Predicting Financial Fragmentation using Machine Learning

Roland Bouillot ROLAND.BOUILLOT@MASTRICHTUNIVERSITY.NL
Department of Macro, International & Labour Economics
Maastricht University
Maastricht, Tongersestraat 53, The Netherlands

Bertrand Candelon BERTRAND.CANDELON@UCLOUVAIN.BE
LFIN
Université catholique de Louvain
Louvain-la-Neuve, Voie du Roman Pays 34, Belgium

Clemens Kool C.KOOL@MAASTRICHTUNIVERSITY.NL
Department of Macro, International & Labour Economics
Maastricht University
Maastricht, Tongersestraat 53, The Netherlands

Abstract

Since the recent surge in inflation and positive real interest rate, financial fragmentation risk is again a matter a fear in the euro area. This paper proposes an evaluation of the fragmentation risk building a novel high-dimensional dataset covering a wide range of fields within 10 European countries over the period 2007-2024. This Big Data dataset has been exploited by a new machine learning technique (XGBoost) to find evidence of the financial fragmentation risk in the Euro Area. It turns out that the predicted long-term yield spreads of peripheral countries rise while those in core countries' rise remain contained or even decrease. This divergence in yield differentials put a lot of scrutiny on the action of the ECB and call for policy guidance to manage a new European sovereign debt crisis.

Keywords: Machine learning, Financial fragmentation risk, XGBoost, Sovereign spreads

Non-technical summary

- The financial fragmentation risk in the Euro area has been rekindled since the COVID-19 pandemic, the European energy crisis and the war in Ukraine. The induced economic turmoil disrupts ECB's mandate to tame inflation as the economic soundness of highly indebted countries is affected by the rate hikes;
- This paper leverages Big Data and machine learning techniques to predict financial fragmentation risk within the Euro Zone;
- The Big Data dataset encompasses several macroeconomic and financial categories of variables from ten Euro-Zone countries from March 2007 to January 2024 (monthly frequency) and gathers more than 900,000 observations overall;
- XGBoost, a novel Gradient Boosted Machine technique, is used as our main Machine Learning technique to extract the most significant features of our dataset in order to predict 10-year sovereign spreads 6-months ahead;
- The main result of our analysis is in line with the recent literature regarding the resurgence of Financial Fragmentation risk: peripheral countries are predicted to have higher long-term yield spreads while core countries yield spreads' rise remain contained or even decreases;
- Our analysis categorizes Austria, the Netherlands and Finland as core countries while peripheral countries are Portugal, Spain, Italy, Ireland, Greece as well as -for the first time ever- France and Belgium;
- The divergence in yield spreads can be explained by "contagion" and/or "flight-to-quality" behaviors from investors, revealing the dynamics between core and peripheral countries;
- In conclusion, financial fragmentation risk is well and alive within the Euro zone, calling for action from macroprudential experts and providing a working ground for academic researchers.

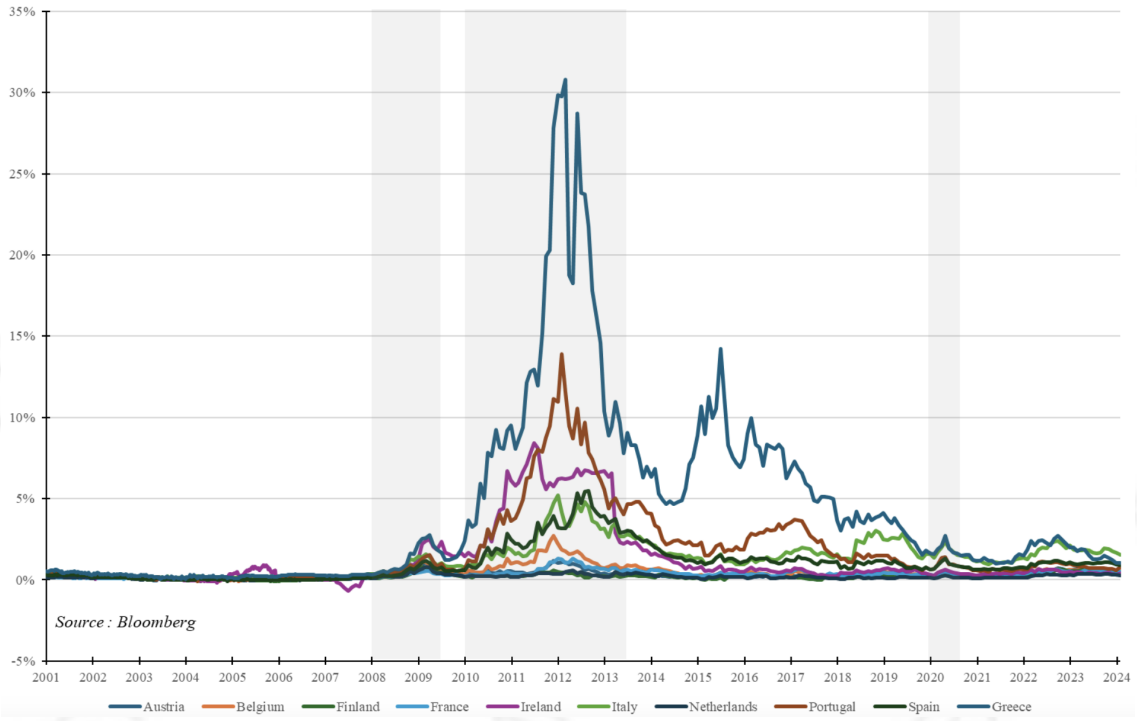
Contents

1 Introduction

The Great Financial Crisis in 2008 and the European Sovereign Debt crisis in 2011 unveiled the Euro area's vulnerabilities. Since then, the Euro area's financial landscape, particularly its sovereign bond market has been a focal point. The introduction of the Euro as a common currency seemed to pave the way towards a convergence, in sovereign bonds across member states. However, as debt and deficit concerns intensified and diverged between Euro countries, investors began discerningly differentiating between sovereign bonds issued by Euro area countries, thereby sowing the seeds of fragmentation among core and peripheral countries.

Financial fragmentation risk appears as the divergence in the sovereign bonds yield spreads across countries. At the apogee of the Euro Area crisis, yield spreads for countries such as Portugal, Spain and Greece skyrocketed. In contrast, those for countries like Austria, the Netherlands or Finland remained relatively stable as shown in Figure 1.

Figure 1: Euro area sovereign spreads



Notes: 10-years sovereign bond spreads for Euro area countries under study (Austria, Belgium, Finland, France, Greece, Ireland, Italy, the Netherlands, Portugal and Spain). The benchmark country is Germany. Weekly data retrieved from Bloomberg. Grey sections indicate the timing of the Great Financial crisis, the European Sovereign Debt Crisis and the COVID-19 crisis.

On the one hand, this divergence reflects "contagion" in sovereign bond markets, where financial stress in one country raises borrowing costs in others (Favero & Missale, 2012).

This relationship arises from investor perceptions of shared risks, such as potential Euro area financial fragmentation, which amplifies spreads across countries. On the other hand, in search of safer havens, investors might flock to bonds perceived as less risky, driving their yields down, implying a "flight-to-quality" behavior from investors (Beber *et al.*, 2009). High-indebted countries running large deficits, and grappling with solvency concerns, witnessed their borrowing costs increase, further straining their fiscal positions. In contrast, countries perceived as economically sound enjoyed reduced borrowing costs, further solidifying their stable economic position. This divergence in borrowing costs reflects heterogeneous risk premia between countries using a unique currency. It therefore generates a risk of the split into "two-speed" Euro area opposing core and peripheral countries, reported as the fragmentation risk. In response to the crises and the fragmentation risk, the European Central Bank (ECB) and other European institutions embarked on a series of measures. The Outright Monetary Transactions (OMT) program, unveiled by the ECB in 2012, aimed at mitigating sovereign bond market fragmentation. By committing to potentially unlimited purchases of sovereign bonds under specific conditions, the ECB sought to reassure market concerns and foster integration.

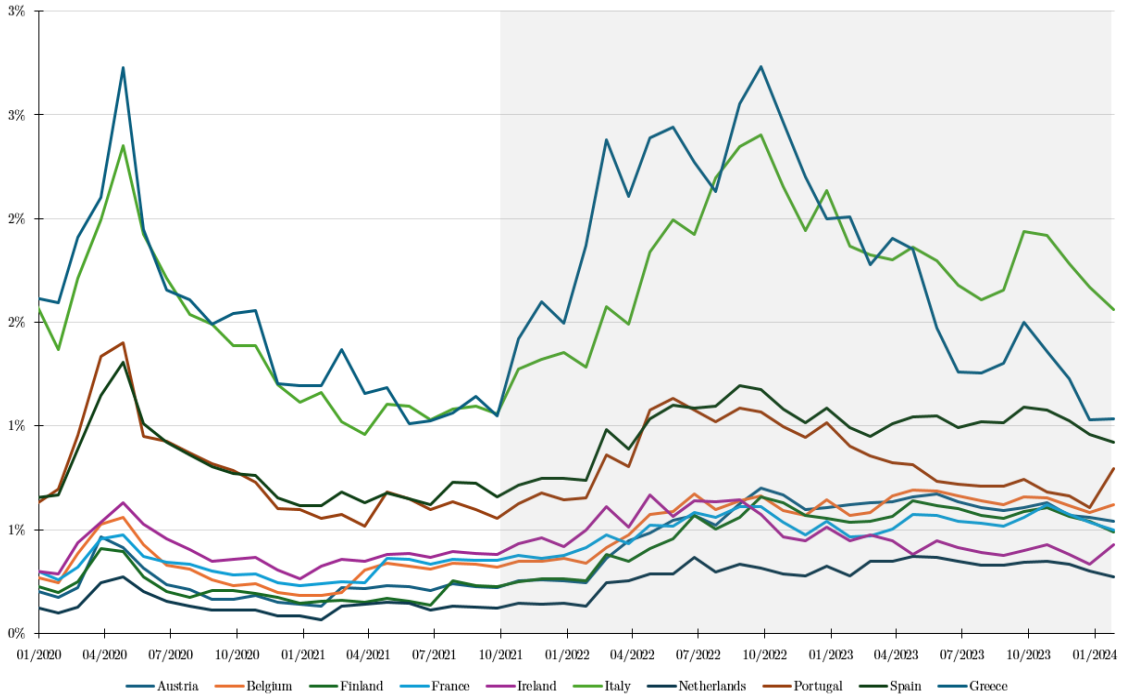
While policy measures have reduced fragmentation risk over the past decade, the path to complete integration remains fraught with challenges. Structural impediments, such as incomplete banking integration, disparities in national regulations, and the absence of a cohesive fiscal policy, continue to undermine the efforts towards a convincing Euro area integration. [1]

This paper proposes to propose novel forecast of the sovereign yield spreads to provide evidence of financial fragmentation risk in the Euro area. The first objective is to identify the divergence of spreads among peripheral and core countries in the Euro area. Highly indebted countries should exhibit rising yield spreads driven by "contagion" whereas sound economies would benefit from "flight-to-quality" and experience decreasing yield spreads. The second objective is to identify which countries are core or peripheral in the Euro area. As the economic and financial soundness of countries may change over time, their belonging to one of the binary clusters is not guaranteed. Indeed, countries with rising debt and large public deficit but being part of the core cluster may shift towards the peripheral cluster. The converse scenario is also possible: a country whose debt stabilizes and whose public deficit is sustainable can shift from the peripheral to the core cluster.

To this aim, we exploit Big Data, specifically a novel high-dimensional dataset. This dataset gathers as much data as available for each of the ten countries under study, leading to a total of 900,000 monthly observations over the March 2007 to January 2024 period. This paper considers every aspect of a country's economic landscape, ranging from labor market indicators to sectorial data, from price indexes to housing & real estate markets, from government finances to demographics. This paper is one of the first we have knowledge of to leverage Big Data to predict financial fragmentation risk, yet such amount of data would not properly fit regular econometric models due to the "Curse of dimensionality" (Bellman, 1966).

1. See Candelon, Luisi & Roccazzella, 2022

Figure 2: Euro area sovereign spreads



Notes: 10-year sovereign bond spreads for Euro area countries under study (Austria, Belgium, Finland, France, Greece, Ireland, Italy, the Netherlands, Portugal, and Spain). The benchmark country is Germany. Weekly data retrieved from Bloomberg.

As dimensions grow (i.e. the number of variables in the model), traditional econometric models become less efficient due to sparse data distribution and difficulties in identifying relevant variable relationships. This sparsity reduces the reliability of parameter estimates and exacerbates over-fitting. In contrast, machine learning models incorporate techniques like regularization or variable selection, allowing them to manage high-dimensional spaces more effectively, to capture complex patterns, and to mitigate the limitations imposed by dimensionality.

In this study, we explore the application of Machine Learning techniques in two critical areas: (i) we conduct a horse race between 14 different Machine Learning techniques to find which algorithm fits best our dataset and hence provide insights into the power of Machine Learning in variable selection. Three main categories of Machine Learning frameworks are selected, namely linear, non-linear and rule based methods. Each algorithm is run for each country, collecting comparable accuracy metrics such as the Root Square Mean Error (RMSE), Mean Absolute Error (MAE) and R^2. We find that the gradient-boosting framework within the rule-based category suits our dataset best. Hence, (ii) using the latest advances in machine learning, and more specifically the XGBoost algorithm, this paper aims at providing insights on the predictability power of Machine Learning applied to the 10-year yield spread of the ten Euro Area countries under study and finding evidence of

the latent financial fragmentation risk. Indeed, our empirical results suggest that the fragmentation dynamics between core and peripheral country is still at work. More worryingly, France and Belgium -two formerly core countries- have swung into the peripheral category, bringing back the fragmentation risk in the frontstage.

This paper is structured as follows : the related literature is presented in Section 2 ; the novel dataset is introduced in Section 3 ; Section 4 provides an in-depth explanation of the different machine learning techniques and the horse race methodology. Finally, empirical results are analyzed in Section 5.

2 Literature Review

2.1 The determinants of sovereign bond yields and spreads

In order to understand the financial fragmentation risk, it is important to understand what the determinants of sovereign bond yields and spreads are. They encompass macroeconomic fundamentals, market dynamics, institutional frameworks and political factors. The interaction of these elements became particularly evident during the European Sovereign Debt (ESD) crisis, where yield spreads became particularly significant, capturing the sentiment of market participants regarding individual sovereign risk and the stability of specific economies within the Euro area.

Macroeconomic fundamentals remain the main determinant of sovereign bond yields. A country's debt-to-GDP ratio, fiscal deficit and economic growth are among the primary indicators analyzed by investors. Bernoth et al. (2004) highlight that deviations from fiscal sustainability, especially in the context of the Maastricht criteria, are strongly correlated with higher yield spreads. High debt ratios not only reflect a greater likelihood of repayment challenges but also signal potential inflationary pressures, particularly in the absence of fiscal discipline. Afonso et al. (2011) extend this analysis by quantifying the impact of fiscal announcements on yields, finding that markets respond more sharply to unexpected deviations from fiscal targets. Similarly, Aizenman et al. (2013) examine the role of macroeconomic imbalances, such as current account deficits and inflation, in driving spreads. Their findings suggest that, beyond fiscal metrics, structural vulnerabilities in the economy exacerbate perceived risks, particularly during global downturns.

Market behavior is another critical determinant. The role of credit ratings in influencing market dynamics is explored by Afonso et al. (2012). Their research shows that downgrades by major credit rating agencies significantly widen yield spreads by signaling increased sovereign risk to investors. Conversely, positive credit rating actions have a muted impact, reflecting asymmetric market responses to good and bad news. Market volatility, often represented by measures such as the VIX or the Euro STOXX 50 index, also plays a pivotal role. Codogno et al. (2003) document how heightened volatility increases risk premiums, disproportionately affecting countries perceived as less stable. This dynamic underscores the importance of maintaining macroeconomic and fiscal stability to buffer against external shocks. Behavioral factors such as investor sentiment and herd behavior also contribute to yield dynamics. Arghyrou and Kontonikas (2012) explore how self-fulfilling prophecies during the ESD crisis led to liquidity crises in countries like Italy and Spain, even in the absence of fundamental solvency issues.

