Sovereign credit default swaps and macroeconomic fundamentals

Daniele Bianchi* Adam Shuaib†

First draft: 1st May 2021 This draft: February 14, 2022

Abstract

We provide evidence in favour of a significant non-linear, time-varying dependence between sovereign credit default swap (CDS) spreads and macroeconomic fundamentals for OECD countries. Macroeconomic conditions alone explain more than 80% of the out-of-sample variation in CDS spreads when non-linear random forest regressions are used. This is almost double the predictive performance of sparse and dense linear statistical learning methods. We test the coherence of the expected CDS spreads across different pure out-of-sample scenarios, e.g. training a random forest regression on EU countries and predicting CDS spreads for non-EU economies. The results suggest that non-linear machine learning methods enable "shadow" sovereign CDS pricing for countries and historical periods in which tradable sovereign CDS contracts are not available.

Keywords: Sovereign default risk, Credit default swap, Machine learning, Non-linearity, Macroeconomic factors

JEL codes: F30, F37, G13, G17, C45

*School of Economics and Finance, Queen Mary, University of London, Mile End Rd, London E1 4NS, UK. E-mail: d.bianchi@qmul.ac.uk Web: whitesphd.com

†Judge Business School, University of Cambridge, Trumpington Street, Cambridge, CB2 1AG, UK. E-mail: as2633@cam.ac.uk

1 Introduction

Observable economic fundamentals suggest the recent European sovereign debt crisis, which originated in peripheral EU economies and later spread around the continent and beyond, was primarily due to the underpricing of credit risk in Greece. In this context, credit default swaps (CDS henceforth) played a crucial role serving as a barometer for the unfolding sovereign debt crisis (see Fig.1).[1] The cost of hedging against sovereign default risk, as proxied by the CDS premium, clearly increases over the 2011/2012 period vis-a-vis European countries.

This paper highlights the role of macroeconomic fundamentals in explaining the relative level of CDS spreads across major OECD economies. More specifically, we use advanced statistical learning methods applied to an otherwise standard predictive framework to ask two simple questions that are still unresolved in the literature. First, we investigate whether the view of CDS spreads being predominantly driven by aggregate economic fundamentals has any foundations in a moderately-sized predictive regression framework. Second, we investigate to what extent non-linear mapping assumptions between economic fundamentals and CDS spreads improve our understanding of the dynamics of sovereign default risk. Equivalently, we ask whether there is any out-of-sample gain in departing from standard sparse and dense linear regression methods such as the lasso, ridge and elastic net in lieu of more flexible non-parametric, non-linear, tree-based regression methods.

Our primary contributions are three-fold. First, we provide a new benchmark for the use of predictive regression models in measuring CDS spreads. This is summarised in two ways. The first is the high out-of-sample predictive R^2_{oos} of non-linear statistical learning methods, in particular random forests, relative to commonly used linear sparse and dense predictive strategies. Second, we provide evidence that tree-based regressions generalise better than linear forecasting strategies across countries, historical periods and aggregate market

[1] A single-name CDS is an over-the-counter credit derivative contract between a seller and a buyer that provides the buyer protection against the default of an underlying entity (corporate or sovereign). The buyer pays the seller a fee called the CDS spread or premium, and in exchange the buyer receives compensation from the seller in the event of a default.

uncertainty conditions.

The second, and perhaps more important contribution of our paper is to provide robust empirical evidence that when used in a non-linear framework, macroeconomic variables alone show significant out-of-sample explanatory power when pricing sovereign credit risk. On an absolute basis, this power is not restricted to specific periods; random forest regressions incorporating solely macroeconomic variables show strong out-of-sample R^2_{oos} across both high-uncertainty and low-uncertainty periods. Contrary to the existing literature, non-linear methods performs well both in peripheral as well as core European economies. In addition, we show that the importance of macroeconomic fundamentals is time varying, with unemployment rates and fluctuations of economic activity around the long term trend consistently ranked among the most important economic factors.

Our third contribution involves a "pure" out-of-sample test of CDS spread pricing. More precisely, we show that once our tree-based regression model has been trained and validated on a suitably large dataset, the model can be used to generate shadow CDS spreads over extended time periods where no existing CDS data is available. Such shadow spreads show good correlation with existing historical measures of sovereign risk, giving weight to their use as an alternative sovereign risk measure. In this respect, our evidence suggests that non-linear machine learning methods open up the possibility of generating shadow CDS spreads for nations with no existing liquid CDS contracts, providing economic agents with a more accurate view on the priced sovereign default risk of such nations.

