Macroeconometrics and Machine Learning Project: Detecting Fiscal Thresholds in Eurozone Convenience Yields

Aleksandra Zachorowska Zhaonan Chu Aditya Srivastava

November 24, 2026

Abstract

This report examines whether Eurozone sovereign bond convenience yields respond nonlinearly to fiscal fundamentals. Using an annual panel of nine countries from 2008–2025 (Germany as the benchmark), we proxy convenience yield spreads with negative CDS spreads relative to Germany and relate them to lagged fiscal variables (debt-to-GDP and the primary surplus spread), controlling for growth and inflation differentials. We compare a linear fixed-effects baseline and a threshold regression to tree-based models (Random Forest and XGBoost). In out-of-sample tests using a time-based split, the tree-based models improve predictive performance over linear benchmarks. Partial dependence results suggest threshold-like behavior: fiscal conditions matter more when debt is high and deficits are large. These findings are descriptive rather than causal, but they indicate that linear models can miss important state dependence in fiscal pricing.

1 Introduction

Government bond convenience yields reflect the non-pecuniary benefits of holding sovereign debt beyond compensation for credit risk. Investors are therefore willing to accept lower returns in exchange for attributes such as liquidity, safety, and the role of government bonds as collateral in financial markets. Convenience yields are thus a key determinant of sovereign borrowing costs and an important object for understanding government bond pricing (Krishnamurthy et al., 2014; Nagel, 2016).

Jiang et al. (2024) document substantial cross-country and time-series variations in convenience yields across Eurozone sovereigns. Using linear panel regressions, they show that fiscal fundamentals – particularly debt-to-GDP and surplus-to-GDP ratios – significantly predict convenience yields even within a currency union. Their analysis establishes that stronger fiscal positions are associated with higher convenience yields and provides a benchmark linear characterization of the relationship between public finances and sovereign bond pricing in the Eurozone.

Crucially, the conceptual framework in Jiang et al. (2024) emphasizes that convenience yields are forward-looking asset prices that depend on both current and expected future fiscal conditions. Using survey-based forecasts, they show that expectations of improving fiscal surpluses are associated with higher convenience yields, consistent with standard asset pricing intuition. Moreover, fiscal conditions affect perceived crisis probabilities, which directly enter the pricing of sovereign bonds. When expectations about fiscal sustainability influence the perceived safety of government debt, small changes in fiscal fundamentals may have limited pricing effects in normal times but disproportionately large effects once fiscal positions deteriorate sufficiently.

This expectation-based mechanism provides a natural rationale for nonlinear and threshold-type relationships between fiscal variables and convenience yields. Models of safe asset demand emphasize that assets perceived as safe can command substantial premia until beliefs about safety shift, at which point prices adjust sharply (Gennaioli et al., 2018). In the context of

sovereign debt, such belief shifts may occur when fiscal conditions approach levels that raise concerns about sustainability or crisis risk. More generally, classic models of fiscal policy with discrete adjustment costs highlight trigger points at which fiscal imbalances are tolerated until they become salient, reinforcing the scope for state-dependent responses to fiscal fundamentals (Bertola and Drazen, 1993). Empirical studies of sovereign bond markets support this view, showing that fiscal variables exert stronger effects on yields when public finances deteriorate or during periods of market stress (Codogno et al., 2003; Ardagna et al., 2007; Égert, 2015).

While traditional econometric approaches can accommodate nonlinearities through threshold or regime-switching models, they typically rely on strong functional-form assumptions and require ex ante restrictions on the number and location of thresholds (Hansen, 1999). In contrast, tree-based machine learning methods offer a flexible, data-driven alternative that can capture nonlinearities and interactions without imposing parametric structure. Through recursive partitioning, these methods naturally identify threshold-type behavior, making them well suited to studying state-dependent fiscal pricing in sovereign bond markets (Athey and Imbens, 2019).

In this project, we extend the analysis of Jiang et al. (2024) by applying Random Forest and XGBoost models to Eurozone sovereign data. Our objectives are threefold: to assess whether the relationship between fiscal ratios and convenience yields deviates from linearity; to identify economically meaningful fiscal thresholds at which markets penalize fiscal weakness more severely; and to compare the predictive performance of machine learning models with standard linear panel regressions. In doing so, we provide new evidence on the nonlinear, expectation-driven pricing of fiscal risk and evaluate the role of machine learning methods as a complement to conventional econometric approaches.