Institutional factors, particularly the role of the European Central Bank (ECB), have significantly shaped sovereign bond yields in the Euro area. The ECB's ability to transmit monetary policy uniformly across the Euro Area is crucial in determining yields and spreads. Grandi (2019) investigates the bank lending channel of monetary policy in mitigating yield divergence, finding that while measures like OMTs and QE have been successful in the short term, persistent structural differences among member states continue to pose chal-

lenges. In addition, the introduction of unconventional monetary policies during the ESD crisis marked a turning point. Altavilla et al. (2015) show that the ECB's Quantitative Easing (QE) measures lowered yields and spreads by reducing fragmentation in financial markets and fostering liquidity. Additionally, Bernoth et al. (2004) highlight the success of the ECB's Securities Markets Programme (SMP) in alleviating pressures on bond markets in peripheral countries. However, the lack of fiscal integration within the Euro area remains a structural issue.

Another important determinant is political stability. Manganelli and Wolswijk (2007) emphasize that political uncertainty—whether from domestic governance issues or broader geopolitical tensions—can lead to significant yield divergence. For instance, the heightened uncertainty surrounding the war in Ukraine and its implications for the Euro area underscore how political events influence bond market dynamics.

2.2 The impact of the Sovereign Debt Crisis (ESD)

The second step to understand the financial fragmentation risk in the Euro area is finding evidence of it. This risk arises from the divergence in yield spreads that can be defined as the persistent gap between the borrowing costs of core countries, such as Germany, Austria and the Netherlands, and peripheral countries, including Greece, Italy and Spain. Hence, this fragmentation risk undermines the principle of financial integration, which is foundational to the European Monetary Union (EMU). The European Sovereign Debt crisis of 2010–2012 marked a pivotal period for the European Monetary Union revealing deep structural vulnerabilities. The perception by the market participants of different levels of risk across countries within the Euro area significantly transformed the dynamics of sovereign bond markets and revealed the financial fragmentation risk to the world.

Before the ESD crisis, sovereign bond yields in the Euro Area were remarkably low and converged closely, reflecting market optimism about the benefits of monetary unification. However, this apparent stability masked underlying risks. Favero and Missale (2012) argue that sovereign bond markets were never fully integrated, even during the pre-crisis period. Their findings demonstrate that despite low and co-moving spreads, market participants did not perceive bonds from different Euro area countries as perfect substitutes. De Grauwe and Ji (2013) argue that market complacency during the pre-crisis period led to insufficient risk differentiation among Euro area countries, particularly between core and peripheral economies. The global financial crisis of 2008–2009 acted as a precursor, exposing vulnerabilities in fiscal sustainability and triggering an initial divergence in spreads as investors reassessed sovereign risks. Candelon et al. (2022) emphasize that even before the crisis formally began, early signs of fragmentation were evident in the differentiated impacts of the 2008 financial crisis. While core economies like Germany and France benefitted from capital inflows, peripheral countries such as Greece, Italy, and Spain faced mounting pressures, with rising debt-to-GDP ratios and deteriorating fiscal balances.

The ESD crisis began in 2010 when Greece revealed significant revisions to its budget deficit, prompting a loss of market confidence. This revelation led to a sharp rise in Greek bond yields and spreads, signaling the onset of a solvency crisis. Unlike Greece, which faced fundamental fiscal insolvency, other high-debt countries such as Spain, Portugal, and Italy encountered liquidity crises driven largely by market panic and contagion effects. The interconnectedness of Euro Area economies means that shocks in one country often have spillover effects on others. This contagion mechanism, as explored by Metiu (2012), includes both direct linkages through trade and financial channels and indirect effects via changes in market sentiment. De Grauwe and Ji (2013) highlight the role of self-fulfilling dynamics, where rising yields increased borrowing costs, further straining public finances and deepening the crisis. Missio and Watzka (2011) document the spread of contagion across the Euro Area, noting how market perceptions of one country's fiscal distress influenced the yields of others, particularly in the periphery. Simultaneously, investors fled to safe-haven assets such as German Bunds, amplifying the divergence in spreads. This phenomenon of "flight-to-quality" was particularly acute during the peak of the crisis in 2011, with Italian bond yields surpassing 7% and Spanish yields following closely. Conversely, the "flight-to-quality" mechanism exacerbates these effects, as capital moves from distressed economies to safer ones. Beber et al. (2009) provide evidence of "flight-to-quality" and "flight-to-liquidity" phenomena during periods of market distress. Their research reveals that investors disproportionately allocate funds to low-risk, highly liquid assets like German Bunds when faced with systemic uncertainty. This behavior amplifies yield spreads, particularly for countries with weaker fiscal positions or less liquid bond markets. The crisis also highlighted the issue of redenomination risk—the fear that some countries might leave the Eurozone and redenominate their debt in a devalued national currency. De Santis (2015) shows how this risk significantly influenced bond spreads during the crisis, particularly in peripheral countries, as markets priced in the possibility of a Euro area breakup. Studies by Zaghini (2016) on corporate bond markets and Gabrieli and Labonne (2018) on interbank lending also highlight how the investors' behavior during the crisis also extends beyond sovereign bond markets.

Policy responses to the crisis evolved over time, beginning with *ad hoc* bilateral loans to Greece and eventually culminating in the creation of the European Stability Mechanism (ESM) in 2012. However, the most decisive interventions came from the European Central Bank (ECB). Mario Draghi's "Whatever it takes" speech in July 2012 marked a turning point, signaling the ECB's commitment to preserving the Euro. This announcement was followed by the launch of the Outright Monetary Transactions (OMT) program, which allowed the ECB to purchase sovereign bonds of distressed countries under strict conditions. Altavilla et al. (2015) demonstrate the effectiveness of the ECB's actions in reducing spreads and restoring market confidence.

Despite the ECB efforts to narrow yield spreads and the fiscal reforms undertaken by several Euro area countries, the COVID-19 pandemic reignited concerns about financial fragmentation within the Euro area. The implementation of large-scale fiscal measures to address the economic impact of the pandemic led to surging debt levels, particularly in peripheral countries. Candelon et al. (2022) emphasize that the divergence between core and

peripheral countries persisted beyond the ESD crisis, albeit at reduced levels, as structural imbalances and high debt levels remained unresolved. Costola and Iacopini (2023) provide further evidence that fragmentation risk remains a pressing issue. Using a time-varying cointegration framework, they propose an indicator to assess the probability of fragmentation in the Eurozone sovereign bond market. Their findings reveal that the probability of fragmentation increases during systemic stress events, such as the COVID-19 pandemic and decreases in response to ECB interventions aimed at stabilizing markets.

This residual fragmentation reflects ongoing concerns about fiscal sustainability, economic resilience and the adequacy of institutional frameworks, especially regarding the lack of fiscal integration within the Euro. Addressing those concerns requires both short-term interventions and long-term structural reforms. In the short term, the ECB's unconventional monetary policies, such as asset purchase programs and targeted longer-term refinancing operations, have been effective in mitigating immediate risks. However, long-term solutions must address the root causes of fragmentation. Proposals for Eurobonds or a centralized fiscal authority aim to pool risks and ensure more uniform market treatment of Euro area sovereigns. Corsetti and Dedola (2016) argue that the absence of such centralized fiscal authority creates asymmetric exposure to market pressures, as individual countries bear the full burden of risk perception. This dynamic was at play during the ESD crisis, when the solvency concerns of weaker economies were not counterbalanced by collective fiscal guarantees. Additionally, the crisis underscored the importance of fiscal integration with ongoing debates about the potential role of Eurobonds and mutualized debt as tools for addressing future crises (Ando et al. (2023) ; Gilbert et al. (2013)).

2.3 Machine Learning

The existing literature on financial fragmentation, while extensive, faces limitations in addressing the complexity and dynamism of global macroeconomic and financial markets. Traditional econometric approaches often fail to account for a broad range of variables, such as demographic trends, survey data and market sentiment, which play critical roles in shaping the investor's behavior. Moreover, by construction, these methods typically struggle to capture the intricate relationships and interactions among these diverse set of variables. In contrast, Machine Learning (ML) techniques —such as LASSO, decision trees or neural networks—excel at modeling such complexities by leveraging their ability to identify non-linear patterns and manage large datasets effectively.

The advent of Big Data has further highlighted the constraints of traditional econometric models, particularly their susceptibility to the "Curse of dimensionality" (Bellman (1966)). This issue arises when an overabundance of predictors (i.e. explanatory variables) introduces high parameter estimation errors, limiting the statistical reliability of the model. Machine Learning algorithms, with their built-in capabilities for variable selection and regularization, effectively address this challenge. By focusing on the most relevant variables and ignoring less informative ones, these methods significantly enhance model precision and reduce overfitting.

The adoption of machine learning techniques has pushed the boundaries of research in financial economics, expanding the scope of data analysis and significantly improving predictive accuracy. Recent literature has demonstrated the potential of ML to redefine research methodologies in the field. For instance, Strader et al. (2020) identify promising directions for applying Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and other artificial intelligence techniques to stock market prediction. Similarly, Gu et al. (2020) show that ML-based approaches, such as decision trees and ANNs, deliver substantial economic gains in asset pricing forecasts, outperforming traditional regression-based models by a significant margin. Moreover, Fouliard et al. (2021) demonstrate the utility of ML in developing early warning indicators for macro-financial crises, further solidifying its role in predictive modeling.

Machine learning has also found applications in sovereign bond markets, particularly in predicting bond spreads. Early contribution by Castellani and Santos (2006) explored the potential of ML for forecasting the monthly yield of 10-year U.S. Treasury bonds. More recently, Kim et al. (2020) compared various ML models, such as recurrent neural networks (RNNs), long short-term memory (LSTM), and support vector regression (SVR), in analyzing credit default swap (CDS) term structures. Their results demonstrated that ML approaches consistently outperformed traditional models, such as the Nelson-Siegel framework, in forecasting future term structures.

The Euro Area has similarly benefited from ML applications in sovereign bond spread analysis. Three recent studies stand out in advancing this field. First, Arakelian et al. (2019) employed recursive partitioning strategies, including Random Forest and regression trees, to stratify European sovereign risk using macroeconomic fundamentals and contagion metrics. Their findings reveal a marked decline in Euro Area CDS contagion after the European Sovereign Debt Crisis (2013–2017), attributed to improved contagion mitigation strategies. Second, Balduzzi et al. (2022) highlight the temporal variability in the relationship between macroeconomic fundamentals and Euro Area CDS spreads. By employing LASSO regression, they identify distinct macro-sensitive regimes, explaining fluctuations in sovereign risk pricing over time. Their study underscores the evolving dynamics of sovereign bond markets and the limitations of static models in capturing such changes. Third, Belly et al. (2022) explore the effectiveness of multiple ML techniques—including XGBoost, SVR, Elastic Net regression, Random Forest, and ANN—in valuing and pricing Euro Area sovereign risk. Their findings demonstrate that ML methods not only outperform traditional Bayesian Model Averaging (BMA) approaches but also capture the multifaceted drivers of sovereign risk. These include macroeconomic indicators, global financial factors, and public sentiment, with the latter derived from innovative data sources such as Google Trends.

In this paper, we acknowledge the previous literature both in the field of financial fragmentation and of machine learning by first running a horse race between 14 machine learning techniques in order to identify the algorithm that fits our dataset best. Then, we predict the 10-year sovereign spread of ten Euro Area countries over a 6-month horizon by leveraging the advanced XGBoost algorithm, providing valuable information for policymakers and investors.

3 Data

This paper introduces a novel dataset that stands out in its dimensionality. Aptly fitting the criteria of "Big Data", this dataset is characterized by its high dimensionality, encompassing a diverse range of categories of variables across ten countries (namely Austria, Belgium, Finland, France, Greece, Ireland, Italy, the Netherlands, Portugal and Spain). The dataset includes monthly time series obtained through Bloomberg, over the March 2007 to January 2024 period, leading to a total of 902,741 individual observations, as detailed in Table 1. All those observations fall within the distinct categories presented in Table 2.

Table 1: Variable and observation count

Country	Number of variables	Number of observations
Austria	193	39,179
Belgium	635	128,905
Finland	467	94,801
France	1,363	276,689
Greece	325	65,975
Ireland	368	74,704
Italy	316	64,148
Netherlands	197	39,991
Portugal	210	42,630
Spain	373	75,719
Total	4,447	902,741

The 10-year sovereign spread is a prevalent metric among market participants and academics for evaluating sovereign risk. We build our 10-year sovereign spread as the difference between a country's 10-year sovereign yield and the risk-free rate. For this purpose, we adopt the German 10-year bond yield as the risk-free rate, reflecting its widespread acceptance as the standard measure in both market practice and academic research for calculating Euro Area spreads.

The vastness and granularity of this dataset not only provide a comprehensive view of each country's economy but also present unique challenges and opportunities for data processing, analysis and modeling. By leveraging the dataset's potential, this paper is able to identify intricate relationships between the idiosyncratic characteristics of countries and their yield spread, potentially generating more accurate and precise predictions. Hence, the introduction of this novel dataset represents a notable contribution to the field, enabling a more refined and comprehensive understanding of Euro area country's yield spread dynamics.

Table 2: Categories of variables for each country

National Accounts	Services Sector	Whole Economy Activity
Surveys & Cyclical Indicators	International Trade	Monetary Sector
Labor Market	Price Indexes	Government Finance & Debt
Retail Sector	Housing & Real Estate	Financial Indicators
Industrial Sector	Household Sector	Demographics

Tables 3 and 4 present descriptive statistics for sovereign bond yields and bond spreads in eleven and ten Euro area countries respectively.For each country, the mean, median, standard deviation, variance, range, interquartile range (IQR), skewness, kurtosis, minimum, and maximum values are reported. This subset of data consists of 203 monthly observations per country, spanning from March 2007 to January 2024, with all figures rounded to two decimals.

Table 3 presents the descriptive statistics for sovereign bond yields across eleven Euro area countries. The mean bond yields reflect notable differences in borrowing costs across the Euro area, with Greece exhibiting a significantly higher mean yield of 7.20%, highlighting heightened risk premium and credit risk perceptions. Portugal, Italy and Ireland also report higher mean yields, above 3%, relative to core Euro Area countries such as Germany, France, Finland, Austria and the Netherlands in which mean yields are below 2% on average. Median values align with mean yields for most countries indicating symmetric distributions. Greece's median yield of 5.32% is notably lower than its mean, hinting at positive skewness driven by periods of elevated yields, such as during the European Sovereign Debt crisis . The standard deviation and variance indicate variability in bond yields, with Greece showing the highest standard deviation at 6.13, signifying high volatility relative to other Euro area nations. This heightened volatility is further evidenced by Greece's wide range of yields (from 0.60% to 32.60%), reflecting episodes of severe market stress during financial crises. Portugal and Ireland also exhibit substantial volatility (standard deviation above 2.85 for both countries), underscoring the uncertainty and risk during their periods of economic turmoil. Conversely, countries like Austria, Finland, France, Germany and the Netherlands demonstrate relatively lower variability as indicated by their lower standard deviations (below 1.6), variance (below 2.5) and narrower ranges (below 540 basis points). The IQR, which measures the difference in yields between the 75th and 25th percentiles of each countries' distribution, further highlights this trend with Greece and Ireland showing the widest interquartile ranges (390 bps and above) while core countries exhibiting narrower IQRs, indicating relative stability. Skewness values reveal the asymmetry in bond yield distributions. Positive skewness in Greece, Portugal and Ireland reflects episodes of sharply increased yields while Italy, Spain, Belgium and France show minimal skewness, suggesting a more symmetric distribution of yields over the sample period. Kurtosis values provide insight into the distributional tails, with Greece having a high kurtosis value indicating

a distribution with frequent extreme values typically associated with financial instability. Other countries, such as Austria and Belgium, display more moderate kurtosis, suggesting a lower occurrence of extreme values.

