FEM21019
Seminar Financial Case Studies
Machine Learning in the Cross-section of Government Bonds

Group ROBECO-C:
Luis Bettencourt 613121
Sunny Hsieh 604455
Leonardo Knol 513082
Shubham Rathi 573379

Erasmus University Rotterdam

Rotterdam, January 13, 2023

Title:

Improving the Prediction of Cross-Sectional Government Bond Returns through Machine Learning

Research question:

Can machine learning techniques be used to improve the prediction of cross-sectional government bond returns?

Background and Literature:

The potential of Machine Learning (ML) in Finance has been under exploration for past few years and it's only gaining more traction as time passes. The premise of ML in Finance is well explained and argued by Israel et. al. (2020) in their paper "Can machines learn Finance?". They put forth that machine learning have "promising" use cases in factor investing, risk management, trading cost management, return prediction etc. Return prediction is one of the hottest topics in asset management and, more generally, in the financial industry. Gu et. al. (2020) show that using Machine Learning methods for equity return prediction gives superior results as compared to linear models which are highly employed in academia and in the industry. They attribute this performance gain of ML over linear models to ML's additional ability to model and capture non-linearity and interaction effects which the linear models, by their structure, cannot.

In the fixed income literature, the use case of ML is not as deeply explored as in equity markets. The potential of machine learning in fixed income has been explored in the corporate bonds domain by Bali et. al. in their paper "Predicting Corporate Bond Returns: Merton Meets Machine Learning" (2022) where they use machine learning to cross-sectionally forecast (corporate) bond returns using ML - on a standalone (statistical) basis and also while combining it with economic theory. To our knowledge, the research on machine learning for government bonds is under-explored as of now with no paper focusing on utilizing machine learning to forecast cross-sectional returns for government bonds. As discussed in preceding section, we think it is an exciting and important topic to research for two reasons. First, government bonds make up 30% of overall market capitalization across asset classes Doeswijk et al., (2020). Secondly, Martens et. al. in their paper "Predicting Bond Returns: 70 years of International Evidence" (2021) show that there is strong and consistent predictability in bond returns that is statistically significant. In Martens et. al. following paper "Factor Investing in Sovereign Bond Markets: Deep Sample Evidence" (2021) they show that bond factors, in particular - momentum, value and low-risk offer consitent and persitent premiums. Our intention in this research is to see if Machine Leanring methods can outperform the linear (factor) model discussed in "Factor Investing in

Sovereign Bond Markets: Deep Sample Evidence" (2021) due to their added flexibility and ability to model interactions and incorporate non-linearity. Our work can be seen as an extension of Gu, Kelly, and Xiu (2020)' work in the government bond space where we explore whether ML based models and strategies can be more profitable than linear (factor) models.

The potential of Machine Learning (ML) in finance has been the subject of research in recent years. Studies have shown that ML has promising use cases in areas such as factor investing, risk management, and return prediction (Israel et al., 2020).
In the area of return prediction, researchers have found that ML methods can achieve superior results when compared to traditional linear models, particularly in equity markets. For example, Gu et al. (2019) in their study "Empirical Asset Pricing via Machine Learning" demonstrated that ML-based models can effectively model and capture non-linearity and interaction effects, which linear models cannot.
In the fixed income segment, the use of ML has not been as extensively explored. Research in the area of corporate bonds has been conducted by Bali et al. (2022) in their study "Predicting Corporate Bond Returns: Merton Meets Machine Learning". However, to the best of our knowledge, research on machine learning for government bonds is currently under-explored.
This is particularly important as government bonds make up 30% of overall market capitalizations across asset classes (Doeswijk, Lam, & Swinkels, 2020), and research has shown strong and consistent predictability in bond returns (Martens et al., 2021). Martens et al. (2021) also demonstrated that bond factors, such as momentum, value, and low-risk, can offer consistent and persistent premiums in sovereign bond markets.
Our research aims to build upon this literature by exploring the potential of ML-based models and strategies in predicting cross-sectional returns for government bonds. Our study can be seen as an extension of the work done by Gu, Kelly, and Xiu (2020) in the government bond space, where we aim to determine if ML-based models can outperform traditional linear (factor) models, as discussed in Martens et al.'s "Factor Investing in Sovereign Bond Markets: Deep Sample Evidence" (2021), due to their added flexibility and ability to incorporate non-linearity and interactions.