2 Dataset

We construct an annual panel dataset for Eurozone sovereigns following the methodology of Jiang et al. (2024). The sample covers the period 2008–2025, focusing on the post–global financial crisis era during which sovereign risk differentials within the Eurozone became economically meaningful. The analysis includes nine member states: Austria, Belgium, Finland, France, Ireland, Italy, the Netherlands, Portugal, and Spain. Germany serves as the benchmark country throughout the analysis.

Sovereign credit default swap (CDS) data are obtained from Markit. Annual CDS observations are constructed using the last available daily quote through April of each year, ensuring alignment with fiscal data releases. Following Jiang et al. (2024), convenience yield spreads are proxied by the negative CDS spread relative to Germany. This measure aims to capture cross-country differences in the non-credit component of sovereign yields and facilitates comparisons within the currency union.

Fiscal variables are obtained from Eurostat and include the government surplus-to-GDP ratio and the debt-to-GDP ratio. The surplus-to-GDP ratio is expressed as a spread relative to Germany, consistent with the construction of convenience yield spreads and capturing cross-country fiscal positions within the union. In contrast, public debt is measured in levels as a ratio to GDP, reflecting the stock nature of government indebtedness. To mitigate concerns about simultaneity between fiscal outcomes and asset prices, all fiscal variables are lagged by one year prior to merging with CDS data.

We additionally include macroeconomic control variables following Jiang et al. (2024): real GDP growth and inflation. Both variables are constructed as log differences of real GDP and the consumer price index, respectively, and are expressed as spreads relative to Germany. Prior to estimation, control variables are mean-centered to facilitate interpretation and interaction effects.

Table 1 reports summary statistics for the main variables used in the analysis. Convenience yield spreads exhibit substantial dispersion across countries and over time, with a standard deviation that is large relative to the mean and extreme negative values during periods of fiscal

Table 1: Summary Statistics

Variable	N	Mean	Std. Dev.	Min	p25	Median	p75	Max
CY Spread vs Germany (bps)	162	-55.42	111.94	-864.43	-53.92	-17.40	-5.37	15.05
Surplus/GDP Spread vs Germany (%)	162	-2.72	3.56	-27.70	-4.27	-2.65	-1.00	7.00
Debt/GDP (%)	162	88.34	28.41	23.90	65.55	88.15	108.02	154.40
GDP Growth Spread vs Germany (%)	162	0.32	3.10	-7.40	-1.21	-0.07	1.39	20.37
Inflation Spread vs Germany (%)	162	-0.14	0.91	-3.63	-0.66	-0.17	0.48	2.67
CDS Spread vs Germany (bps)	162	55.42	111.94	-15.05	5.37	17.40	53.92	864.43

and financial stress (which can be seen in Figure 1). This heterogeneity indicates that sovereign bonds within the Eurozone differ in their perceived safety and liquidity, particularly in adverse market conditions.

Fiscal fundamentals also display considerable variation. The surplus-to-GDP spread relative to Germany has a negative mean and a wide range, suggesting that many countries run persistent deficits relative to the benchmark while occasionally experiencing sharp fiscal deteriorations. Public debt levels vary substantially across countries, ranging from relatively low-debt sovereigns to highly indebted economies with debt-to-GDP ratios well above the sample average. Taken together, the summary statistics document pronounced cross-country and time-series heterogeneity in both fiscal fundamentals and convenience yield spreads, suggesting that the relationship between fiscal conditions and sovereign bond pricing may differ across fiscal states.