Table 3: Descriptive statistics for Euro area sovereign bond yields

Country	Mean	Median	Std. Dev	Variance	Range	IQR	Skew.	Kurt.	Min	Max
Austria	1.83	1.73	1.57	2.47	5.33	2.68	0.24	1.68	-0.44	4.89
Belgium	2.04	2.08	1.63	2.66	5.36	2.96	0.15	1.56	-0.39	4.97
Finland	1.72	1.54	1.52	2.31	5.24	2.60	0.33	1.78	-0.43	4.80
France	1.88	1.99	1.48	2.20	5.22	2.49	0.18	1.71	-0.41	4.81
Germany	1.46	1.31	1.48	2.19	5.32	2.34	0.42	2.03	-0.70	4.62
Greece	7.20	5.32	6.13	37.60	32.01	5.19	2.06	7.82	0.60	32.60
Ireland	3.06	2.59	2.86	8.16	11.76	3.90	0.90	2.92	-0.31	11.45
Italy	3.20	3.48	1.56	2.44	6.49	2.71	-0.02	1.87	0.54	7.03
Netherlands	1.69	1.63	1.53	2.35	5.36	2.48	0.32	1.84	-0.55	4.81
Portugal	3.87	3.21	2.96	8.77	15.64	2.84	1.31	4.97	0.03	15.67
Spain	2.84	2.98	1.79	3.21	6.79	2.87	0.12	1.78	0.04	6.83

Notes: Bond yields across 11 Euro area countries under study, based on 203 monthly observations per country from March 2007 to January 2024.

Table 4: Descriptive statistics for Euro area sovereign bond spreads

Country	Mean	Median	Std. Dev	Variance	Range	IQR	Skew.	Kurt.	Min	Max
Austria	0.37	0.29	0.24	0.06	1.13	0.27	1.24	4.25	0.04	1.18
Belgium	0.57	0.44	0.43	0.18	2.64	0.37	2.16	8.56	0.06	2.70
Finland	0.26	0.24	0.16	0.03	0.79	0.15	1.01	3.56	-0.01	0.78
France	0.42	0.36	0.22	0.05	1.28	0.22	1.63	6.85	0.04	1.31
Greece	5.74	3.68	6.04	36.43	30.58	6.47	2.09	7.86	0.21	30.79
Ireland	1.60	0.60	2.13	4.54	9.09	1.26	1.75	4.65	-0.66	8.43
Italy	1.73	1.56	0.94	0.88	5.01	0.90	1.18	4.73	0.20	5.20
Netherlands	0.23	0.20	0.13	0.02	0.78	0.18	1.29	5.48	0.03	0.81
Portugal	2.40	1.47	2.54	6.44	13.77	2.40	2.01	7.03	0.11	13.88
Spain	1.38	1.08	1.06	1.13	5.46	0.69	1.75	5.94	0.04	5.50

Notes: Bond spreads across 10 Euro area countries under study, based on 203 monthly observations per country from March 2007 to January 2024. Germany is used as the benchmark for each other country.

The descriptive statistics in Table 4 offer insights into the properties of the sovereign bond spreads of each Euro area country under study relative to Germany, which serves as the benchmark country (risk-free). The mean values reveal variation in spread levels across countries. Greece exhibits the highest average spread at 5.74%, while countries such as Austria, Belgium, Finland, France or the Netherlands maintain lower average spreads.

Countries like Ireland, Italy, Portugal and Spain are in between, with spreads above 1% on average but below 3%. Median values align again with mean values for most countries, suggesting generally symmetric distributions with Greece being an exception due to its significantly higher median and a broad range of spreads. The standard deviation and variance provide insights into the volatility of spreads with Greece, Ireland and Portugal showing the highest values. Conversely, countries like Austria, Finland, France and the Netherlands display lower volatility, highlighting their relative stability. The range further accentuates the variability among countries with Greece, Ireland, Italy, Portugal and Spain showing extensive ranges (above 500 bps), whereas Finland and the Netherlands exhibit narrower ranges (below 100 bps). The interquartile range (IQR) follows similar patterns with Greece and Portugal standing out due to their wider IQRs. Skewness and kurtosis provide information on the distributional shape of spread yields. Positive skewness values across most countries indicate right-skewed distributions, suggesting the presence of periods with abnormally high spreads, particularly in countries such as Greece, Belgium and Portugal (above 2), where the risk premia were markedly higher during periods of financial distress. High kurtosis values for several countries, especially Belgium, Greece and Portugal indicate heavier tails, suggesting frequent extreme values or outliers in their spread distributions.

These descriptive statistics highlight the substantial heterogeneity in bond yields and spreads behaviors across Euro area countries with distinct contrasts in average levels, volatility and distributional characteristics. These differences may give hints on whether a country belongs to the core or the peripheral country clusters. These differences highlight the divergent divergent credit risk perceptions and market conditions that have characterized Euro area sovereign bond markets over the past decade, reflecting both structural differences and the impact of economic and financial crises on specific member states.

4 Comparing Machine Learning techniques : a Horse Race

In this section, we begin by detailing each of the 14 Machine Learning techniques included in our horse race. Subsequently, we present the results of this horse race.

4.1 Overview of the Machine Learning techniques

In order to find evidence of Financial Fragmentation, this paper exploits our novel high-dimensional dataset. With so much data, conventional econometric models are not suitable. This is why we employ machine learning. Indeed, machine learning offers a plethora of techniques for variable selection, aiding in the identification of the most informative predictors in high-dimensional datasets. In this paper, we perform a horse race of linear, non-linear and rule-based techniques to find which machine learning methodology best fits our data. The 14 Machine Learning techniques competing in this horse race are : (1) Standard LASSO, (2) Boosted LASSO, (3) Group LASSO, (4) Sparse-Group LASSO, (5) Ridge regression, (6) Elastic Net regression, (7) Support Vector Regression, (8) Feedforward Neural Network, (9) Decision Trees, (10) Regularized Trees, (11) Boosted Trees, (12) Stochastic Gradient Boosting, (13) Extreme Gradient Boosting and (14) LightGBM, as presented in Table 5.

4.1.1 Linear Techniques

The group of linear techniques encompasses the LASSO, the Ridge and the Elastic Net regressions. Before examining each linear technique in detail, it is worth noting that they suffer from scale dependency. The algorithm's effectiveness in selecting variables and assigning coefficients is influenced by the scale of the predictor variables. Linear techniques penalize the absolute size of coefficients, shrinking less important coefficients toward zero. However, the magnitude of each variable's coefficient depends on the variable's scale. For example, in a dataset with two predictors, one measured in thousands and the other in single units, LASSO would penalize the larger-scaled variable more heavily because its raw coefficient value would typically be larger, regardless of its relative importance. This imbalance can lead LASSO to shrink coefficients unevenly, selecting or discarding variables based on their scale rather than their true contribution to the model. To mitigate this effect, we standardize the predictor variables (mean of zero and standard deviation of one). By standardizing our data, we ensure that each variable's scale is comparable, allowing LASSO to more accurately apply penalties based on variable importance rather than scale differences.

Table 5: Machine Learning Techniques and Accuracy Metrics

Machine Learning Techniques	**Nb**	**Accuracy metrics**
Linear		
Standard LASSO	1	Cross-Validation using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R^2 metrics
Boosted LASSO	2	
Group-LASSO	3	
SG-LASSO	4	
Ridge	5	
Elastic Net	6	
Non-linear		
Support Vector Regression (SVR)	7	Cross-Validation using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R^2 metrics
Feedforward Neural Network (FNN)	8	
Rule-based		
Trees		
Decision Trees	9	Cross-Validation using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R^2 metrics
Regularized Trees	10	
Boosted Trees	11	
Stochastic Gradient Machines (SGM)		
SG Boosting	12	Cross-Validation using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R^2 metrics
XGBoost	13	
LightGBM	14	

Notes : Cross-validation is a technique for assessing model performance by splitting data into training and validation sets (80% and 20% respectively in our models). Root Mean Squared Error (RMSE) measures the average magnitude of prediction errors, penalizing larger errors, calculated as the square root of MSE. Mean Absolute Error (MAE) measures the average absolute difference between predictions and actual values, reflecting typical prediction error without penalizing large errors. R-squared quantifies how well the model explains variance in the target, ranging from 0 (no fit) to 1 (perfect fit).

1 The LASSO models

First, the LASSO regression is a statistical method that performs both variable selection and regularization in order to enhance the prediction accuracy of the statistical model it produces. It modifies the OLS objective function by adding a penalty equivalent to the absolute value of the magnitude of coefficients. Mathematically, the LASSO objective function can be expressed as:

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j|, \tag{1}$$

Where, y_i is the observed response for the i-th observation, $\hat{y}_i$ is the predicted response, calculated as $\hat{y}_i = X_i\beta$, where X_i is the variable vector for the i-th observation, and β is the vector of coefficients, n is the number of observations, p is the number of predictors and λ_1 is the tuning parameter that controls the strength of the penalty imposed on the coefficients. By shrinking the coefficients, LASSO controls the model's complexity, ensuring that it captures the underlying pattern in the data without being overly sensitive to the training set. The regularization parameter λ_1 plays a critical role here; as λ_1 increases, the flexibility of the LASSO model decreases, leading to less complex models. The second term of the equation is called the $L1$ penalty term.

Over the years, the regular LASSO methodology has been improved. New variants of LASSO have been developed, notably the Boosted LASSO, the Group LASSO and the Sparse-Group LASSO. Boosted LASSO merges the concepts of LASSO and boosting, a powerful ensemble technique. Boosting involves building a model from the training data, then creating a second model that attempts to correct the errors from the first model. By combining this with LASSO's regularization and its selection capabilities, Boosted LASSO can handle complex datasets with intricate structures more effectively than regular LASSO. Group LASSO extends the LASSO technique to situations where predictors can be naturally grouped, and the groups, rather than individual predictors, are of interest. In this variant, the $L1$ penalty term is applied to the norms of the coefficients of the groups, rather than to the individual coefficients. Group LASSO ensures that either all the coefficients in a group are zero or none of them are, thus respecting the intrinsic grouping in the data. Finally, Sparse-Group LASSO is a sophisticated hybrid that combines the ideas of both group and sparse LASSO. It allows for both group-wise and individual variable selection, making it a powerful tool for models where both individual and grouped variable effects are of interest. This variant is particularly useful in models where some groups of variables are known to be important but there is also a need to identify individual variables of significance outside these groups. Sparse-Group LASSO applies a mixed penalty, combining the $L1$ penalty term to encourage sparsity at the individual level and a group-level $L2$ penalty term.

2 The Ridge model

Second, the Ridge regression is a technique used to analyze multiple regression data that suffer from multicollinearity. When multicollinearity occurs, least squares estimates are unbiased but their variances are large so they may be far from the true value. Ridge regression stabilizes the regression estimates in such a way that it reduces the standard errors. The key idea of Ridge regression is to add a degree of bias to the regression estimates, which in return reduces the standard errors. It modifies the OLS objective function by adding a penalty equivalent to the square of the magnitude of the coefficients. The Ridge regression objective function can be written as:

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda_2 \sum_{j=1}^{p} \beta_j^2. \tag{2}$$

As for the LASSO method, λ_2 is the tuning parameter that controls the amount of shrinkage: the larger the value of λ_2, the greater the amount of shrinkage. Through λ_2, the Ridge method applies a penalty to the size of coefficients. However, unlike LASSO which can eliminate some coefficients altogether by setting them to zero, Ridge regression only shrinks the coefficients. This is particularly useful when dealing with data where all variables are important and should be retained in the model. The second term of the equation is called $L2$ penalty term.

3 The Elastic Net model

Third and last among the linear methods, Elastic Net is a regularization and variable selection method that combines the properties of both LASSO and Ridge regressions. Elastic Net aims to overcome limitations of both Ridge and LASSO by combining their penalty terms. Hence, the objective function of Elastic Net can be written as:

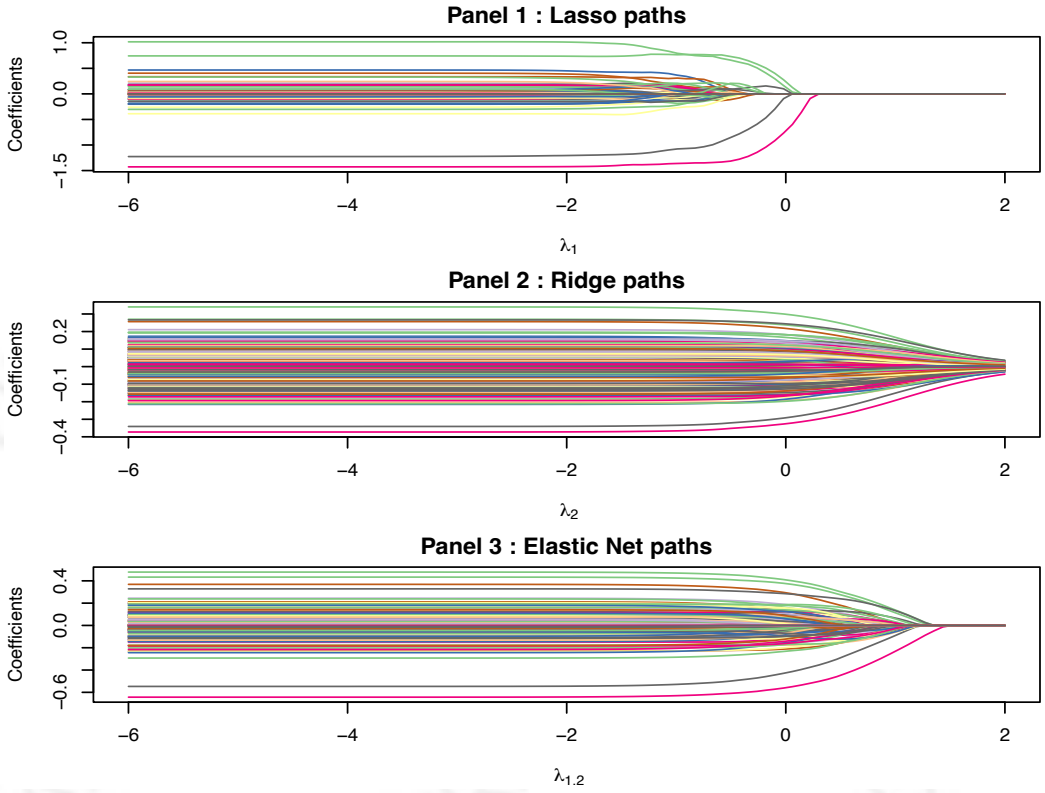
$$\min_{\beta} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j| + \lambda_2 \sum_{j=1}^{p} \beta_j^2. \tag{3}$$

Here, λ_1 and λ_2 are the parameters that control the $L1$ (LASSO) and $L2$ (Ridge) penalty terms, respectively. Elastic Net thus balances the properties of LASSO and Ridge regression methods. The $L1$ penalty term facilitates variable selection, making the model simpler and more interpretable. The $L2$ penalty term shrinks the coefficients towards zero, but not exactly zero, which stabilizes the model estimation process. In our benchmark, we put the same weight on the $L1$ and $L2$ penalty terms, that is 50% on each penalty term. One of the significant advantages of Elastic Net is its ability to handle correlated predictors effectively. In situations where there are groups of correlated variables, LASSO tends to select only

one variable from a group and ignore the others. Elastic Net, by contrast, tends to select groups of correlated variables. In addition, Elastic Net provides a continuum of solutions that range from Ridge to LASSO depending on the parametrization of the λ. By tuning the parameters λ_1 and λ_2, one can explore a range of models between the two extremes.

To visualize the LASSO, Ridge and Elastic Net methods in action, it is possible to plot the regularization paths (the coefficients) against the regularization parameter (λ). Figure 3 presents coefficient paths for the three types of regularization methods using simulated data (50 observations for 200 variables with only 10 contributing to the response). Each panel in the figure visualizes how the coefficients for various variables evolve as the regularization strength, represented by λ, changes. The x-axis in each plot is on a logarithmic scale, indicating increasing values of λ from left to right. This corresponds to a transition from weak to strong regularization. The y-axis denotes the values of the coefficients (or weights) associated with each variable, where each colored line represents the path of a single variable's coefficient across different values of λ.