Objectives:

The main objective of this research is to explore the potential of machine learning techniques to improve the prediction of cross-sectional government bond returns. To achieve this objective, the following specific research questions will be addressed:

- What are the most important predictor variables for government bond returns, as identified by different Machine Learning algorithms?
- How do the predictive accuracy and economic gains of Machine Learning methods compare to those of traditional techniques in the context of government bond returns?

- How does the choice of Machine Learning method affect the prediction of government bond returns?
- Can portfolio strategies based on Machine Learning techniques outperform traditional benchmarks, under various measures of portfolio performance?

Data:

The data for this research proposal was provided by Robeco. The data provided consists of 85 predictor variables and 42 bonds for the target variable for each month, from 1980 until 2022. The target variable includes bonds of 6 different maturities for 7 different countries. Furthermore, a significant number of missing values are to be seen in the target variable. Further examination revealed that the majority of missing values are due to bonds with a maturity of 15 years or more. However, other maturities also portray missing values prior to May 1990. To address this issue, all bonds with a maturity of 15+ years from the 7 countries are going to be removed from the dataset, and all data prior to May 1990 will be excluded.

Additionally, there are also missing values in the features related to the equity (bond) characteristics. To handle these missing values, the approach proposed by Gu et al. (2020) will be implemented. This method replaces the missing values with the cross-sectional median characteristic of each stock (bond) for each month.

This approach preserves the overall distribution of the data and accounts for any outliers or extreme values. It has been widely used in similar studies and found to provide reliable results. The modified dataset with replaced missing values for the equity characteristics will be used for the analysis. The data quality and reliability will be confirmed by robustness checks and sensitivity analysis in the proposed research. So from now on we will use only 35 bonds in our research.

Methods:

The research will be conducted by means of empirical analysis. A large dataset of 35 government bond prices and other relevant variables will be collected and analyzed using a range of machine learning algorithms, including linear and generalized linear models, dimension reduction techniques, tree-based methods, and neural networks. The predictive accuracy and economic gains of these methods will be compared to those of traditional techniques, such as cross-sectional regressions and time-series regressions.
Throughout this research, both Linear and Machine Learning models will be used in this analysis. Although the main focus is on the prediction capabilities and dynamics of the Machine Learning methods, the Linear models will serve as a benchmark.

- Multivariate Regression (all features). A linear regression that uses all available features to predict the vector of the Bond's returns. This method is expected to perform poorly but will serve to outline the shortcomings of Linear models on non-linear relations.

- Multivariate Regression (Bond Factors). The same regression method described earlier is applied, but considering the factor model from Baltussen, Martens, Penninga (2021), with the predictors: Momentum, Value, and Low-Risk. Unlike the previous model, this one implies a proven multi-factor strategy, thus serving as a proper benchmark for the performance of the Machine Learning models.

- Regularization (Elastic Net). Introducing the regularization terms to a linear model allows for variable selection among the features, thus improving the out-of-sample stability. It will provide the first insights into the importance of each feature.

- Principal Components Regression (PCR). This method also tackles dimension reduction by keeping a set of linear combinations of the data from performing principal component analysis, and zeros the coefficients of the components with low variance.

- Generalized Linear Model (GLM). It is the first method that takes into account non-linear interactions in this exercise. Similarly Gu, Kelly, and Xiu (2020), we chose the group lasso penalization function.

- Random Forests (RF). The first of the Machine Learning methods with a high potential for economic gain. Unlike GLM there is no established assumption on the non-linear interactions, making it more flexible.

- Support Vector Machine (SVM). Another Machine Learning method that is expected to deliver economic gains, SVM is chosen as a parallel to the Neural Networks of Gu, Kelly, and Xiu (2020), but taking into the shorter amount of data used.

In order to evaluate each model's performance, the out-of-sample R^2 for the individual excess stock returns. During the course of the research, other methods proposed by literature may be considered, namely in regard to ranking the algorithms, addressing feature importance, or portfolio strategy performance.