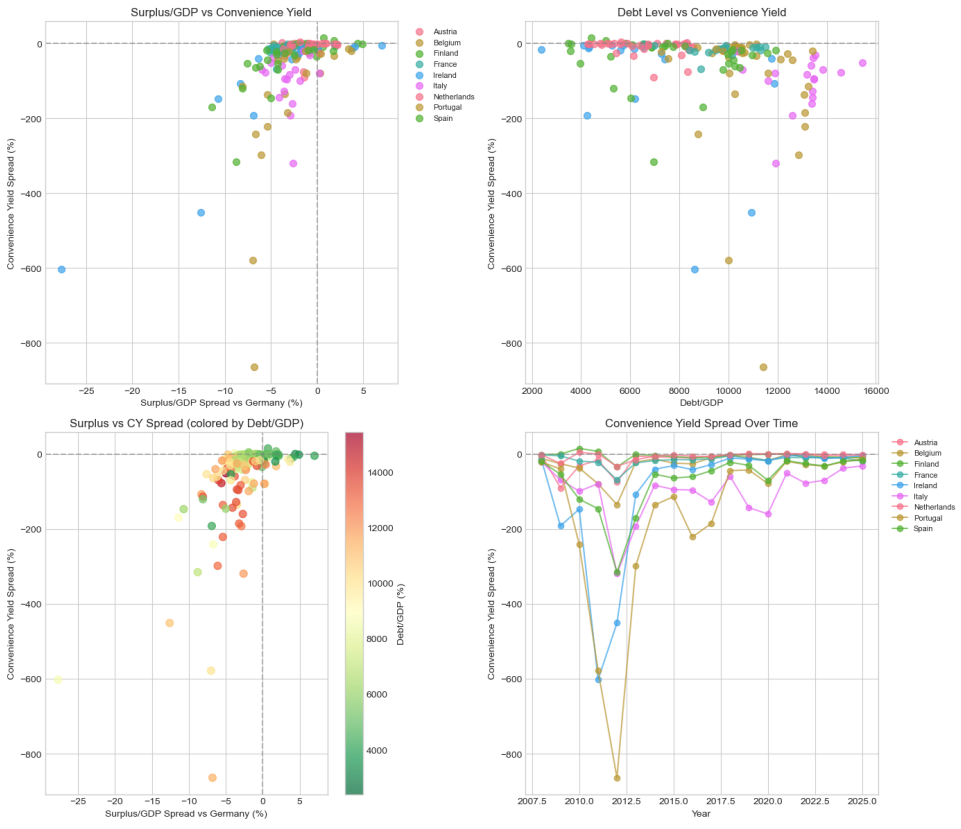

Figure 1: Descriptive Evidence

Figure 1 presents descriptive evidence from the raw data on the relationship between fiscal fundamentals and convenience yield spreads. The scatter plots indicate that this relationship is not well approximated by a simple linear pattern. Countries with stronger fiscal positions—characterized by higher surplus-to-GDP ratios and lower levels of public debt—tend to exhibit higher convenience yield spreads, whereas fiscally weaker countries are associated with lower convenience yield spreads.

The interaction between fiscal surplus and public debt is particularly evident in the panel that plots surplus-to-GDP spreads against convenience yield spreads while coloring observations by debt levels. For similar surplus positions, higher public debt is associated with systematically lower convenience yield spreads, suggesting that fiscal improvements are priced differently depending on the underlying stock of public debt.

The time-series panel further illustrates that sharp declines in convenience yield spreads coincide with episodes of fiscal and financial stress, most notably during the Eurozone sovereign debt crisis of 2010–2012. High-debt countries such as Portugal and Ireland experience deeper and more persistent declines during these periods. Overall, the descriptive evidence highlights pronounced heterogeneity and state dependence in the relationship between fiscal fundamentals and sovereign bond pricing, motivating the use of empirical approaches that allow for interactions and nonlinear effects.

3 Methodology

This section outlines the empirical framework used to study the relationship between fiscal fundamentals and sovereign bond convenience yield spreads. We begin with a linear fixed-effects regression as a benchmark, extend it using a threshold regression to allow for state-dependent effects, and finally apply tree-based machine learning models to capture nonlinearities in a data-driven manner.

3.1 Baseline Linear Regression

Our baseline specification is inspired by Jiang et al. (2024) but is not a direct replication. In particular, key variables are expressed as spreads relative to Germany rather than in levels, which is natural in a monetary union setting and allows us to abstract from common Eurozone-wide shocks. The baseline regression is given by:

$$CY_{i,t} = \alpha_i + \beta_1 Surplus_{i,t-1} + \beta_2 Debt_{i,t-1} + \gamma X_{i,t-1} + \varepsilon_{i,t}, \tag{1}$$

where $CY_{i,t}$ denotes the convenience yield spread relative to Germany, $Surplus_{i,t-1}$ is the lagged surplus-to-GDP spread relative to Germany, $Debt_{i,t-1}$ is the lagged debt-to-GDP ratio, and $X_{i,t-1}$ includes GDP growth and inflation spreads. All fiscal variables are lagged by one year to mitigate simultaneity concerns.