The main empirical analysis is based on a sample of 29 OECD countries over the February 2011 to November 2019 period. Albeit short, the sample covers the bulk of the European sovereign debt crisis and the relativly calm period that followed. In other words, the sample covers two distinct market phases characterised by different trading dynamics and underlying economic activity. As far as CDS spreads are concerned, we look at the 3, 5, 7 and 10-year CDS premiums for each country. Country-specific economic activity is summarised by 13 macroeconomic indicators covering employment/labour markets, growth, monetary policy, fiscal health and foreign trade. The covariates are utilised individually to capture "level" effects

and subsequently interacted with the maturity of each CDS contract to capture "slope" effects.

Methodologically, we evaluate and compare a variety of sparse and dense penalised regression methods and principal component regressions, as well as generalised linear models and non-linear tree-based methods. While several of these strategies may appear overly simplistic, they have been shown to perform well in forecasting asset returns and risk premiums out of sample (see e.g. Rapach et al., 2013, Kelly and Pruitt (2013; 2015), Rossi and Timmermann, 2015, Freyberger et al., 2020, Feng et al., 2020, Kozak et al., 2020, Gu et al., 2019, Kelly et al., 2019, Sirignano et al., 2016, Chen et al., 2019, and Bianchi et al., 2021).

The main empirical results lend strong support for the use of macroeconomic fundamentals coupled with non-linear regression tree-based methods. More specifically, the evidence is largely in favour of using random forest regression methods and simple regression trees. In addition, random forests highlight the heterogeneous role of macroeconomic fundamentals both across countries (e.g. EU vs non-EU) and over time (e.g. during and after the EU sovereign debt crisis). All in all, unemployment rates, consumption expenditures and the output gap seem to be the strongest determinants of CDS spreads.

To interpret the economic validity of our results, we delve further into the model-implied CDS spreads and investigate (1) the reliability of the forecasts within a pure out-of-sample framework, that is, by training the models on specific groups of countries and forecasting the CDS premiums on the residual economies, and (2) by looking at the correlation between the model-implied CDS spreads and aggregate measures of market and economic policy uncertainty. We highlight two main results: first, the random forest generalises best in our pure out-of-sample exercise - it proxies the CDS spreads for countries which are not in the training sample better than competing strategies. Second, we show the CDS premiums implied by the random forest estimates strongly correlate with both the uncertainty on sovereign debt economic policies and aggregate market uncertainty measures. This is true both in-sample (2011-2019) and out-of-sample (comparing generating shadow CDS spreads over the pre-2001 period to available macroeconomic indicators over the same period).

To summarise, our empirical results support theoretical models that feature both time variation in risk prices and time-varying risk based on aggregate economic conditions as in, e.g., Pesaran et al. (2006), Longstaff et al. (2011), Ang and Longstaff (2013), Doshi et al. (2013), Jeanneret (2015), Kim et al. (2017), Oehmke and Zawadowski (2017), Doshi et al. (2017), Augustin et al. (2021), among others. This echoes and extends the results in Alessi et al. (2019); they show that for 11 Eurozone countries over the 2009-2013 period, the market price of economic risk for sovereign CDS is time varying and possibly increases when investor attention to economic fundamentals becomes extreme. However, based on linear sparse regression methods, they show that past crisis periods there is a disconnect between market developments and macroeconomic fundamentals. We extend this literature by showing that when acknowledging the existence of latent non-linearities between economic growth, labour market conditions, fiscal capacity and credit risk premiums, one can show that economic fundamentals alone explain a great deal of time series and cross-sectional variation in sovereign CDS premiums. In this respect, our paper, by means of non-linear statistical learning methods, strengthens the argument that the market price of sovereign risk is tightly connected to economic fundamentals.