Expected results:

It is expected that the results of this research will provide new insights into the use of Machine Learning for the prediction of cross-sectional government bond returns. In particular, it is expected that the results will identify the most important predictor variables for government bond returns, and will demonstrate the relative performance of different Machine Learning algorithms in terms of predictive accuracy and economic gains. It is also expected that the results will shed light on the sensitivity of government bond return prediction to the choice of Machine Learning method, and on the potential for combining Machine Learning with economic theory to better understand the mechanisms underlying bond returns. Furthermore, this research intends to form portfolios as done in Kelly et al., 2020, with the purpose of identifying possible enhancements in the performance of the Machine Learning methods used.

Potential impact:

This research has the potential to contribute to the understanding of government bond returns and to the development of more effective prediction methods in this context. By demonstrating the effectiveness of Machine Learning in this context, the research may also encourage the wider adoption of these methods in empirical asset pricing more generally.

References:

Bali, T. G., Goyal, A., Huang, Y., Jiang, W., & Wen, L. (2022). Predicting Corporate Bond Returns: Merton Meets Machine Learning. Journal of Financial Economics.

Doeswijk, R., Lam, K., & Swinkels, L. (2020). A review of the global bond market. BIS Quarterly Review, 53-68.

Gu, J., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. Journal of Financial Economics, 137(1), 1-38.

Israel, R., Kelly, B., & Moskowitz, T. J. (2020). Can machines learn finance? Journal of Financial Economics, 136(1), 1-22.

Martens, M., Patel, J., & Vasios, M. (2021). Factor Investing in Sovereign Bond Markets: Deep Sample Evidence. Journal of Portfolio Management.

Martens, M., Petajisto, A., & van der Zwan, J. (2021). Predicting Bond Returns: 70 years of International Evidence. Journal of Financial Economics, 138(1), 1-19.

Appendix

Additional possible questions to incorporate

1. Can machine learning techniques be used to identify patterns in government bond return predictors that are not detectable by traditional methods?

2. How do the predictive abilities of machine learning methods for government bond returns change over time, and how can these changes be explained by economic and market conditions?

3. Can machine learning methods be used to develop more accurate risk measures for government bond investments, such as value-at-risk or expected shortfall?

4. How do the results of machine learning-based government bond return prediction compare to those obtained from other asset classes, such as equities or commodities?

5. Can machine learning techniques be used to improve the prediction of government bond yield spreads, and how do these predictions compare to those obtained using traditional techniques?

6. How does the inclusion of macroeconomic variables in the predictor set affect the prediction of government bond returns using machine learning methods?

7. Can machine learning techniques be used to predict the direction of government bond returns, or are they only effective at predicting the level of returns?
8. How do the results of machine learning-based government bond return prediction vary across different bond issuers and markets, and how can these differences be explained?
9. Can machine learning methods be used to predict the impact of government bond supply and demand on bond returns, and how does this impact vary across different bond issuers and markets?
10. How do the results of machine learning-based government bond return prediction change when considering different time horizons, such as short-term, medium-term, or long-term returns?

Tables 1 and 2 depict the format whereby the results and descriptive statistics are intended to be displayed respectively. Furthermore, the portfolios that are formed will be displayed as done in Table 3.

Table 1: Monthly out-of-sample cross-sectional bond prediction performance. (Note that the performance is given in terms of R^2_{oos})

	OLS$_1$	**OLS$_2$**	**Reg**	**PCA**	**GLM**	**RF**	**NN1**	**NN3**	**NN5**	**SVM**
Results										

Table 2: Cross-sectional statistics over the sample period of 1990-2022.

				Percentile			
	Mean	**Median**	**SD**	**5$_{th}$**	**25$_{th}$**	**75$_{th}$**	**95$_{th}$**
Return (%)							
Carry							
Momentum							
Value							
GDP							

Table 3: Portfolio Performance Across the Various Models

	R^2	Sharpe Ratio		
Model	**Cross-section**	**Equal-weight**	**Value-weight**	**Description**
$\mathbf{OLS_1}$				
$\mathbf{OLS_2}$				
GLM				
RF				
NN				