Country fixed effects α_i capture time-invariant heterogeneity across sovereigns. While this linear model provides a useful benchmark, coefficient interpretation may be complicated by multicollinearity among fiscal variables and by the assumption of constant marginal effects across fiscal states.

3.2 Threshold Regression

To allow for state-dependent effects, we estimate a threshold regression in which the marginal effect of fiscal variables differs across regimes defined by a threshold variable. The model is specified as:

$$CY_{i,t} = \begin{cases} \alpha_i + \beta^{(L)} Z_{i,t-1} + \gamma' X_{i,t-1} + \varepsilon_{i,t}, & q_{i,t-1} \leq \tau, \\ \alpha_i + \beta^{(H)} Z_{i,t-1} + \gamma' X_{i,t-1} + \varepsilon_{i,t}, & q_{i,t-1} > \tau, \end{cases} \tag{2}$$

where $CY_{i,t}$ denotes the convenience yield spread of country i relative to Germany at time t. The variable $Z_{i,t-1}$ represents the fiscal variable of interest (either the surplus-to-GDP spread relative to Germany or the debt-to-GDP ratio), while $q_{i,t-1}$ denotes the threshold variable that determines the regime. In our application, the threshold variable coincides with the fiscal variable under consideration. The vector $X_{i,t-1}$ includes macroeconomic control variables, namely GDP growth and inflation spreads relative to Germany, and α_i denotes country fixed effects.

The threshold value τ is selected to minimize the residual sum of squares over a grid of candidate values. Coefficients $\beta^{(L)}$ and $\beta^{(H)}$ are interpreted as regime-specific marginal effects, allowing the relationship between fiscal fundamentals and convenience yield spreads to differ across fiscal states.

3.3 Tree-Based Machine Learning Models

We complement the econometric analysis with two tree-based machine learning models. Tree-based models are particularly well suited to this setting because the underlying economic mechanisms seem to be state-dependent and potentially imply discontinuous pricing responses. When fiscal conditions affect expectations of future surpluses or perceived crisis probabilities, the marginal effect of fiscal variables on convenience yields may change sharply across fiscal states. Decision-tree methods naturally capture such threshold-type behavior by partitioning the data into regimes without imposing ex ante assumptions on the number or location of thresholds.

3.3.1 Random Forest

A Random Forest (Breiman, 2001) constructs a "forest" of decision trees. Each tree is then trained on a bootstrap sample of training data with a random subset of features at each split. The final prediction aggregates the trees' outputs by averaging across them. This method provides a number of advantages for threshold detection, which interests us for the sake of this project. First, it automatically captures non-linear relationships and interactions. Moreover, the decision trees split naturally, thereby identifying threshold values. Finally, Random Forest is robust to outliers and does not require feature scaling. However, Random Forest reduces variance by averaging many trees, which may lead to an understatement of thresholds. In order to avoid this issue on the object of our interest in this project, we consider another method as well.

3.3.2 XGBoost(Gradient Boosting)

XGBoost (developed by Chen and Guestrin (2016)) consists in building decision trees sequentially, where each tree corrects the errors of previous trees so as to minimize the loss. It is a gradient boosting method that adds some regularisations and penalties on model complexity (e.g., penalties on number of leaves per tree) and the magnitude of leaf weights. These restrictions can help limit overfitting and improve out-of-sample performance, especially in smaller samples, as compared to standard gradient boosting techniques.

3.4 Model Evaluation

Model performance is evaluated using an out-of-sample prediction framework. The dataset is split into training and test samples along the time dimension. The training sample covers the period 2008–2021 (126 observations), while the test sample covers 2022–2025 (36 observations), corresponding to an approximate 80/20 split (77.8%/22.2%). Splitting the data by year avoids mixing observations from crisis and non-crisis periods across samples and better reflects a forecasting-style evaluation in which models are assessed on their ability to predict future outcomes. We compare the out-of-sample predictive performance of four models: a linear fixed-effects regression serving as a benchmark, a threshold linear regression, a Random Forest model, and an XGBoost model. All models are trained on the same training sample and evaluated on the same test sample to ensure comparability.