While non-linear machine learning methods have been explored in-depth in respect to binary sovereign default classification, research into the efficacy of machine learning for sovereign credit risk pricing is sparse. Manasse et al. (2003) provide one of the first approaches to sovereign default forecasting by means of a logistic regression alongside classification trees. The paper identifies macroeconomic variables reflecting solvency and liquidity factors that predict a debt-crisis episode one year in advance. Similarly, Fioramanti (2008) use standard artificial neural networks to understand if a nation's economic status in one year may help to predict a sovereign debt crisis the following year. Basu and Perrelli (2019) generate crisis lists for 169 countries over 27 years and evaluate the performance of signal extraction vs. machine learning techniques for the prediction of such crises. Silva et al. (2019) use random forests to classify the sovereign ratings of various countries, using PCA and clustering as a form of variable selection. Unlike these studies, our objective is not to provide an "early warning" system based on recession indicators and/or financial crsis, but rather to explain the high-frequency variation in sovereign

spreads or credit default swap (CDS) premiums by means of economic fundamentals alone. As highlighted by Aizenman et al. 2013; Bernoth and Erdogan 2012; Pasquariello 2014; Augustin and Tédongap 2016; Chernov et al. 2020, there is still no clear understanding on how market prices incorporate information on country-specific fundamentals over time to price default risk. We contribute to this still-open debate by showing that within the context of a reduced form regression method, once accounting for non-linear pricing relationships, economic fundamentals indeed explain the vast majority of the pricing dynamics of default risk on a higher frequency basis.

2 A simple motivating framework

In this section, we provide a simple motivation for the use of machine learning methods in the analysis of sovereign CDS based on macroeconomic fundamentals. Here we outline the broad reasoning, leaving the technical details to Appendix A.

The default intensity is assumed to depend on some economic factors $\boldsymbol{x}_t$ that affect the default-free term structure. In its simplest case, the intensity rate λ_t is linear in the parameters, such that is $\log \lambda_t = \boldsymbol{\beta}' \boldsymbol{x}_{t-1}$ (see, e.g., Gourieroux et al., 2006). Assuming a constant recovery rate and standard no-arbitrage conditions (see, e.g., Augustin and Tédongap, 2016; Doshi et al., 2013), the log of the CDS spread s_t can be defined as

$$s_t = \text{constant} + \lambda_t \propto \boldsymbol{\beta}' \boldsymbol{x}_{t-1}, \tag{1}$$

However, at least a priori, there is no reason why the pricing mechanism embedded in the default intensity rate (i.e. the functional form of $\log \lambda_t$) should necessarily be linear. For example, Doshi et al. (2013) and Doshi et al. (2017) show that by assuming the default intensity rate is a quadratic function of a set of latent state variables, i.e., $\lambda_t = (\boldsymbol{\beta}' \boldsymbol{x}_{t-1})^2$, the pricing kernel of CDS spreads is quadratic. Similarly, Augustin and Tédongap (2016) specify a non-linear dynamic for economic fundamentals which leads to a state-dependent CDS spread.[2] In

[2] Along the same lines, Galil et al. (2014), Jeanneret (2015), Blommestein et al. (2016), Lahiani et al. (2016)

this paper, we expand the existing literature and remain agnostic as to which functional form of the default intensity rate best fits the CDS spread dynamics, that is,

$$s_t \propto f\left(\boldsymbol{x}_{t-1}\right), \tag{2}$$

with $f\left(\boldsymbol{x}_{t-1}\right)$ a function of the economic fundamentals which is left unspecified. Although assuming the mapping between s_t and $\boldsymbol{x}_{t-1}$ as unknown a priori seems perhaps too general, a crude inspection of the data supports the idea of significant non-linear relationships between credit default spreads and macroeconomic variables. Figure 2 reports the 5-year CDS premiums against pairs of macroeconomic factors for a sample of countries, namely the US, UK, France and Japan. We compare the fitted value of a linear model (flat surface) against the fitted value from a third-order polynomial (curve). As an example, we consider some of the main macroeconomic indicators used in the empirical analysis such as industrial production, harmonised unemployment, the consumer price index and the composite leading indicator.[3]

Three main facts emerge: first, for a number of macroeconomic pairs and countries there is a great deal of non-linearity in relationship to the CDS spread. As an example, business cycle fluctuations and inflation have a highly non-linear interaction with the 5-year CDS for Japan. Similarly, there is evidence of a highly non-linear correlation between business cycle fluctuations and industrial production. Second, the nature of the non-linearity of credit risk pricing is far from obvious and possibly heterogeneous across macroeconomic variables. That is, specifying a particular functional form for a predictive model may be too restrictive, even if it is indeed non-linear. Third, a linear model may still successfully capture some local correlation between macroeconomic fundamentals and credit risk premiums. For instance, a linear function seems to approximate fairly well the relationship between the composite leading indicator, inflation and the 5-year CDS for France, especially when the macroeconomic factor pairs are in the region close to the sample mean.

show empirically that non linearity in the exposure of corporate default risk to economic conditions could arise because of endogenous structural breaks such as major credit and market events.

