model Validation: A Guide for Practitioners*. Charlottesville, VA: CFA Institute Research Foundation.

Tao, X., B. Wang, J. Wang, and C. Wu. 2022. "Economic Policy Uncertainty and the Cross-Section of Corporate Bond Returns." *Journal of Fixed Income* 32 (1): 6–44.

Wu, B., K. J. DiCiurcio, B. Yeo, and Q. Wang. 2022. "Forecasting US Equity and Bond Correlation—A Machine Learning Approach." *Journal of Financial Data Science* 4 (1): 76–86.

Named Endowments

CFA Institute Research Foundation acknowledges with sincere gratitude the generous contributions of the Named Endowment participants listed below.

Gifts of at least US$100,000 qualify donors for membership in the Named Endowment category, which recognizes in perpetuity the commitment toward unbiased, practitioner-oriented, relevant research that these firms and individuals have expressed through their generous support of CFA Institute Research Foundation.

Ameritech
Anonymous
Robert D. Arnott
Theodore R. Aronson, CFA
Asahi Mutual Life Insurance Company
Batterymarch Financial Management
Boston Company
Boston Partners Asset Management, L.P.
Gary P. Brinson, CFA
Brinson Partners, Inc.
Capital Group International, Inc.
Concord Capital Management
Dai-Ichi Life Insurance Company
Daiwa Securities
Mr. and Mrs. Jeffrey Diermeier
Gifford Fong Associates
John A. Gunn, CFA
Investment Counsel Association of America, Inc.
Jacobs Levy Equity Management
Jon L. Hagler Foundation
Long-Term Credit Bank of Japan, Ltd.
Lynch, Jones & Ryan, LLC
Meiji Mutual Life Insurance Company
Miller Anderson & Sherrerd, LLP
John B. Neff, CFA
Nikko Securities Co., Ltd.
Nippon Life Insurance Company of Japan
Nomura Securities Co., Ltd.
Payden & Rygel
Provident National Bank
Frank K. Reilly, CFA
Salomon Brothers
Sassoon Holdings Pte. Ltd.
Scudder Stevens & Clark
Security Analysts Association of Japan
Shaw Data Securities, Inc.
Sit Investment Associates, Inc.
Standish, Ayer & Wood, Inc.
State Farm Insurance Company
Sumitomo Life America, Inc.
T. Rowe Price Associates, Inc.
Templeton Investment Counsel Inc.
Frank Trainer, CFA
Travelers Insurance Co.
USF&G Companies
Yamaichi Securities Co., Ltd.

Senior Research Fellows

Financial Services Analyst Association

For more on upcoming CFA Institute Research Foundation publications and webcasts, please visit www.cfainstitute.org/research/foundation.

CFA Institute Research Foundation

Available online at rpc.cfainstitute.org

ISBN 978-1-952927-55-3