Model performance is assessed using the mean squared error (MSE) and the root mean squared error (RMSE) computed on the test sample. In addition, we report percentage improvements in test MSE relative to the linear fixed-effects benchmark. RMSE is reported in the same units as the dependent variable, facilitating interpretation of prediction errors. We also verify

the predicted versus actual test data points visually in a scatter plot (included in the Jupyter notebook).

This evaluation focuses exclusively on predictive performance. The machine learning models are not designed for causal inference, and differences in predictive accuracy should not be interpreted as evidence of causal relationships between fiscal variables and convenience yields.

4 Empirical Results

4.1 Predictive Performance Comparison

Table 2: Out-of-Sample Model Performance Comparison

Model	Train MSE	Test MSE	Train RMSE	Test RMSE	Test MSE Improvement (%)
Random Forest	0.583	0.115	0.764	0.339	84.82
XGBoost	0.026	0.157	0.162	0.396	79.30
Linear Threshold FE	0.725	0.241	0.852	0.490	68.26
Linear FE–OLS	0.601	0.758	0.775	0.871	0.00

Table 2 reports out-of-sample prediction performance for all models considered, with the linear fixed-effects regression as a benchmark. The results show substantial differences in predictive accuracy across models. Among all specifications, the Random Forest model achieves the lowest test MSE and RMSE, reducing test MSE by approximately 85% relative to the linear benchmark. The XGBoost model also performs strongly, achieving a test MSE improvement of approximately 79%. These results indicate that tree-based machine learning models substantially outperform linear specifications in terms of out-of-sample predictive accuracy.

The threshold regression improves upon the linear fixed-effects model, reducing test MSE by approximately 68%. This suggests that allowing for regime-dependent effects captures some nonlinear structure in the data. However, the threshold model remains less accurate than the tree-based machine learning approaches, which flexibly capture nonlinearities and interactions without imposing a predefined threshold structure.

Finally, the linear fixed-effects regression exhibits the weakest out-of-sample performance, with a test RMSE of 0.871. While this specification provides a useful benchmark and point of comparison, its relatively high prediction error suggests that constant marginal effects across fiscal states may be too restrictive in this setting.

4.2 Threshold Regression Results

To further investigate nonlinear patterns suggested by the predictive results, we examine Partial Dependence Plots or PDP plots (Friedman, 2001) for the two main fiscal variables of interest: the surplus-to-GDP ratio and the debt-to-GDP ratio. Figure 2 compares partial dependence functions derived from the Random Forest and XGBoost models to the corresponding linear fixed-effects and threshold regression approximations.

The PDPs reveal clear deviations from linearity for both fiscal variables. For the debt-to-GDP ratio, the machine learning models exhibit a noticeable change in slope at debt levels around 85%. Below this threshold, variations in public debt are associated with relatively gradual changes in predicted convenience yield spreads, with the partial dependence functions displaying a relatively flat profile. Beyond this point, predicted convenience yields decline more sharply as debt increases, indicating a stronger association between higher debt levels and lower convenience yield spreads. The estimated threshold from the linear threshold regression, shown by the vertical dashed line, lies close to this region and captures part of the observed nonlinearity.

For the surplus-to-GDP ratio, the PDPs indicate a nonlinear relationship characterized by a steep deterioration in predicted convenience yield spreads at very large deficits, followed by a flatter