Figure 3: LASSO, Ridge and Elastic Net paths



The first panel displays the coefficient paths for the LASSO model. As λ_1 increases and regularization strengthens, most coefficients are progressively driven to zero, with only a select few remaining non-zero at the highest values of λ_1. This behavior is a consequence of the LASSO method's $L1$ penalty term, which induces sparsity in the model. For small values of λ_1, the penalty term has minimal impact, allowing many coefficients to remain

active. As λ_1 grows, however, the $L1$ penalty becomes more influential, eliminating the smaller coefficients while retaining only those that are most impactful for the model.

The second panel illustrates the coefficient paths for the Ridge model. Here, the trends differ from those observed in the LASSO paths. Although the Ridge model similarly causes coefficients to diminish as λ_2 increases, none of the coefficients are forced to zero; instead, they decrease smoothly across the entire range of λ_1. Unlike LASSO, Ridge regression employs an $L2$ penalty, which penalizes the square of the coefficients, leading to gradual shrinkage without zeroing out any of the coefficients. This penalty structure allows all variables to remain active, with their magnitudes reduced according to the strength of the regularization.

The third panel shows the coefficient paths for the Elastic Net model, which combines the penalty terms of both LASSO ($L1$) and Ridge ($L2$). This panel reflects a mix of sparsity and smooth shrinkage: as $\lambda_{1,2}$ increases, some coefficients shrink to zero as in LASSO, while others decrease gradually as in Ridge. The parameter α controls the balance between both $L1$ and $L2$ penalty terms. Here, $\alpha = 0.5$, so the $L1$ and $L2$ penalties contribute equally, merging LASSO's sparsity with Ridge's smooth shrinkage.

4.1.2 Non-Linear Techniques

Previous machine learning techniques were linear, but in a complex network introducing non-linear approaches could be useful. Among these, Support Vector Regression (SVR) and Neural Networks (NN) stand out for their versatility and power.

1 The Support Vector Regression (SVM) model

On one hand, Support Vector Regression (SVR) works by finding a curve that best fits the data while allowing some flexibility in error tolerance. Indeed, unlike traditional regression models, SVR allows a tolerance ε around the predictions, within which errors are not penalized. This tolerance is known as the ε-tube or support vectors. Observations within this tube are considered close enough and don't add to the error, while observations outside the tube are penalized based on how far they fall from it. This approach helps SVR capture the general trend of the data without over-fitting to noise. Since observations must share the same scale to form the ε-tube, we standardize our data sample (mean of zero and standard deviation of one).

To achieve this, SVR minimizes the coefficients (represented as the weight vector $\|\mathbf{w}\|$) of the function that best fits the data, ensuring a flatter, simpler model that generalizes well to big data. Simultaneously, the model introduces slack variables (i.e. measures of constraint deviation) ξ_i and ξ_i^* for each observation i, representing deviations outside the ε-tube. These slack variables allow for some observations to lie outside the margin when necessary, especially when the data cannot be perfectly fit within the ε-tube.

Combining those terms yields the SVR objective function and can be formulated as follows:

$$\underset{\mathbf{w},b,\xi,\xi^*}{\text{minimize}} \quad \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*), \tag{4}$$

Where C is the regularization parameter that controls the trade-off between the model's complexity with the amount of error tolerated outside the ε-tube. A larger C places more emphasis on reducing deviations outside the margin, while a smaller C prioritizes a simpler model with potentially higher deviations.

The constraints for this optimization problem enforce that the residuals of each prediction lie within the ε-tube, with exceptions allowed only for the slack variables:

$$y_i - (\mathbf{w}\cdot\mathbf{x}_i + b) \leq \varepsilon + \xi_i, \tag{5}$$

$$(\mathbf{w}\cdot\mathbf{x}_i + b) - y_i \leq \varepsilon + \xi_i^*, \tag{6}$$

Where $\xi_i, \xi_i^* \geq 0$. These constraints ensure that, for data points within the ε-tube, no penalty is incurred, while points outside the tube are penalized in proportion to their distance from the tube boundary.

To solve this problem, SVR can be reformulated to allow the "kernel trick" -a method to implicitly map data into higher-dimensional spaces without explicitly computing these transformations- to be applied. This is the reason is why SVR is interesting to us since it is is particularly powerful for regression tasks where a non-linear relationship may exist. The problem then becomes:

$$\underset{\alpha,\alpha^*}{\text{maximize}} \quad -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(\mathbf{x}_i,\mathbf{x}_j) + \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)y_i - \varepsilon\sum_{i=1}^{n}(\alpha_i + \alpha_i^*), \tag{7}$$

subject to

$$\sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0, \quad 0 \leq \alpha_i, \alpha_i^* \leq C. \tag{8}$$

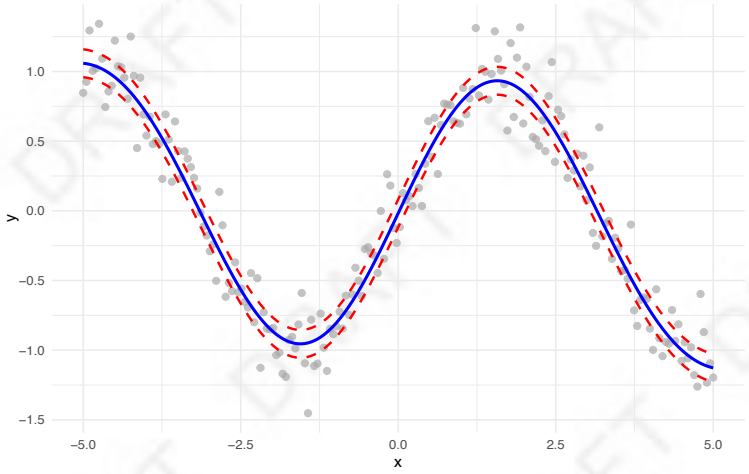
Here, $K(\mathbf{x}_i, \mathbf{x}_j)$ is our kernel function allowing us to unlock the model's non-linearity. In addition, α_i and α_i^* are the Lagrange multipliers representing how much each observation influences the position and shape of the regression function. The solution to this formulation results in the regression function that depends only on a subset of observations : those outside the ε-tube.

Finally, the SVR prediction for a new observation is calculated using the learned regression function. The final prediction function is:

$$f(\mathbf{x}) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b, \tag{9}$$

Where α_i and α_i^* are the optimized Lagrange multipliers, $K(\mathbf{x}_i, \mathbf{x}_j)$ is our kernel function and b is the error term. Figure 4 presents the result of a SVR prediction using simulated data and a Gaussian kernel to illustrate the model's non-linearity capability. The blue line depicts the SVR prediction while the grey points are the original and simulated data. The red dashed line is the ε-tube.

Figure 4: Support Vector Regression prediction with ε-tube and Gaussian kernel



Notes: The blue line depicts the SVR prediction while the grey points are the original and simulated data. The red dashed line is the ε-tube. This SVR model uses 200 simulated observations and a Gaussian kernel.

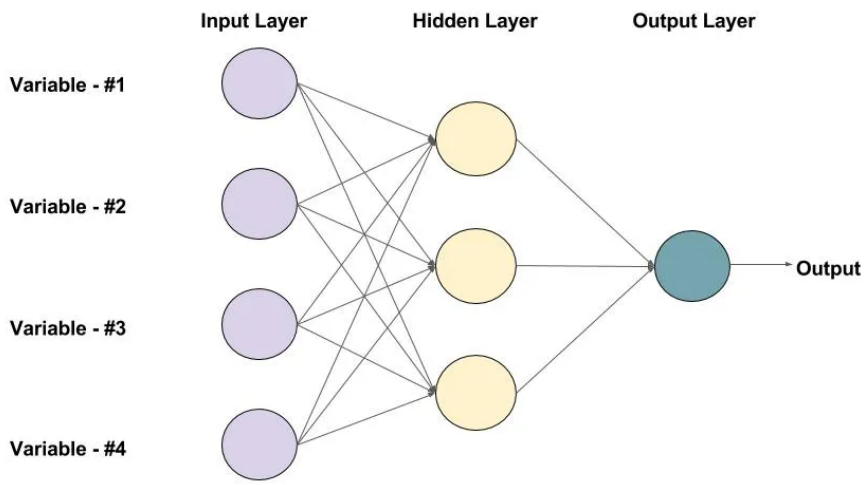
2 The Feedforward Neural Network (FNN) model

On the other hand, Neural Networks (NN) are a series of algorithms, modeled loosely after the human brain, that are designed to recognize patterns and make predictions. They consist of layers of interconnected nodes (or "neurons") that process input data in a way that allows the network to learn and make complex decisions. Among the wide range of Neural Network sub-models, we are interested in Feedforward Neural Networks (FNN) since they are effective with time series, can learn non-linear relationships and make accurate regression predictions. The working of a feedforward neural network involves two phases: (i) the Feedforward phase and (ii) the Backpropagation phase.

In the Feedforward phase, our data (non-standardized) is introduced into the input layer of the FNN model and propagates through each hidden layer sequentially until reaching the final output layer. In each hidden layer, the network computes a weighted sum of the inputs and applies an activation function, introducing non-linearity to enhance the model's capacity for pattern recognition. This forward propagation continues until the FNN generates a prediction at the final output layer. Figure 5 illustrates the functioning of a FNN with one hidden layer and 3 neurons. Our FNN model is designed to process the data through 10 hidden layers with 5 neurons.

Then, the Backpropagation phase begins. After a prediction is generated, the FNN calculates the error, defined as the difference between the predicted output and the actual target value. This error is then propagated backward through the network layers, with the goal of adjusting the weights to minimize the error. Weight adjustment is achieved through a gradient descent optimization algorithm, which systematically reduces the error by modifying the weights in proportion to their contributions to the prediction error. This iterative refinement of weights enhances the model's accuracy over successive training cycles.

Figure 5: Architecture of a Feedforward Neural Network (FNN)



Let's consider a FNN with L layers, where each layer l contains a certain number of neurons. Each neuron in a layer computes a weighted sum of its inputs, adds a bias, and then applies a non-linear activation function. A single neuron j in layer l receives input signals $x_1, x_2, \ldots, x_n$ from the neurons in the previous layer, each associated with a weight $w_{1j}, w_{2j}, \ldots, w_{nj}$. The neuron computes a weighted sum of these inputs:

$$z_j = \sum_{i=1}^{n} w_{ij} x_i + b_j, \tag{10}$$

Where b_j is a bias term that allows the neuron to adjust its output independently of the input. This weighted sum z_j is then passed through an activation function $\phi(z)$, which introduces non-linearity to the model and enables the network to approximate non-linear patterns in the data.

In our model, we use the Tanh activation function as our activation function in the hidden layers, defined as:

$$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}, \tag{11}$$

Where z is the weighted sum of the inputs plus the bias term, as described in Equation (10). The Tanh function maps values to a range between -1 and 1, which centers the data around zero, reduces bias in activations and improves convergence during training.

In a FNN, data is processed through multiple layers by iteratively applying the weighted sums, bias adjustments and Tanh activations in each hidden layer. For layer l, we denote the transformation from the previous layer's output $\mathbf{a}^{(l-1)}$ as:

$$\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}, \tag{12}$$

Where $\mathbf{W}^{(l)}$ represents the weight matrix for layer l, $\mathbf{b}^{(l)}$ is the bias vector, and $\mathbf{z}^{(l)}$ represents the weighted input to layer l. The output of layer l is then calculated by applying the Tanh activation function element-wise to each value in $\mathbf{z}^{(l)}$:

$$\mathbf{a}^{(l)} = \tanh(\mathbf{z}^{(l)}), \tag{13}$$

This feedforward process is repeated through each hidden layer, allowing the network to learn non-linear transformations of the data. After processing through our 10 hidden layers, the data reaches the final output layer for prediction.

In prediction tasks, the output layer uses a linear activation function, as it allows the network to output a continuous range of values. This is particularly important for predicting values like bond spreads, which can vary widely depending on the country under study. Unlike the Tanh activation function that restricts values between $[-1, 1]$, the linear activation function in the final output layer simply outputs the weighted sum of the inputs without additional transformation. The final prediction, $\hat{y}$, is given by:

$$\hat{y} = \mathbf{W}^{(L)}\mathbf{a}^{(L-1)} + b^{(L)}, \tag{14}$$

Where L is the output layer, $\mathbf{W}^{(L)}$ is the weight matrix, and $b^{(L)}$ is the bias term.

The feedforward phase ends by obtaining the predicted value and allows to enter the second phase : training the model. Training involves backpropagation, an optimization method that calculates gradients of a loss function with respect to each weight and bias in the FNN. For n number of samples, we measure the difference between the predicted output $\hat{y}$ and the actual target value y using the following loss function :

$$\mathcal{L}(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \tag{15}$$

During backpropagation, the network updates weights by moving in the opposite direction of the gradient to minimize the loss. For each weight $w_{ij}^{(l)}$ in layer l, the update rule is:

$$w_{ij}^{(l)} \leftarrow w_{ij}^{(l)} - \eta \frac{\partial \mathcal{L}}{\partial w_{ij}^{(l)}}, \tag{16}$$

Where η is the learning rate, a parameter that controls the size of each adjustment. This iterative update continues until the network converges to an optimal set of weights that minimizes the loss and provides a definitive final prediction.

4.1.3 Rule-based Techniques

Rule-based techniques involve methods that infer logical rules from the data. Among those techniques, tree-based methods are able to capture complex interactions among variables through their branching structure and are inherently explainable because each split in the tree follows a decision criterion. Tree models include Decision Trees, Regularized Trees and Boosted Trees. In addition to tree models, a second category of methods has emerged : the Stochastic Gradient Machines. They encompass a range of machine learning techniques that optimize a cost function iteratively. This category includes Stochastic Gradient Boosting (SGB), XGBoost and LightGBM.

1 Tree-based models

Tree-based models are a category of machine learning algorithms structured around hierarchical decision-making, where predictions are generated through a series of binary decisions that split the data into increasingly refined subsets. Originally designed for classification tasks, tree-based models have been effectively adapted for prediction.

Decision Tree models have the simplest setup. When designed for predicting values, they operate by recursively partitioning the data based on variable values, creating a hierarchical structure of splits to approximate the target variable. The tree begins with a root node that represents the entire dataset. At each node, the model selects the variable and corresponding threshold that best divides the data to minimize prediction error. The decision

tree model seeks to minimize the variance of the target variable within each subset (i.e the datasets after a split) by choosing splits that reduce the overall error. For a given node, the model evaluates candidate splits on variable x to find the split $x \leq s$ that minimizes the Mean Squared Error (MSE) across the resulting subsets. For a split that divides the data D into subsets D_{left} and D_{right} based on a threshold s, the optimal split minimizes the objective:

$$\text{MSE} = \frac{|D_{\text{left}}|}{|D|}\text{Var}(D_{\text{left}}) + \frac{|D_{\text{right}}|}{|D|}\text{Var}(D_{\text{right}}), \tag{17}$$

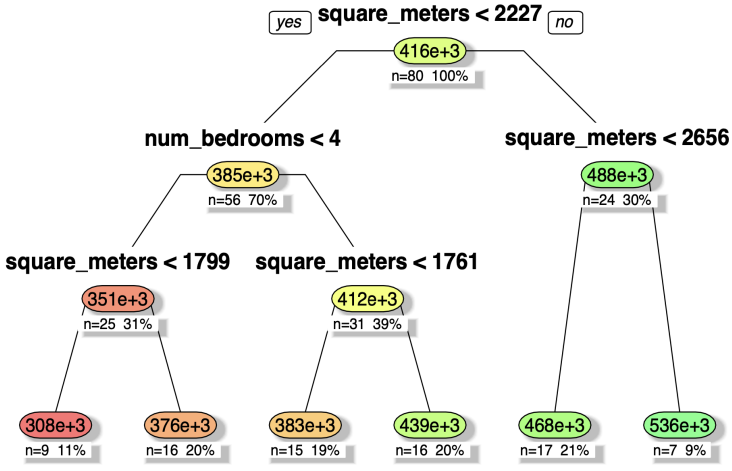
Where $\text{Var}(D_{\text{left}})$ and $\text{Var}(D_{\text{right}})$ represent the variance of the target variable in the left and right subsets, respectively, and $|D|$ denotes the number of samples in the original dataset. The split that results in the lowest weighted average variance is chosen, as it best reduces the prediction error at that node.