[3]The composite leading indicator (CLI) is designed to provide early signals of turning points in business cycles showing fluctuation of the economic activity around its long term potential level. CLIs show short-term economic movements in qualitative rather than quantitative terms.

Motivated by this literature and the simple empirical evidence from Figure 2, we investigate the possibility that a more precise measurement of sovereign CDS based on economic fundamentals alone (i.e. a more precise estimate of $f(\boldsymbol{x}_{t-1})$) can be obtained by using flexible, non-linear transformations of the data, an avenue that has been advocated by Stock and Watson (2002) within the context of forecasting macroeconomic time series. Differently from the Gaussian linear/quadratic models in Doshi et al. (2013),Doshi et al. (2017) and Chernov et al. (2020), we do not postulate a specific functional form connecting CDS spreads and state variables; instead we use various statistical techniques such as trees and ensemble modeling to tease out the relationship. Besides being agnostic about the functional form between CDS premiums and macroeconomic variables, the use of tree-based methods has two additional advantages relative to dense and sparse linear regression methods which have been often adopted (see, e.g., Alessi et al., 2019 and the references therein).

First, there is non-negligible correlation across different macroeconomic indicators, and the overall correlation structure is not necessarily stable over time. Figure 3 shows this case in point. Macroeconomic variables tend to be quite highly correlated. Apart from increasing the variance of the predictions, standard forecasting methods often break down when predictors are redundant and/or carry time-varying signals (see, e.g., Bühlmann et al., 2013 and Gu et al., 2020). In this respect, penalised regression methods such as lasso (see, e.g., Tibshirani, 1996), ridge (see, e.g., Hsiang, 1975), and elastic net (see, e.g., Zou and Hastie, 2005) are helpful in reducing the model dimension at the cost of leaning towards selecting one variable from a group of correlated or nearly linearly dependent predictors.[4] On the other hand, tree-based methods are more robust towards redundant predictors (see Hastie et al., 2009). In fact, decision trees make no assumptions on relationships between features. If features A, B are heavily correlated, little information can be gained from splitting on B after having split on A. As a result, a splitting rule would typically move to a third variable C without the

[4]Extensions such as the group-lasso proposed by Yuan and Lin (2006) may help in accounting for some economic structure in the data since they do consider a group-specific penalization term. However, they often do not make explicit the correlation across groups of predictors; they assume a priori group sparsity in the economic structure of the data and also assume no time-varying interdependence among groups of regularized predictors; all features which seem contradicted by the data as Figure 3 suggests.

aggregate loss function/accuracy being affected. There may still be some danger in over-fitting the model if the level of correlation between features in the training set does not generalize to unseen data. However, this is less of an issue for tree ensembles, especially when feature bagging (independent random samples of feature subsets used to build each tree) is employed, e.g., Random Forests.

Second, both sparse and dense linear predictive methods typically entail either an implicit or explicit reduction of the model space with the intention to arbitrarily lower model complexity in order to minimize predictive loss. In penalised regressions, increasing the tuning parameter (i.e., increasing shrinkage) leads to a higher bias. By utilizing cross-validation, the researcher aims to balance the bias-variance trade-off by adjusting the tuning parameter. Similarly, in factor models the optimal number of latent common components is chosen by using some information criterion to reduce the model variance at the cost of increasing bias (see, e.g., Bai and Ng, 2002). On the other hand, tree-based methods (and ensemble trees in particular) retain all information and balance the bias-variance trade-off via reducing the complexity of the tree structure through cost complexity pruning rather than reducing the features/input space. We show that this is relevant when forecasting CDS premiums over time.

The existing literature on sovereign credit risk has vastly ignored the potential capability of machine learning techniques to address the issue of non-linearity and model regularization. Arguably, this comes at the expense of not fully capturing the full extent to which credit risk is priced based on economic fundamentals alone. This is the focus of our paper.

3 Research design

In this section, we outline the research design for our empirical analysis. We begin with an overview of our data collection strategy, as well as outlining the various linear and non-linear frameworks implemented in the main empirical analysis. We conclude with a short discussion on the estimation strategy. We leave most technical and computational details to the appendix for the interested reader.