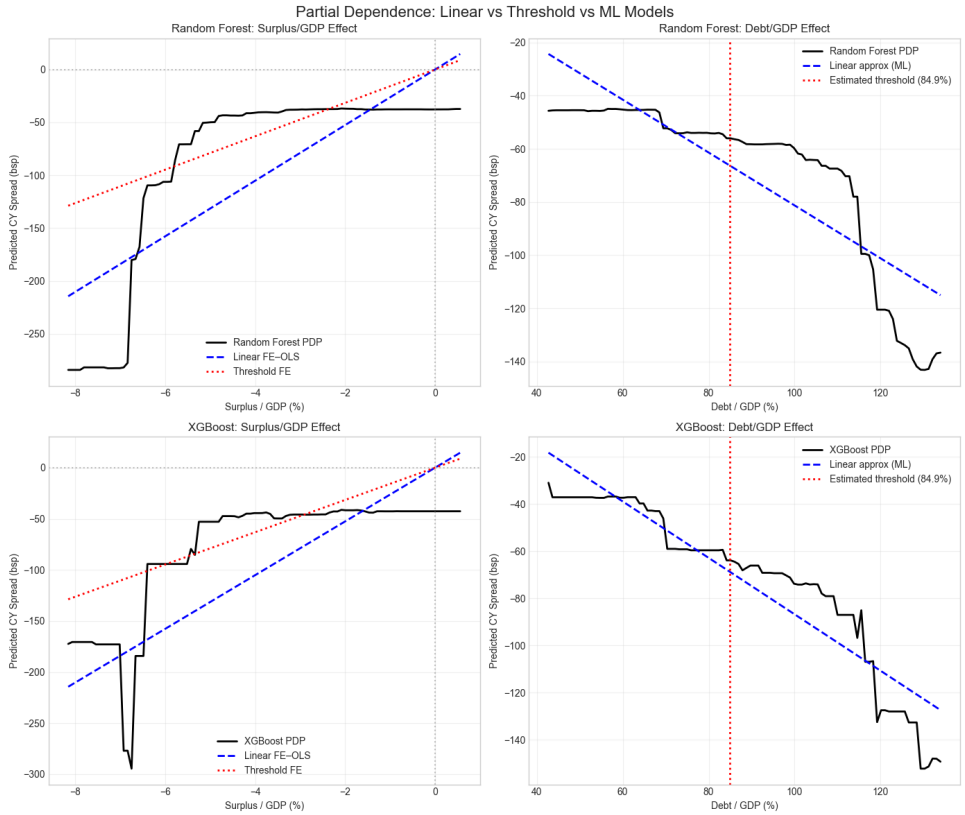

Figure 2: PDP Plots

response for moderate deficit levels. In particular, both Random Forest and XGBoost show pronounced changes when surplus-to-GDP ratios fall below approximately −6 percent, while the marginal effect becomes more muted as deficits improve. The threshold regression approximates this pattern by allowing for a discrete change in slope, whereas the linear specification imposes a constant marginal effect across the entire range.

Overall, the partial dependence analysis highlights state-dependent and nonlinear associations between fiscal fundamentals and convenience yield spreads. While the precise shape of these relationships differs across models, the results consistently suggest that linear specifications may mask important changes in the relationship between fiscal variables and sovereign bond pricing across fiscal regimes. These findings are descriptive in nature and complement the out-of-sample performance results by illustrating how flexible models capture nonlinear patterns in the data.

5 Interpretation

The empirical results suggest that the relationship between fiscal fundamentals and sovereign bond convenience yield spreads is nonlinear and state-dependent. While linear specifications impose constant marginal effects across fiscal states, both the threshold regression and the machine learning models indicate that the association between fiscal variables and convenience yields changes once fiscal conditions deteriorate beyond certain levels.

The partial dependence analysis reveals pronounced nonlinearities in the relationship between fiscal fundamentals and convenience yield spreads. Predicted convenience yields decline relatively gradually at low and moderate debt levels, but deteriorate much more sharply once debt-to-GDP ratios exceed approximately 85 percent. This pattern is only partially captured by the threshold regression and is more flexibly represented by the tree-based machine learning models. A similar state-dependent relationship emerges for the surplus-to-GDP ratio, where very large deficits are associated with disproportionately lower predicted convenience yield spreads, while marginal

improvements at more moderate deficit levels have smaller predictive effects.

These nonlinear patterns help explain the superior out-of-sample performance of flexible models. Linear fixed-effects regressions average over fiscal states and therefore mask changes in the strength of the association between fiscal fundamentals and sovereign bond pricing across regimes. Allowing for regime-dependent effects improves predictive accuracy, and fully flexible machine learning models perform best by capturing nonlinearities and interactions without imposing a predetermined functional form.