Once the tree structure is fully established, each path from the root node to a terminal leaf node represents a set of conditions based on variable values that lead to a final prediction. At each leaf node, the model makes a prediction based on the mean value of the target variable within that subset:

$$\hat{y} = \frac{1}{|D_{\text{leaf}}|} \sum_{i \in D_{\text{leaf}}} y_i, \tag{18}$$

Where D_{leaf} is the set of instances that reach that leaf node, and y_i represents the observed value of the target variable for each instance i. This average provides a stable prediction based on the samples that meet the conditions defined by the path through the tree.

Figure 6: Decision Tree applied to housing prices



The decision tree stops growing when it reaches a pre-set maximum depth or when further splits do not meaningfully reduce the prediction error. This structure allows decision trees to capture non-linear relationships and interactions between variables without requiring any assumptions about the underlying data distribution. using simulated data, Figure 6 illustrates a decision tree model predicting the prices of 100 houses based on two variables: the amount of square meters and the number of bedrooms. At each node, the tree makes a split based on a threshold value for one of these variables, dividing the dataset in subsets to reduce prediction error. Each terminal leaf node shows a predicted house price (in thousands) depending on the two variables. The "n" value at each node indicates the number of samples in that subset, and the percentage reflects the proportion of the dataset in that group. A house with less than 2,227 square meters (first node), more than 4 bedrooms (second node) and less than 1,761 square meters (third and terminal leaf node) is predicted to have a price of 383,000 euros.

Regularized Trees are enhanced decision trees that incorporate constraints to prevent overfitting, thereby improving the model's ability to generalize. In a standard decision tree, splits are made to minimize the variance in each subset using Mean Squared Error (MSE) as the splitting criterion. However, regularized trees add penalties to this process, balancing prediction accuracy with model simplicity. One common regularization approach is pruning, which either restricts tree growth during training (pre-pruning) or removes branches after the tree is fully grown (post-pruning). Pre-pruning stops the tree from splitting further when a stopping criterion is met (e.g., maximum depth). Post-pruning, on the other hand, begins with a fully grown tree and systematically removes branches that contribute minimally to reducing prediction error. This can be formulated as minimizing an objective function that includes a complexity term:

$$\text{Objective} = \text{MSE} + \alpha \times \text{Complexity(T)}, \tag{19}$$

Where α is the regularization parameter and Complexity(T) represents the complexity of the tree T (e.g., the number of leaf nodes or tree depth). This penalty term discourages overly complex trees, ensuring that the final model avoids capturing noise in the training data.Another regularization technique is to penalize splits based on certain criteria, effectively adding a threshold for variance reduction. By setting this threshold, splits are only allowed if they result in a substantial improvement in the target variable's prediction accuracy, reducing unnecessary splits that may lead to overfitting.

However, a single decision tree may suffer from high variance and overfit the training data. To address these limitations, ensemble methods like Stochastic Gradient Machines are employed. Methods such as Boosted Trees, Stochastic Gradient Boosting (SGB), XGBoost or LightGBM sequentially build trees that correct the errors of previous trees and then combine the average of the predictions of the final trees to create a more accurate model.

This is the case for Boosted Trees, which combine multiple decision trees to form a more accurate predictive model by iteratively focusing on errors. Boosted Trees work through an

additive process, where each new tree attempts to correct the residual errors of the preceding trees. This approach involves sequentially building trees and updating the model with each iteration, ultimately creating an ensemble of trees. The boosted model at iteration t is represented as:

$$F_t(x) = F_{t-1}(x) + \eta \cdot h_t(x), \tag{20}$$

Where $F_{t-1}(x)$ is the model from the previous iteration, $h_t(x)$ is the new tree trained on the residuals (errors) of $F_{t-1}(x)$, and η is the learning rate, the parameter controlling the contribution of each new tree to the overall model. By adjusting η, the model can be more cautious in its updates, which helps prevent overfitting.

After a specified number of iterations T, the final prediction for any input x is the cumulative sum of the initial prediction and all adjustments from each tree in the ensemble:

$$\hat{y} = F_T(x) = F_0(x) + \sum_{t=1}^{T} \eta \cdot h_t(x). \tag{21}$$

2 The Stochastic Gradient Machine (SGM) models

Stochastic Gradient Boosting (SGB) is similar to Boosted trees but the difference lies in the introduction of randomness in the training process. Indeed, each new tree is trained on a random subsample of the data, rather than the entire dataset. In Stochastic Gradient Boosting, each tree $h_t(x)$ is trained on a random subsample $S_t \subseteq D$, where D is the entire dataset. Using equation (20) as a basis, the SGB model at iteration t is can be formulated as:

$$F_t(x) = F_{t-1}(x) + \eta \cdot h_t(S_t). \tag{22}$$

After all iterations are complete, the final prediction $\hat{y}$ for a given input x is the cumulative output of all the trees in the ensemble:

$$\hat{y} = F_T(x) = F_0(x) + \sum_{t=1}^{T} \eta \cdot h_t(S_t), \tag{23}$$

Where T is the total number of iterations (trees). This cumulative approach leverages each tree to incrementally reduce error by addressing the residuals of the previous trees, resulting in a stronger predictive model for the target variable.

XGBoost, or eXtreme Gradient Boosting, is an optimized version of gradient boosting that incorporates additional regularization. It minimizes a regularized loss function that

includes both a predictive loss (e.g., Mean Squared Error) and a regularization term to penalize model complexity. The objective function in XGBoost is:

$$\text{Obj}(F_t) = \sum_{i=1}^{n} L(y_i, F_{t-1}(x_i) + h_t(x_i)) + \Omega(h_t), \tag{24}$$

Where $L(y_i, \hat{y}_i)$ is the loss function measuring the difference between the actual and predicted values for the i-th observation, h_t is the new tree, and $\Omega(h_t) = \gamma T + \frac{1}{2}\lambda \sum_j w_j^2$ is a regularization term. Here, T is the number of leaves in the tree, w_j represents the weights of each leaf, and γ and λ are regularization parameters that control the complexity of each tree.

XGBoost optimizes the objective function by using a second-order Taylor expansion of the loss function, which incorporates both the gradient (first derivative) and the Hessian (second derivative) to calculate the optimal weights for each leaf. This makes the updates more precise and allows the model to converge faster than standard gradient boosting. Mathematically, for a given loss $L(y, \hat{y})$, the gradients g_t and Hessians h_t at iteration t are:

$$g_t = \frac{\partial L(y, F_{t-1}(x))}{\partial F_{t-1}(x)}, \quad h_t = \frac{\partial^2 L(y, F_{t-1}(x))}{\partial F_{t-1}(x)^2}. \tag{25}$$

After a new tree is built to minimize the sum of the gradients and Hessians, XGBoost uses this tree to update the model's predictions. Each new tree, $h_t(x)$, is added to the cumulative prediction function $F_{t-1}(x)$, which represents the predictions from all previous trees. This addition is scaled by the learning rate, η. The updated prediction function after adding the new tree at iteration is the same as equation (20). The function $F_t(x)$ now incorporates the refined adjustments made by the latest tree to improve accuracy. By adding $h_t(x)$, which has been trained to correct the remaining residual errors (gradients) from previous iterations, the overall model incrementally improves its fit to the target values. XGBoost then repeats this process iteratively, building additional trees that continue to correct residual errors from previous predictions. The final prediction for a given input x after T iterations (trees) is:

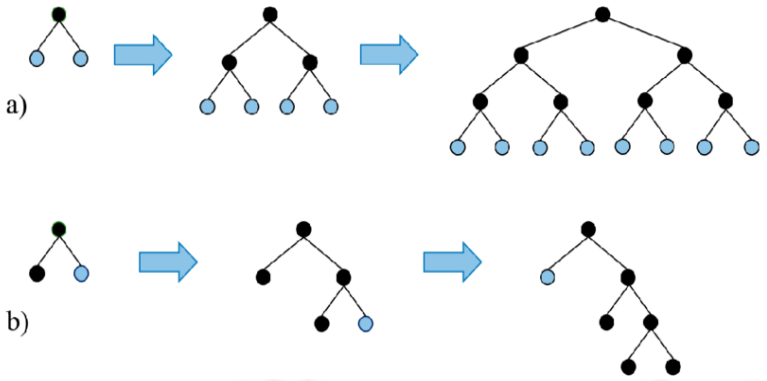
$$\hat{y} = F_T(x) = F_0(x) + \sum_{t=1}^{T} \eta \cdot h_t(x), \tag{26}$$

Where $F_0(x)$ is the initial prediction (the mean of the target values) and the cumulative sum represents the incremental improvements made by each tree. Note that in this paper, XGBoost is ahead other machine learning methods in our horse race.

Finally, LightGBM (Light Gradient Boosting Machine) is an advanced gradient boosting framework. It is particularly optimized for large datasets and high-dimensional data, using

a unique approach called "Leaf-wise growth" which distinguishes it from XGBoost's "Level-wise tree growth". Those two growth patterns are illustrated in Figure 7. In panel (a), we see XGBoost's level-wise tree growth approach in which the tree grows layer by layer (from top to bottom), with all nodes at the same level being split before moving on to the next layer. This level-wise strategy ensures that the tree remains balanced, as all branches grow at the same rate. However, this approach can be less efficient because it splits some nodes that contribute minimally to reducing the overall error. Panel (b) illustrates LightGBM's leaf-wise tree growth approach. LightGBM grows the tree by expanding the leaf node with the highest potential to reduce the error, as determined by a splitting criterion. This method focuses on optimizing the most informative branches first, which results in a more asymmetric, unbalanced tree structure. By growing deeper on the most promising leaves, the model can capture patterns more efficiently, leading to faster convergence. However, leaf-wise growth may also lead to overfitting if the tree grows too deep in certain areas, especially without constraints like maximum depth.

Figure 7: Leaf-wise and Level-wise tree growth



Although their leaf growth strategy is different, the objective function of LightGBM is the same as the one of XGBoost, presented in equation (24). However, the key distinction is how LightGBM builds trees. The Leaf-wise approach involves evaluating each leaf node's potential splits and selecting the one that maximally reduces the objective function, rather than splitting all nodes in parallel.

On a more technical touch, LightGBM also includes a Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB), which improve computational efficiency. GOSS is a sampling method that retains a higher proportion of samples with large gradients (high error) while randomly sampling from smaller gradients, ensuring the model focuses on samples where prediction error is highest. EFB combines mutually exclusive features into a single feature, reducing dimensionality and enhancing processing speed. These optimizations make LightGBM highly efficient for large and sparse datasets.

4.2 The Horse Race

Given the specific characteristics of our database and the empirical objectives of this paper, comparing various Machine Learning techniques is an indispensable practice which provides insights into the efficacy and applicability of different algorithms. This sub-section aims to systematically compare the selected Machine Learning techniques. The necessity behind this comprehensive evaluation is rooted in the diverse nature of these methods, each harboring unique strengths and limitations that may affect the selection accuracy of the variables in our dataset.

4.2.1 The race setup

In order to compare accurately all Machine Learning models, the horse race employs three accuracy metrics : the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE) and R^2. Those metrics have been chosen since they are easily comparable between models and provide the level of accuracy needed to settle the winner of the horse race. This horse race aims at highlighting the best-performing algorithm, that is the one characterized by the lowest RMSE and MAE and the highest R^2.

RMSE measures the average magnitude of prediction errors, with a particular sensitivity to larger errors. It is calculated by taking the square root of the mean of the squared differences between predicted and actual values. The formula for RMSE is :

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{27}$$

where y_i represents the actual value, $\hat{y}_i$ the predicted value, and n the number of observations. Because errors are squared before averaging, RMSE penalizes large errors more than smaller ones, making it particularly useful in contexts where large deviations are undesirable. RMSE is expressed in the same units as the target variable, providing an interpretable measure of the model's average prediction error.

MAE, on the other hand, represents the average of the absolute differences between predicted and actual values, without squaring the errors. It is calculated as :

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|. \tag{28}$$

Unlike RMSE, MAE does not disproportionately penalize large errors, which makes it a more direct indicator of the average prediction error across all observations. This attribute makes MAE useful in situations where all errors, regardless of magnitude, should be treated equally. As with RMSE, MAE is expressed in the same units as the target variable, offering

an easily interpretable measure of the average absolute error in predictions.

R-squared (R^2), quantifies the proportion of variance in the target variable that is explained by the model. It is calculated using the formula :

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{29}$$

where $\bar{y}$ is the mean of the actual values. R^2 ranges from 0 to 1, with 1 indicating that the model perfectly explains the variance in the target variable and 0 indicating that the model does no better than a simple mean prediction. In some cases, R^2 can be negative, implying that the model performs worse than a baseline model that predicts the mean. Unlike RMSE and MAE, which directly measure error magnitudes, R^2 provides insight into the overall explanatory power of the model, making it a useful indicator of model fit in relation to the variability in the data.

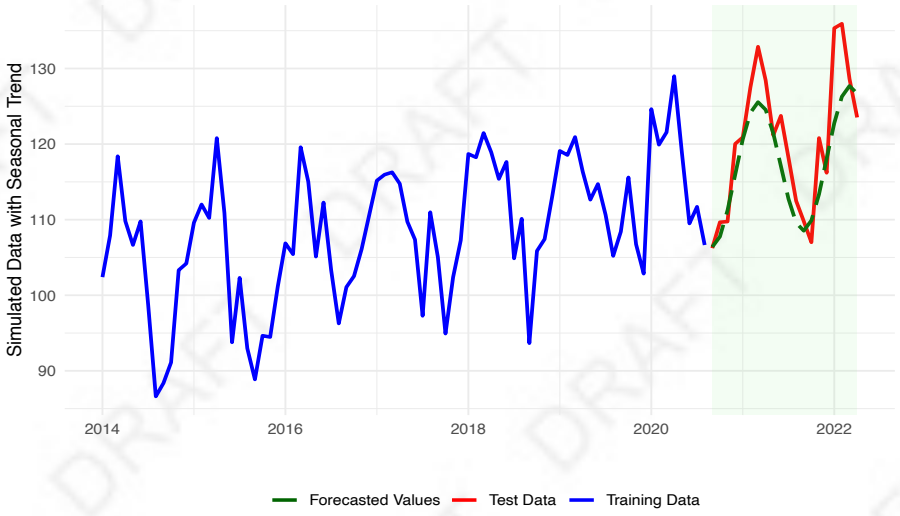
Prior to unleashing the horses on the racetrack, some parameters are chosen to ensure comparability and maximizing each of the techniques performances. First, the predictors are standardized before fitting the models. Standardizing is particularly important for models using penalties on the size of the coefficients, such as LASSO, Ridge or Elastic net models. Standardization also improves the interpretability of coefficients and the ability to rank the coefficient's importance by the relative magnitude of its post-shrinkage coefficient estimate. Second, we use cross-validation both as a method for model selection and parameter tuning as well as a tool for forecast performance evaluation. This method is a widely-used approach in machine learning, that divides the dataset into two distinct subsets: a training set used to fit the model and a test set used to evaluate its performance on unseen data (i.e. realized data points that are hidden from the model). In this paper, cross-validation is applied with 80% of the dataset being used for training and 20% for testing.

For model selection and parameter tuning, this train/test cross-validation approach allows for the optimization of hyperparameters by repeatedly evaluating the model's forecasting performance on the test set across different parameter settings. By systematically adjusting hyperparameters and observing their impact on the test set performance, we can identify configurations that enhance model accuracy and prevent overfitting. This approach is particularly relevant for regularized models, such as LASSO and Ridge regression, as well as for models like neural networks, decision trees and stochastic gradient machines, where precise hyperparameter tuning is essential to balance the model's complexity and generalization. Regularized models benefit from cross-validation by selecting the optimal regularization strength, represented by the parameter λ, which prevents overfitting by penalizing excessive complexity. This is crucial because an overly strong penalty can lead to underfitting, while a weak penalty can result in overfitting. Cross-validation helps find the balance, enabling these models to generalize effectively. In models such as neural networks, decision trees or stochastic gradient machines, cross-validation plays an important role in hyperparameter tuning. For neural networks, which are prone to overfitting due to their non-linear

flexibility, cross-validation assists in setting parameters like the learning rate and the network depth to achieve optimal generalization. Similarly, for decision trees and stochastic gradient machine models, cross-validation aids in tuning parameters such as tree depth and learning rate to optimize model complexity. By using cross-validation to test various configurations, these models can capture underlying trends without overfitting.

In terms of forecast performance evaluation, this train/test cross-validation framework provides a direct assessment of how the model will perform forecasts on unseen data, reflecting its accuracy and reliability. By reporting metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), or R^2, we obtain an overview of the model's accuracy and generalization capability. Figure 8 presents an example of cross-validation using simulated data with an upward seasonal trend on an Elastic Net model (50% $L1$ and 50% $L2$ penalty terms). The model is trained using data represented by the blue line in order to forecast the test data, represented by the red line. The model's forecasted values are represented by the green dashed line. Cross-validation helps not only finding the optimal value for λ (i.e. the strength of the regularization on the penalty terms) but also estimate the forecasting accuracy of the model with RMSE (5.07), MAE (3.87) and R^2 (0.68) metrics.

Figure 8: Cross-validation example with simulated data



Notes : Simulated time series data with a seasonal upward trend. The blue and red lines represents realized data points used to train and test the model, respectively. The green dashed line represents forecasted values. The model used is Elastic Net, with a distribution of 50% L1 and 50% L2 penalty terms.

4.2.2 The race results

In this horse race, each machine learning technique represents a competing horse, with each country serving as an individual race. The race consists of three heats, corresponding to each of the accuracy metrics used: RMSE, MAE, and R^2. This setup results in a total of 30 distinct races for 14 competing models. The overall winner will be the model that performs best across the most races, thereby demonstrating its ability to generate accurate forecasts and its reliability when applied to the countries' diverse datasets. The results for each metric are provided in tables 6 to 8.

1 The RMSE results

Table 6 compiles the Root Mean Square Error (RMSE) values. Across the RMSE table, XGBoost consistently achieves lower RMSE scores relative to other models, winning 70% of the races. In order to understand the magnitude of XGBoost's RMSE value, we compare it to a linear technique (Elastic Net) and to a non-linear technique (Support Vector Regression) since they are usually the techniques displaying the lowest RMSE in their respective categories. In particular, XGBoost achieves the lowest RMSE values in Belgium (0.142), France (0.078), Greece (1.639), Italy (0.277), The Netherlands (0.043), Portugal (0.664) and Spain (0.248). Other techniques win in Austria (SGBoosting), Finland (Regularized trees) and Ireland (LightGBM).

In peripheral countries, other models exhibit higher RMSE values, such as in Italy, where Elastic Net records an RMSE of 0.381, 37% higher than XGBoost, while Support Vector Regression (SVR) reaches 0.474, a 71% increase. Similarly, in Portugal, Elastic Net's RMSE is 55% higher than XGBoost, and SVR's is 89% higher. In Greece, XGBoost achieves an RMSE of 1.639, by comparison, Elastic Net records a significantly higher RMSE of 6.947, which is 324% above XGBoost's level, indicating a substantial gap in performance. Similarly, Support Vector Regression in Greece has an RMSE of 2.575, which is 57% higher than XGBoost, further emphasizing XGBoost's effectiveness in controlling error magnitude in this country. In Spain, XGBoost achieves an RMSE of 0.248, again setting a low benchmark. Here, Elastic Net's RMSE reaches 0.303, a 22% increase over XGBoost, while SVR records an RMSE of 0.325, 31% higher than XGBoost's value.

In core countries such as Belgium, XGBoost achieves an RMSE of 0.142. Comparatively, Elastic Net records an RMSE of 0.219, which is 54% higher than XGBoost. Support Vector Regression (SVR) in Belgium shows an even larger discrepancy with an RMSE of 0.274, representing a 93% increase over XGBoost's benchmark value. In France, XGBoost achieves an RMSE of 0.078, establishing the lowest error model. Elastic Net's RMSE is 0.113, which is 45% higher than XGBoost while SVR records an RMSE of 0.178, an increase of 128% over XGBoost. In the Netherlands, XGBoost reaches an RMSE of 0.043, sharing the lowest error value with another Stochastic Gradient machine technique : SGBoosting. Elastic Net achieves a higher RMSE of 0.059, which is 37% above XGBoost, while SVR records an RMSE of 0.046, showing a 7% increase over XGBoost.

Overall, the RMSE results highlight XGBoost as the most effective model for minimizing prediction errors across the countries under study, consistently outperforming linear models such as LASSO, Ridge and Elastic Net as well as surpassing non-linear models like Support Vector Regression and Feedforward Neural Networks. Although XGBoost emerges as the winning model in the majority of cases, it encounters some competition from other techniques within its own category, such as SG Boosting, LightGBM and Regularized Trees. These models occasionally achieve better results, demonstrating that -while XGBoost is the most reliable overall- other gradient-boosting and tree-based techniques can also be effective.

2 The MAE results

Table 7 presents the results for the Mean Absolute Error (MAE) values. Again, XGBoost consistently establishes itself as the most effective model in the countries under study, winning 80% of the races. Similarly to the analysis of the RMSE results, XGBoost serves here as the benchmark model and its performance is compared against the ones of Boosted LASSO as a representative of linear models and Support Vector Regression (SVR) as a representative of non-linear models, since they are the best performers. Overall, , XGBoost achieves the lowest MAE values in Austria (0.065), Belgium (0.086), Finland (0,051), France (0.057), Greece (0.884), The Netherlands (0.031), Portugal (0.346) and Spain (0.170). The LightGBM technique wins by a slight margin in Ireland (0,153) and Italy (0,203).

In peripheral countries like Greece, XGBoost achieves an MAE of 0.884, setting a relatively low benchmark in a country where MAE values are generally high : above 1 in 12 out of 14 models. In comparison, SVR records a significantly higher MAE of 1. 414, which is 60% above XGBoost's value, indicating a substantial difference in error reduction capability between these two models. Boosted LASSO performs even less effectively, with an MAE of 1. 627, which is 84% higher than XGBoost. In Portugal, XGBoost maintains its position as the most accurate model with an MAE of 0.346. Boosted LASSO records an MAE of 0.430, which is 24% higher than XGBoost, while SVR, with an MAE of 0.674, exhibits a 95% increase over XGBoost's. In Spain, XGBoost achieves an MAE of 0.170. In comparison, Boosted LASSO has an MAE of 0.220, which is 29% higher than XGBoost's value and SVR records an MAE of 0.239, 41% higher than XGBoost.

In core countries, such as Austria, XGBoost achieves an MAE of 0.065 while Boosted LASSO records an MAE of 0.072, approximately 11% higher than XGBoost, indicating that while Boosted LASSO performs reasonably well, it still falls short of XGBoost's precision in minimizing error. SVR in Austria shows an even larger deviation with an MAE of 0.078, which is 20% higher than XGBoost. In Belgium, XGBoost achieves an MAE of 0.086, once again establishing itself as the model with the lowest error. Boosted LASSO records an MAE of 0.108, which is 26% higher than XGBoost. SVR, with an MAE of 0.120, performs even less favorably, showing a 40% increase over XGBoost. In Finland, XGBoost attains an MAE of 0.051 while Boosted LASSO achieves 0.052, which is just 2% higher than XGBoost, indicating that Boosted LASSO performs quite closely to XGBoost in this particular country. Nevertheless, XGBoost still maintains a slight edge in minimiz-

ing errors. SVR, on the other hand, records an MAE of 0.061, which is 20% higher than XGBoost's value. In France, XGBoost records an MAE of 0.057, maintaining its position as the most accurate model. Boosted LASSO, by contrast, shows an MAE of 0.065, which is 14% higher than XGBoost. SVR records a considerably higher MAE of 0.101 in France, which is 77% above XGBoost's value. In the Netherlands, XGBoost achieves an MAE of 0.031. Boosted LASSO records an MAE of 0.051, which is 65% higher than XGBoost while SVR, with an MAE of 0.033, performs more closely to XGBoost, showing only a 6% increase.

In summary, XGBoost consistently achieves the lowest MAE values across Austria, Belgium, Finland, France, and the Netherlands, demonstrating superior predictive accuracy relative to both linear and non-linear models. While Boosted LASSO occasionally performs closely to XGBoost, particularly in Finland, it generally exhibits higher error rates. SVR, although competitive in some countries, particularly in the Netherlands, also consistently falls short of XGBoost's level of precision. The most competitive model comes again from the same category as XGBoost : LightGBM. The latter wins in Ireland and Italy, providing evidence that gradient-boosting models are more adequate than linear and non-linear ones.

3 The R^2 results

Finally, table 8 presents the results for the R^2 values. For a third time, XGBoost performs particularly well, winning 60% of the races. Again, XGBoost serves as the benchmark and is compared to Boosted LASSO and Support Vector Regression, both of which are the top performers within their respective categories. Overall, XGBoost claims the highest R^2 values in Belgium (0.916), France (0.896), Italy (0.912), the Netherlands (0.804), Portugal (0.947) and Spain (0.929). Other rule-based techniques contribute to additional wins, with Regularized Trees performing best in Finland (0.68), SG Boosting in Austria (0.857), and LightGBM in Ireland (0.988). For the first time, a linear technique wins a race; Boosted LASSO in Greece (0.937).

In peripheral countries, like Greece, XGBoost must concede its first defeat in a race against a linear technique. Indeed, XGBoost achieves an R^2 value of 0.937, while Boosted LASSO records an R^2 of 0.938. This race is particularly close, indicating that XGBoost performs nearly on par with Boosted LASSO in Greece. However, SVR shows a substantially lower R^2 of 0.845, representing a 10% reduction in explained variance compared to XGBoost. In Italy, XGBoost attains an R^2 of 0.912 while Boosted LASSO records an R^2 of 0.890, only 2% lower than XGBoost, showing that Boosted LASSO remains a strong linear alternative. SVR, on the other hand, achieves an R^2 of 0.744, which is 18% lower than XGBoost, indicating a notable gap in explanatory power. In Portugal, XGBoost records an R^2 of 0.947, the highest among all models but Boosted LASSO's R^2 of 0.940 is only 1% lower. However, SVR's R^2 in Portugal is 0.810, which is 14% below XGBoost's. In Spain, XGBoost achieves an R^2 of 0.929, setting a high benchmark for model performance. Boosted LASSO records an R^2 of 0.867, which is 7% lower than XGBoost while SVR, with an R^2 of 0.879, performs somewhat better than Boosted LASSO but still falls short to XGBoost by around 5%.

In core countries, such as Belgium, the R^2 value for XGBoost is 0.916, setting a high benchmark for model effectiveness. By comparison, Boosted LASSO records an R^2 of 0.812, which is 11% lower than XGBoost. This discrepancy suggests that Boosted LASSO, while performing reasonably well, lacks the capacity to explain the variance as effectively as XGBoost in the Belgian dataset. SVR, with an R^2 of 0.691, performs even less favorably, showing an R^2 value that is 25% lower than XGBoost. In France, XGBoost achieves an R^2 value of 0.896, again establishing itself as the most effective model in terms of variance explanation. Boosted LASSO, by contrast, records an R^2 of 0.847, which is 5% lower than XGBoost, indicating that although Boosted LASSO performs relatively well, it does not match XGBoost's predictive capacity. SVR records a significantly lower R^2 of 0.466 in France, which is 48% below XGBoost's value. In the Netherlands, XGBoost attains an R^2 of 0.804 while Boosted LASSO records an R^2 of 0.547, which is 32% lower, reflecting a considerable gap in predictive performance. On the other hand, SVR records an R^2 of 0.780, which is slightly lower than XGBoost by 3%, indicating that SVR performs competitively in this particular country.

The R^2 analysis demonstrates XGBoost's superiority in capturing data variance, establishing it as the most effective model. Consistently achieving the highest R^2 values, XGBoost outperforms both Boosted LASSO and Support Vector Regression (SVR). In countries like Belgium, Greece, Ireland, Italy, Portugal and Spain, XGBoost explains a larger proportion of the variance, reflecting its capacity for accurate predictions. While Boosted LASSO occasionally performs competitively, particularly in Greece, Ireland and Portugal, it generally records slightly lower R^2 values, revealing limitations in its explanatory power relative to XGBoost. SVR, despite being the best non-linear alternative, exhibits more significant deficits in R^2, particularly in Belgium, Finland and France.

In conclusion, synthesizing the insights from RMSE, MAE, and R^2 metrics make XGBoost emerge as the most accurate and reliable model by winning 70% of the races (21 out of 30). It consistently demonstrated low error rates (both RMSE and MAE) and high explanatory power (R^2) across the countries' datasets. Although other models, such as SG Boosting or LightGBM, performed comparably or even better in certain countries, XGBoost's stability and performance across all three metrics make it the definitive winner in this horse race. In light of those results, XGBoost's consistent ability to balance error minimization with high explanatory variance suggests this model would be the preferred choice for predicting Euro area sovereign bond spreads. Hence, we choose XGBoost as our model for the remainder of the paper.

Table 6: Horse Race - Root Mean Square Error (RMSE)

Machine Learning Techniques	**Country**									
Linear	*Austria*	*Belgium*	*Finland*	*France*	*Greece*	*Ireland*	*Italy*	*Netherlands*	*Portugal*	*Spain*
Standard LASSO	0.094	0.212	0.083	0.113	6.082	0.393	0.376	0.060	1.002	0.303
Boosted LASSO	0.105	0.214	0.076	0.095	1.627	0.374	0.311	0.065	0.707	0.340
Group-LASSO	0.094	0.332	0.075	0.840	3.738	0.477	0.313	0.091	1.822	0.309
SG-LASSO	0.095	0.278	0.073	0.141	2.996	0.477	0.455	0.082	1.868	0.307
Ridge	0.165	0.290	0.088	0.165	3.176	0.782	0.663	0.062	1.820	0.493
Elastic Net	0.102	0.219	0.084	0.113	6.947	0.407	0.381	0.059	1.030	0.303
Non-linear										
Support Vector Regression (SVR)	0.117	0.274	0.076	0.178	2.575	0.392	0.474	0.046	1.256	0.325
Feedforward Neural Network (FNN)	0.199	0.278	0.121	0.141	2.996	0.477	0.943	0.082	1.868	0.922
Rule-based										
Trees										
Decision Trees	0.209	0.212	0.097	0.127	2.271	0.292	0.336	0.059	1.162	0.639
Regularized Trees	0.092	0.184	**0.069**	0.090	1.909	0.333	0.304	0.044	0.816	0.246
Boosted Trees	0.095	0.185	0.073	0.104	1.813	0.298	0.322	0.046	0.718	0.326
Stochastic Gradient Machines (SGM)										
SG Boosting	**0.089**	0.196	0.073	0.098	1.797	0.288	0.299	0.043	0.833	0.310
XGBoost	0.097	**0.142**	0.069	**0.078**	**1.639**	0.337	**0.277**	**0.043**	**0.664**	**0.248**
LightGBM	0.100	0.211	0.076	0.097	2.185	**0.220**	0.279	0.046	0.941	0.295

Table 7: Horse Race - Mean Absolute Error (MAE)

Machine Learning Techniques					**Country**					
Linear	*Austria*	*Belgium*	*Finland*	*France*	*Greece*	*Ireland*	*Italy*	*Netherlands*	*Portugal*	*Spain*
Standard LASSO	0.071	0.116	0.066	0.075	2.180	0.233	0.283	0.045	0.613	0.244
Boosted LASSO	0.072	0.108	0.052	0.065	0.896	0.182	0.217	0.051	0.430	0.220
Group-LASSO	0.072	0.204	0.059	0.401	1.736	0.282	0.244	0.075	1.129	0.216
SG-LASSO	0.071	0.177	0.057	0.108	1.594	0.282	0.328	0.068	1.150	0.214
Ridge	0.109	0.164	0.075	0.098	1.835	0.567	0.502	0.046	1.115	0.357
Elastic Net	0.073	0.122	0.065	0.076	2.359	0.250	0.287	0.044	0.641	0.244
Non-linear										
Standard SVR	0.078	0.120	0.061	0.101	1.414	0.262	0.336	0.033	0.674	0.239
Feedforward NN	0.128	0.177	0.095	0.108	1.594	0.282	0.696	0.068	1.150	0.672
Rule-based										
Trees										
Decision Trees	0.118	0.128	0.072	0.087	1.441	0.206	0.247	0.041	0.737	0.379
Regularized Trees	0.064	0.094	0.053	0.063	1.256	0.156	0.211	0.032	0.419	0.164
Boosted Trees	0.071	0.097	0.057	0.075	1.134	0.172	0.235	0.034	0.407	0.223
Stochastic Gradient Machines (SGM)										
SG Boosting	0.068	0.091	0.055	0.070	1.037	0.160	0.213	0.031	0.421	0.211
XGBoost	**0.065**	**0.086**	**0.051**	**0.057**	**0.884**	0.154	0.205	**0.031**	**0.346**	**0.170**
LightGBM	0.074	0.109	0.058	0.073	1.303	**0.153**	**0.203**	0.032	0.467	0.218