Importantly, the estimated threshold locations do not align exactly with formal European Union fiscal rules, such as the Maastricht debt criterion or the Stability and Growth Pact deficit limit. Instead, nonlinearities emerge at higher debt levels and larger deficits, suggesting that the pricing of fiscal risk may respond more strongly to sustained or severe fiscal deterioration rather than to mechanical rule-based cutoffs. These results therefore point to nonlinear pricing patterns rather than sharp policy-triggered effects.

Finally, the analysis is subject to important limitations. The sample size restricts the complexity of models that can be reliably estimated, and the machine learning methods employed are designed for prediction rather than causal inference. Consequently, the results should be interpreted as descriptive evidence of nonlinear associations between fiscal variables and convenience yield spreads, rather than as estimates of causal effects. Despite these limitations, the findings highlight the importance of accounting for state dependence and nonlinearities when analyzing the fiscal determinants of sovereign bond pricing.

6 Conclusion

This project shows that flexible, tree-based machine learning models provide substantial gains in predicting sovereign bond convenience yield spreads relative to standard linear specifications. In particular, Random Forest and XGBoost significantly outperform the linear fixed-effects benchmark in out-of-sample prediction, indicating that the relationship between fiscal fundamentals and convenience yields might be nonlinear.

The results further suggest that fiscal variables are associated with convenience yields in a state-dependent manner. Partial dependence analysis highlights nonlinearities, with sharper deteriorations in predicted convenience yields at high levels of public debt and under severe fiscal deficits – patterns that are not well captured by linear models with constant marginal effects. Threshold regressions capture some of this structure but remain less flexible than machine learning approaches.

While the analysis is predictive rather than causal, the findings complement existing linear evidence by demonstrating that fiscal determinants of convenience yields operate differently across fiscal states. Future work could explore time-varying nonlinearities, alternative fiscal indicators, or extend the analysis to a broader set of sovereign issuers beyond the Eurozone.

References

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785-794.

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

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232.

Jiang, Z., Lustig, H., Van Nieuwerburgh, S., and Xiaolan, M. Z. (2025). Bond Convenience Yields in the Eurozone Currency Union. *NBER Working Paper*.

Krishnamurthy, A., Nagel, S., and Orlov, D. (2014). Sizing up repo. *The Journal of Finance*, 69(6):2381–2417.

Nagel, S. (2016). The liquidity premium of near-money assets. *The Quarterly Journal of Economics*, 131(4):1927–1971.

Bertola, G. and Drazen, A. (1993). Trigger points and budget cuts: Explaining the effects of fiscal austerity. *American Economic Review*, 83(1):11–26.

Cole, H. L. and Kehoe, T. J. (2000). Self-fulfilling debt crises. *Review of Economic Studies*, 67(1):91–116.

Gennaioli, N., Shleifer, A., and Vishny, R. (2018). *A Crisis of Beliefs*. Princeton University Press.

Codogno, L., Favero, C., and Missale, A. (2003). Yield spreads on EMU government bonds. *Economic Policy*, 18(37):503–532.

Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? *Journal of Banking & Finance*, 45:182–198.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Hansen, B. E. (2011). Threshold regression models. *Encyclopedia of Statistics in Quality and Reliability*.

Ardagna, S., Caselli, F., and Lane, T. (2007). Fiscal discipline and the cost of public debt service: Some estimates for OECD countries. *The B.E. Journal of Macroeconomics*, 7(1).

Égert, B. (2015). Public debt, economic growth and nonlinear effects: Myth or reality? *Journal of Macroeconomics*, 43:226–238.

Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics*, 93(2):345–368.

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *American Economic Review: Papers & Proceedings*, 109:1–9.

Gennaioli, N., Martin, A., and Rossi, S. (2018). The safe asset shortage conundrum. *The Journal of Finance*, 73(1): 1–40.

European Council (1992). Treaty on European Union (Maastricht Treaty). Official Journal of the European Communities.

Buiter, W., Corsetti, G., and Roubini, N. (1993). Excessive deficits: Sense and nonsense in the Treaty of Maastricht. *Economic Policy*, 8(16):57–100.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer, New York.

Appendix

All code used in this analysis is provided in the accompanying Jupyter Notebook. The notebook contains full data construction pipeline, model estimation procedures and all figures and tables reported in the paper, ensuring full reproducibility of the results.