Table 8: Machine Learning Techniques - R-Squared (R^2)

Machine Learning Techniques	**Country**									
Linear	*Austria*	*Belgium*	*Finland*	*France*	*Greece*	*Ireland*	*Italy*	*Netherlands*	*Portugal*	*Spain*
Standard LASSO	0.840	0.815	0.534	0.783	0.135	0.962	0.839	0.625	0.879	0.895
Boosted LASSO	0.798	0.812	0.603	0.847	**0.938**	0.966	0.890	0.547	0.940	0.867
Group-LASSO	0.838	0.548	0.615	-10.880	0.673	0.944	0.888	0.127	0.600	0.890
SG-LASSO	0.838	0.682	0.639	0.663	0.790	0.944	0.764	0.285	0.580	0.892
Ridge	0.503	0.653	0.470	0.541	0.764	0.849	0.499	0.589	0.602	0.720
Elastic Net	0.813	0.802	0.517	0.784	-0.128	0.959	0.834	0.636	0.872	0.895
Non-linear										
Standard SVR	0.753	0.691	0.606	0.466	0.845	0.962	0.744	0.780	0.810	0.879
Feedforward NN	0.284	0.682	-0.004	0.663	0.790	0.944	-0.014	0.285	0.580	0.022
Rule-based										
Trees										
Decision Trees	0.210	0.814	0.354	0.727	0.879	0.979	0.871	0.636	0.838	0.531
Regularized Trees	0.847	0.860	**0.680**	0.864	0.915	0.973	0.895	0.800	0.920	0.930
Boosted Trees	0.836	0.858	0.637	0.818	0.923	0.978	0.882	0.777	0.938	0.877
Stochastic Gradient Machines (SGM)										
SG Boosting	**0.857**	0.842	0.639	0.839	0.924	0.980	0.898	0.803	0.916	0.889
XGBoost	0.827	**0.916**	0.674	**0.896**	0.937	0.971	**0.912**	**0.804**	**0.947**	**0.929**
LightGBM	0.818	0.817	0.606	0.840	0.888	**0.988**	0.911	0.780	0.893	0.900

5 Empirical Results

This empirical section builds on the results of the horse race, where XGBoost emerged as the most accurate and reliable model for forecasting Euro area bond spreads. The horse race not only demonstrated XGBoost's superiority over 13 competing models, including LASSO models, Support Vector Regression and other rule-based techniques, but also showed its reliability in capturing the relationships among the hundreds of variables supplied in our novel big dataset. Leveraging this predictive power, we first conduct a country-specific empirical analysis offering a detailed examination of the predicted yield spreads and their implications for financial fragmentation within the Euro Area. Then, we provide an evaluation of bond spread dynamics by creating a correlation heatmap with a dissimilarity measure which sheds light on the persistent core-periphery divide in the Euro area. Through this methodology, the empirical section not only confirms XGBoost's capabilities but also provides insights into the financial risks and market trends shaping the future of the European Monetary Union.

5.1 The country-specific analysis

For the ten Euro area countries included in this analysis, we use XGBoost models to predict the 10-year sovereign bond yield spread six months into the future. Figures 9 to 18 illustrate the time series of observed and predicted spread values for each country. Overall, the predictions align with the initial hypothesis, underscoring the heterogeneous dynamics of sovereign bond yield spreads across the Euro area. Specifically, Austria, Finland and the Netherlands are expected to have significantly lower yield spreads while Italy and Greece are predicted to face notably higher spreads. Meanwhile, the outlook for France, Belgium, Spain, Ireland and Portugal is projected to be relatively stable, although their predicted yield spreads slightly exceed the most recently observed data points.

The observed yield spreads for Austria, Finland and the Netherlands (figures 9, 10, and 11, respectively) display a notable increase at the start of 2022. This rise can be attributed to the heightened financial and economic uncertainty stemming from the war in Ukraine, coupled with the energy crisis that exacerbated inflationary pressures and prompted the European Central Bank (ECB) to adopt a tighter monetary policy stance. Following this initial spike, the spreads gradually stabilized over the course of 2022 and 2023. The six-month-ahead predictions for these countries reveal a distinct and pronounced downward trend in yield spreads. This outcome aligns with the expectation that Austria, Finland and the Netherlands, as core Euro area countries, benefit from "flight-to-quality" behavior. During periods of heightened uncertainty, investors tend to shift toward safer assets. This "flight-to-quality" behavior leads to an increased demand for assets associated with core countries, which are perceived as more financially stable and thus less risky. Consequently, this shift in the investors' preferences results in a contraction of yield spreads for the core economies, reflecting their status as financial safe havens during times of economic and geopolitical instability.

On the other hand, for Italy and Greece (figures 17 and 18), the actual bond spreads show significant volatility during the early 2010s, peaking during the European Sovereign Debt Crisis, followed by a gradual decline in subsequent years. The narrower spreads can be attributed to measures such as the ECB's introduction of unconventional monetary policies, including the Outright Monetary Transactions (OMT) program which played a critical role in reassuring investors of the Euro area's commitment to safeguarding financial stability. The establishment of the European Stability Mechanism (ESM) and fiscal consolidation efforts also contributed to mitigating sovereign risk perceptions. Although those developments reduced investor uncertainty, fostering a period of relative market normalization and lower bond spreads, the recent period exhibits renewed volatility and higher spreads. The XGBoost model predicts a notable upward movement in spreads over the six-month horizon, indicating revived concerns about Greece and Italy's fiscal sustainability and economic soundness. This predicted rise in spreads hence aligns with the broader pattern of "contagion" behavior observed among peripheral Euro area countries. Contagion refers to the phenomenon where economic or financial instability in one country spreads to others. This behavior is particularly relevant within the Euro Area, where shared monetary policies and deeply integrated financial systems create channels through which shocks in one country can quickly influence others. In the case of Greece and Italy, both nations have historically exhibited vulnerabilities that make them susceptible to such contagion effects. Greece's high public debt levels, structural deficits and history of fiscal crises amplify its exposure to shifts in investor confidence. Similarly, Italy's significant public debt and prolonged economic stagnation leave it vulnerable to risk repricing in global markets. When investors perceive heightened risks in one peripheral country, such as Greece, the fear of similar vulnerabilities in Italy can trigger a chain reaction. This results in rising risk premiums across peripheral countries. For Greece and Italy, this "contagion" behavior amplifies the divergence in spreads compared to core Euro area countries, further reinforcing financial fragmentation. Furthermore, the magnitude of the predicted increase in yield spreads needs to be considered, particularly in the case of Greece. The Greek yield spreads are predicted to reach levels not observed in the past five years, a development that raises significant concerns. Indeed, the combination of higher bond yields with structural domestic deficit imbalances leads to increased pressure on Greek finances, heightening the risk of a new European Sovereign Debt crisis.

Finally, the five remainder countries –France, Belgium, Portugal, Ireland and Spain- represent a diverse spectrum of the Euro area's financial structure, with France and Belgium usually considered as core countries while Portugal, Ireland and Spain often categorized as peripheral countries. In France and Belgium (figures 12 and 13), the observed spreads display relative stability following the resolution of the ESD crisis, maintaining a narrow range with slight fluctuations. On the contrary, Spain, Ireland and Portugal (figures 14, 15 and 16, respectively) experienced significant financial pressures during the crisis, with spreads peaking sharply before decreasing and stabilizing in the years that followed the ESD crisis, benefiting from the same risk-mitigating measures as Italy and Greece. However, the recent period of economic uncertainty and geopolitical instability rekindled the financial risk associated to each of the countries, highlighted by an increase in yield spreads in 2022

and then stabilized. The XGBoost models for each country project a continuation of this stable trend, with predicted yield spreads showing minimal variation over the six-month horizon. The results suggest that those countries benefits from an intermediary position within the Euro area, attracting moderate investor confidence while avoiding the extremes spread movements experienced by core or peripheral countries. Ireland's economic recovery, underpinned by robust fiscal consolidation as well as Portugal and Spain's reduced fiscal vulnerabilities have contributed to their relative stable spread levels. Although they remain sensitive to external investor sentiment due to their classification as peripheral countries, they have transitioned to a more stable position. More worryingly, the fiscal and economic situation in France and Belgium seems to have deteriorated, moving both countries closer to this intermediate status, straddling the core-periphery divide. While still benefiting from lower bond yields compared to peripheral countries, they also display residual investor concerns, leading to flatter predicted spread trends relative to the downward trends experienced by core countries.

Figure 9: Austria

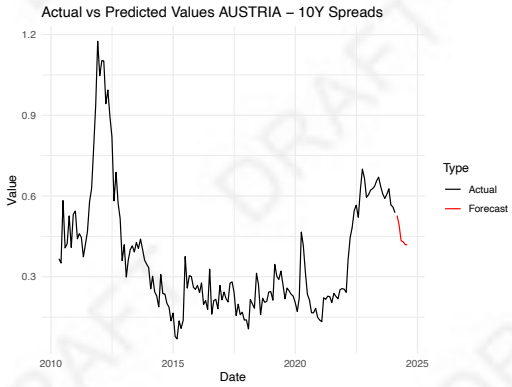


Figure 10: Finland

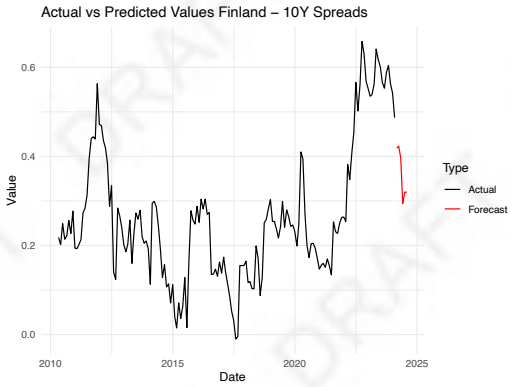


Figure 11: The Netherlands

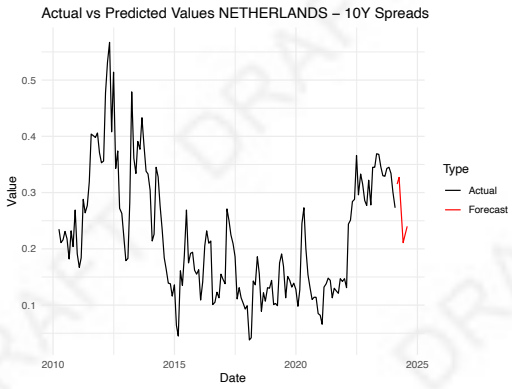


Figure 12: France

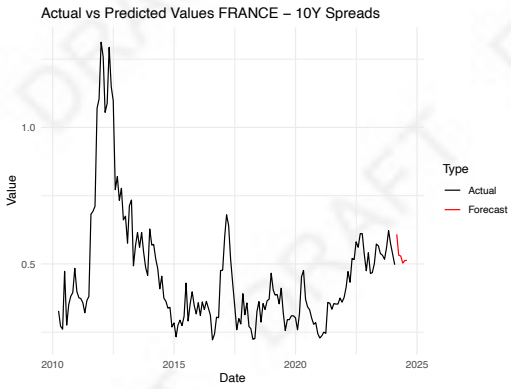


Figure 13: Belgium

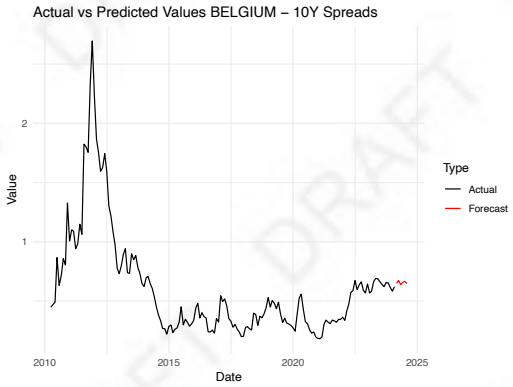


Figure 14: Spain

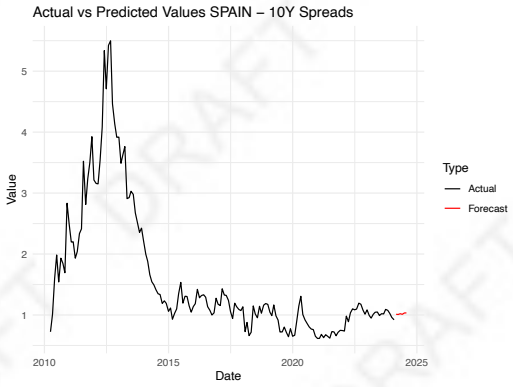


Figure 15: Ireland

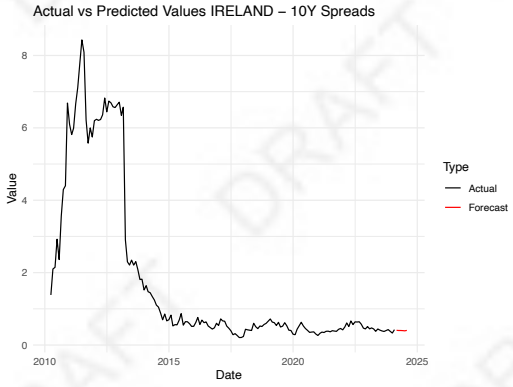


Figure 16: Portugal

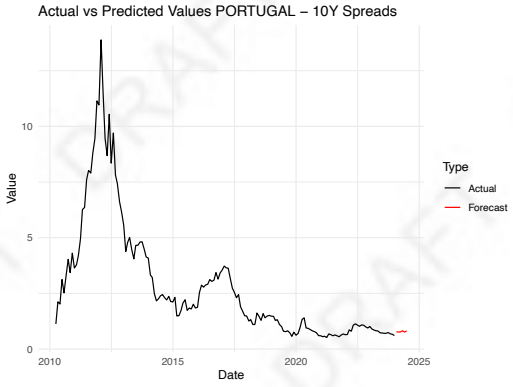


Figure 17: Italy

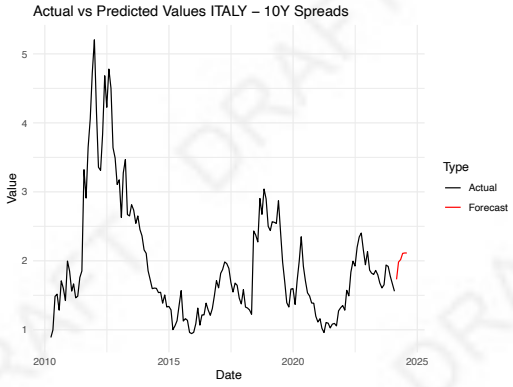
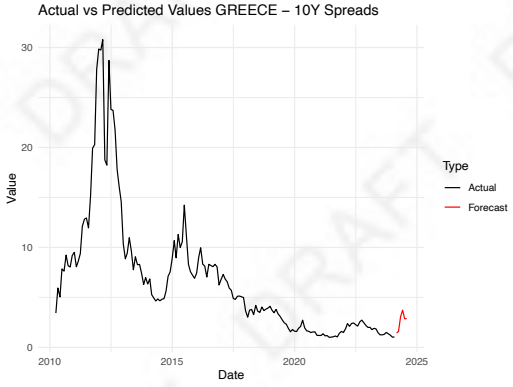


Figure 18: Greece



5.2 The Euro area's yield spread dynamic

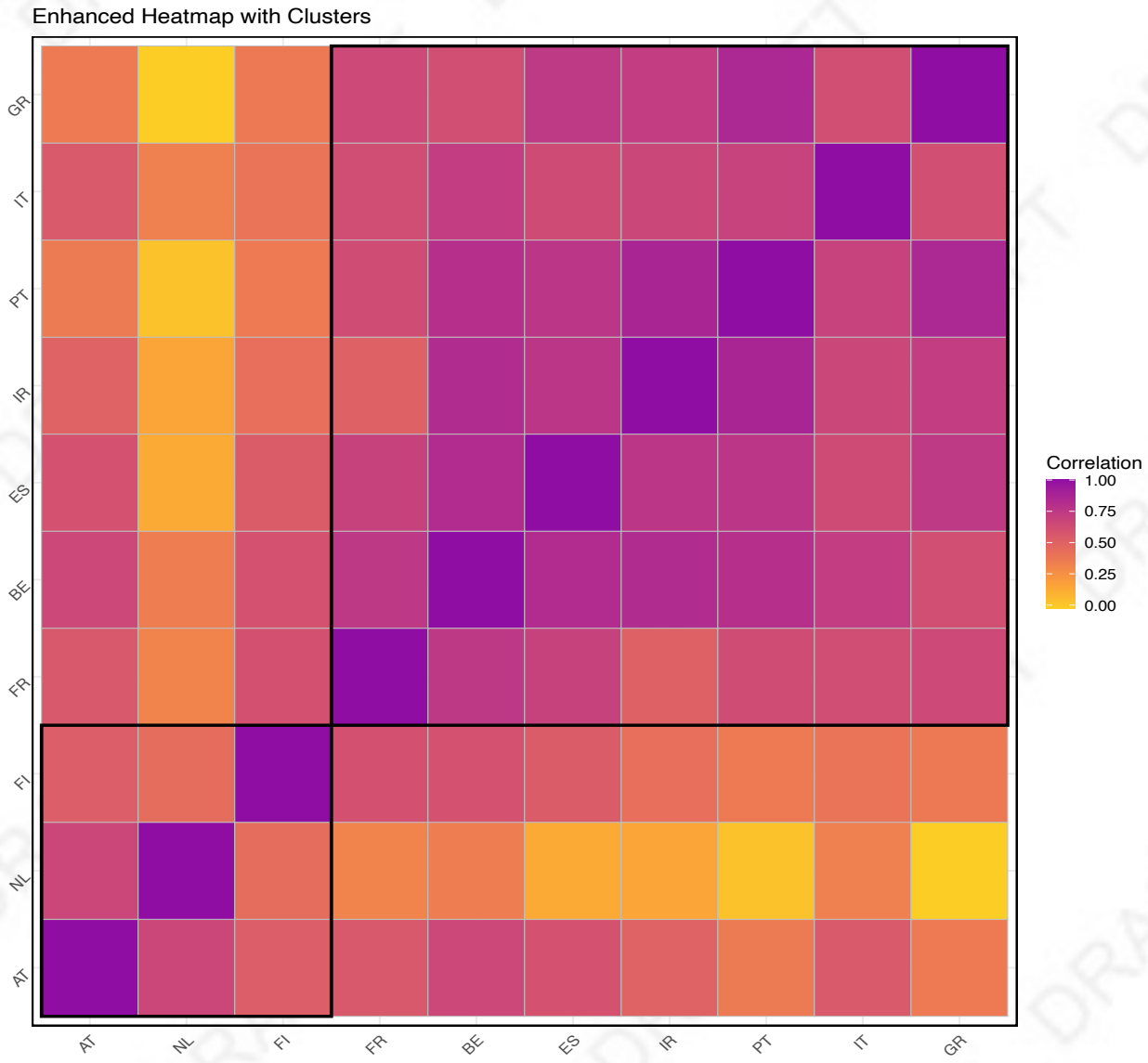
The correlation heatmap presented in figure 19 visualizes the pairwise relationships between the predicted 10-year bond spreads of the countries under study, structured into two distinct clusters. The heatmap is based on the dissimilarity measure $1-\rho_{i,j}$, where $\rho_{i,j}$ represents the pairwise average correlation coefficient between the predictions for countries i and j. Higher correlation values (closer to 1) indicate stronger similarities in predicted spread dynamics, while lower values reflect greater divergence. The two clusters are demarcated by rectangles, emphasizing the dichotomy between the peripheral and core countries within the Euro Area.

The first cluster, predominantly encompassing Greece, Italy, Portugal, Ireland and Spain, aligns with the group of usual peripheral countries in the literature. This grouping reflects high levels of interdependence in their predicted spread dynamics, suggesting that these countries are subject to similar investor behavior patterns. These patterns stem from shared vulnerabilities, such as higher debt levels, fiscal deficits and susceptibility to contagion effects during periods of financial instability.

The second cluster, consisting of Austria, Finland and the Netherlands, represents the core countries within the Euro area. These nations exhibit high intracluster correlations in their predicted spreads, reflecting their shared status as safe-haven economies benefiting from "flight-to-quality" behavior. The low correlation between this cluster and the peripheral countries highlights the financial fragmentation risk within the Euro area, underscoring the persistent divergence in investor perceptions between these two groups of countries.

Interestingly and for the first time to our knowledge, France and Belgium have switched sides and joined the peripheral countries, reflecting their struggling economic and financial situations. Considered as core countries in the literature because of their strong fiscal positions, stable economic growth and high investor confidence, they benefited from low bond yields and acted as anchors of stability within the area, aligning with other core countries. In contrast, peripheral countries often face higher debt burdens, greater fiscal vulnerabilities and increased susceptibility to market pressures such as "contagion", leading to higher risk premiums and borrowing costs. Thus, the reclassification of Belgium and France as peripheral countries would signal a significant and worrisome shift in the region's economic and financial landscape and could exacerbate the financial fragmentation risk within the Euro Area. Indeed, as a large, systemically important country, France's relegation to the periphery would strain the Euro area's cohesion, raising concerns about the effectiveness of shared monetary policies in stabilizing the economic bloc. Similarly, Belgium's transition would further reduce the pool of core countries, leaving fewer nations capable of absorbing market shocks and reinforcing investor confidence. This would heighten vulnerabilities across the Euro Area, increasing the risk of contagion and threatening the stability of the entire monetary union.

Figure 19: Correlation Heatmap with 2 Clusters



6 Conclusion

This paper proposes a novel high-dimensional dataset covering a wide range of fields within 10 European countries over 2007-2024 to forecast the sovereign yield spread in the Euro area. It proposes to compare the different ML techniques to provide the most accurate forecasts and shows that XGBoost presents the best predictive ability among the ML techniques. Forecasts suggest that the predicted long-term yield spreads of peripheral countries rise while those in core countries' rise remain contained or even decrease. It therefore suggest a increase in the fragmentation risk in the Euro Area.

The implications of the findings of this paper are manifold. The consequences induced by the financial fragmentation risk go well beyond financial markets and impact the real economy. Elevated borrowing costs for peripheral countries can curtail public and private investment, potentially leading to subdued growth trajectories and exacerbated unemployment levels, exacerbating financial fragmentation. It also endangered the strategy toward sustainable public debt in peripheric countries. The stability of the Euro Area necessarily requires financial fragmentation. Discourses surrounding the completion of the banking union, regulatory harmonization, public deficit limitations and potential strides towards a fiscal union are key elements to reduce the risk of fragmentation.

From a policy perspective, the evidence of financial fragmentation risk necessitates a re-evaluation of the existing European financial frameworks and mechanisms, especially in the context of cross-country cooperation and support. Governments in peripheral countries need to be acutely aware of their nation's vulnerabilities and should consider implementing measures to stabilize their public deficit and to bolster their financial systems. Conversely, for core countries, the results serve as a reminder of the interconnected nature of global finance. While they might exhibit resilience in the face of shocks, the repercussions in peripheral nations could eventually ripple back, affecting not only their stability but all of the Euro Area's.

Although the effects of the ECB's policies to bridge the chasm between core and peripheral nations has been a success in the past decade, the ECB must not overlook this financial fragmentation risk, especially in the current context of high inflation, high amount of government debt and worsening public deficits.

References

António Afonso, Pedro Gomes, and Philipp Rother. Short- and long-run determinants of sovereign debt credit ratings. *International Journal of Finance & Economics*, 16(1):1–15, 2011. doi: 10.1002/ijfe.416.

António Afonso, Davide Furceri, and Pedro Gomes. Sovereign credit ratings and financial markets linkages: Application to european data. Working Paper Series 1347, European Central Bank, 2012. URL https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1347.pdf. Working Paper Series No. 1347, June 2011.

Joshua Aizenman, Michael Hutchison, and Yothin Jinjarak. What is the risk of european sovereign debt defaults? fiscal space, cds spreads and market pricing of risk. *Journal of International Money and Finance*, 34:37–59, 2013. doi: 10.1016/j.jimonfin.2012.11.011. URL https://doi.org/10.1016/j.jimonfin.2012.11.011.

Carlo Altavilla, Giacomo Carboni, and Roberto Motto. Asset purchase programmes and financial markets: Lessons from the euro area. Working Paper Series 1864, European Central Bank, 2015. URL https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1864.en.pdf. Working Paper Series No. 1864, November 2015.

Sakai Ando, Giovanni Dell'Ariccia, Pierre-Olivier Gourinchas, Guido Lorenzoni, Adrian Peralta-Alva, and Francisco Roch. Debt mutualization in the euro area: A quantitative exploration. Technical Report WP/23/59, International Monetary Fund, 2023. URL https://www.imf.org/en/Publications/WP/Issues/2023/03/17/Debt-Mutualization-in-the-Euro-Area-A-Quantitative-Exploration-530384.

Veni Arakelian, Petros Dellaportas, Roberto Savona, and Marika Vezzoli. Sovereign risk zones in europe during and after the debt crisis. *Social Science Research Network (SSRN)*, 2019. URL https://ssrn.com/abstract=3217047.

Manfred Arghyrou and Björn Kontonikas. Rating agencies, self-fulfilling prophecies and multiple equilibria: An empirical model of the european sovereign debt crisis. Discussion Paper 2012-15, University of St. Gallen, School of Economics and Political Science, 2012. URL http://www.fgn.unisg.ch.

Pierluigi Balduzzi, Roberto Savona, and Lucia Alessi. Anatomy of a sovereign debt crisis: Machine learning, real-time macro fundamentals, and cds spreads. *Social Science Research Network (SSRN)*, 2022. URL https://ssrn.com/abstract=3548727.

Alessandro Beber, Michael W. Brandt, and Kenneth A. Kavajecz. Flight-to-quality or flight-to-liquidity? evidence from the euro-area bond market. *The Review of Financial Studies*, 22(3):925–957, 2009. doi: 10.1093/rfs/hhm088. URL https://doi.org/10.1093/rfs/hhm088.

Richard Bellman. *Dynamic programming*, volume 153. American Association for the Advancement of Science, 1966.

Guillaume Belly, Lukas Boeckelmann, Carlos Mateo Caicedo Graciano, Alberto Di Iorio, Klodiana Istrefi, Vasileios Siakoulis, and Arthur Stalla-Bourdillon. Forecasting sovereign risk in the euro area via machine learning. *Social Science Research Network (SSRN)*, 2022. URL https://ssrn.com/abstract=3974515.

Kerstin Bernoth, Jürgen von Hagen, and Ludger Schuknecht. Sovereign risk premia in the european government bond market. Working Paper Series 369, European Central Bank, 2004. URL https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp369.pdf.

Bertrand Candelon, Angelo Luisi, and Francesco Roccazzella. Fragmentation in the european monetary union: Is it really over? *Journal of International Money and Finance*, 122: 102545, 2022. doi: 10.1016/j.jimonfin.2021.102545. URL https://doi.org/10.1016/j.jimonfin.2021.102545.

Marco Castellani and Emanuel Augusto dos Santos. Forecasting long-term government bond yields: An application of statistical and ai models. Technical report, Faculdade Ciências e Tecnologia, Universidade Nova Lisboa, 2006. URL http://centria.fcsh.unl.pt.

Lorenzo Codogno, Carlo Favero, and Alessandro Missale. Yield spreads on emu government bonds. *Economic Policy*, 18(37):503–532, 2003. doi: 10.1111/1468-0327.00114_1. URL https://doi.org/10.1111/1468-0327.00114_1.

Giancarlo Corsetti and Luca Dedola. The mystery of the printing press: Monetary policy and self-fulfilling debt crises. Technical Report 2016/035, ADEMU Working Paper Series, 2016. URL http://www.ademu-project.eu/publications/working-papers.

Michele Costola and Matteo Iacopini. Measuring sovereign bond fragmentation in the eurozone. *Finance Research Letters*, 51:103354, 2023. doi: 10.1016/j.frl.2022.103354. URL https://doi.org/10.1016/j.frl.2022.103354.

Paul De Grauwe and Yuemei Ji. Self-fulfilling crises in the eurozone: An empirical test. *Journal of International Money and Finance*, 34:15–36, 2013. doi: 10.1016/j.jimonfin.2012.11.003. URL https://doi.org/10.1016/j.jimonfin.2012.11.003.

Roberto A. De Santis. A measure of redenomination risk. Working Paper Series 1785, European Central Bank, 2015. URL https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1785.en.pdf.

Carlo Favero and Alessandro Missale. Sovereign spreads in the euro area: Which prospects for a eurobond? *Economic Policy*, 27(70):231–273, 2012. doi: 10.1111/j.1468-0327.2012.00286.x. URL https://doi.org/10.1111/j.1468-0327.2012.00286.x.

Jeremy Fouliard, Michael Howell, Hélène Rey, and Vania Stavrakeva. Answering the queen: Machine learning and financial crises. Working Paper Series 28302, National Bureau of Economic Research, 2021. URL https://www.nber.org/papers/w28302.

Silvia Gabrieli and Claire Labonne. Bad sovereign or bad balance sheets? euro interbank market fragmentation and monetary policy, 2011-2015. Working Paper Series 687, Banque de France, 2018. URL https://publications.banque-france.fr/en.

Niels Gilbert, Jeroen Hessel, and Silvie Verkaart. Towards a stable monetary union: What role for eurobonds? Working Paper Series 379, De Nederlandsche Bank, 2013. URL https://www.dnb.nl/en/publications/.

Pietro Grandi. Sovereign stress and heterogeneous monetary transmission to bank lending in the euro area. *European Economic Review*, 119:251–273, 2019. doi: 10.1016/j.euroecorev.2019.07.011. URL https://doi.org/10.1016/j.euroecorev.2019.07.011.

Shihao Gu, Bryan Kelly, and Dacheng Xiu. Empirical asset pricing via machine learning. Working Paper Series 25398, National Bureau of Economic Research, 2020. URL https://www.nber.org/papers/w25398.

Won Joong Kim, Gunho Jung, and Sun-Yong Choi. Forecasting cds term structure based on nelson–siegel model and machine learning. *Complexity*, 2020:1–23, 2020. doi: 10.1155/2020/2518283. URL https://doi.org/10.1155/2020/2518283.

Simone Manganelli and Guido Wolswijk. Market discipline, financial integration and fiscal rules: What drives spreads in the euro area government bond market? Technical Report 745, European Central Bank, 2007. URL https://ssrn.com/abstract=978373.

Norbert Metiu. Sovereign risk contagion in the eurozone. *Economics Letters*, 117(1):35–38, 2012. doi: 10.1016/j.econlet.2012.04.074. URL https://doi.org/10.1016/j.econlet.2012.04.074.

Sebastian Missio and Sebastian Watzka. Financial contagion and the european debt crisis. Technical Report 3554, CESifo Working Paper Series, 2011. URL https://ssrn.com/abstract=1920642.

Troy J. Strader, John J. Rozycki, Thomas H. Root, and Yu-Hsiang John Huang. Machine learning stock market prediction studies: Review and research directions. *Journal of International Technology and Information Management*, 28(4):63–80, 2020. doi: 10.58729/1941-6679.1435. URL https://scholarworks.lib.csusb.edu/jitim/vol28/iss4/3.

Andrea Zaghini. Fragmentation and heterogeneity in the euro-area corporate bond market: Back to normal? *Journal of Financial Stability*, 23:51–61, 2016. doi: 10.1016/j.jfs.2016.01.009. URL https://doi.org/10.1016/j.jfs.2016.01.009.