3.1 Data

The empirical analysis is based on a set of global macroeconomic data from the OECD Revision Analysis Data set, an online facility documenting monthly revisions to OECD Main Economic Indicators (MEI) across 36 member states. Macroeconomic data is available on a monthly basis from February 1999, with historical time-series data provided as far back as 1960 for each variable in each monthly revision release. In order to maximise the size of the data set and include the largest possible number of countries we focus on the period from February 2011 to November 2019. Such a time frame allows the inclusion of multiple economic shocks and high uncertainty periods across a range of different OECD countries. Although relatively short, Figure 1 shows that the sample period covers periods of intense variation both in the cross section (left panel) and the time series (right panel) of CDS spreads. As a matter of fact, the left panel shows that both the cross-sectional average and volatility of CDS spreads significantly increase throughout the unfolding of the European sovereign debt crisis in 2011/2012. The right panel shows that the variation primarily comes from European economies and emerging economies, e.g., Turkey, whereas developed non-European nations tend to have relatively stable CDS premiums.

As far as macroeconomic indicators are concerned, the OECD Revision Analysis dataset provides a standardised list of variables covering employment/labour markets, growth, monetary policy, fiscal health and foreign trade. Due to idiosyncratic revision reporting by member states, many of these standardised variables are only available for a small handful of countries. In order to ensure consistent monthly availability for macroeconomic variables across each in-sample country, we choose a representative subset of 13 variables available across all countries and months in our sample. Table 1 provides a description of the 13 macroeconomic indicators used in the main empirical analysis. We consider the current account balance (as % of GDP) as a fiscal capacity measure. Labour market conditions are captured by the harmonized unemployment rate (in %). Consumer prices index (% YoY) proxies aggregate inflation and international trade in goods – both exports and imports (% monthly growth) – captures

international commercial trade. Finally, a variety of GDP components (all % growth rate) including government consumption expenditures, gross capital formation, private consumption expenditures and deviations from the output trend (i.e. composite leading indicator) proxy for economic output. Given the choice of time frame and variable selection, our final monthly data set encompasses 104 months of data across 29 OECD member states.[5]

Table 2 reports the sample average and standard deviations of all measures across all of the countries in our data set. Similar to the CDS premiums, there is substantial cross sectional variation in the macroeconomic indicators. Not surprisingly, countries with weaker economic fundamentals tend to have higher CDS spreads on average. EU peripheral countries such as Italy and Spain tend to have higher (lower) unemployment (industrial production growth), which is matched by a relatively higher CDS premiums. On the other hand, countries with relatively more robust economic fundamentals such as Germany, USA and the Nordic states have lower (higher) unemployment (industrial production growth), which is matched by lower-than-average CDS premiums. This is in line with the conventional wisdom that sovereign default risk is linked to economic fundamentals (see, e.g., Pesaran et al., 2006, Longstaff et al., 2011, Jeanneret, 2015, Kim et al., 2017, and Alessi et al., 2019). However, Figure 2 shows that the mapping between unemployment, industrial production and CDS spreads is far from linear.

Daily sovereign CDS spreads are obtained via IHS Markit, who aggregate marks from sell-side contributors to generate composite daily spreads. We focus on daily 3yr, 5yr, 7yr and 10yr CDS spreads - i.e. for a given country in our data set, each day contains four CDS premiums.

3.1.1 Interpolating macroeconomic variables

While CDS premiums are available on a daily basis, the set of macroeconomic variable is available monthly and therefore more sparse along the time series dimension. One obvious alternative could be to aggregate CDS premiums monthly and match the frequency of the macroeconomic indicators. However, a training sample relying solely on monthly data would

[5]N.B. July and August 2016 have been omitted from our sample due to lack of data availability.

not be expansive enough to allow a careful out-of-sample evaluation of each model across different market conditions, that is, the sample would be too shallow and sparse to measure accurately the mapping between economic conditions and CDS premiums. In other words, given the limited amount of data, an empirical analysis based on monthly observations would limit the possibility to recursively estimate and assess the out-of-sample performances, which is particularly relevant given the significant time variation in sovereign CDS spreads.

To address this issue, we follow some of the existing literature in macroeconomic forecasting and adopt an over-sampling method to match the monthly frequency of the macroeconomic data with the higher frequency of the CDS premiums (see, e.g., Marcellino, 2004, English et al., 2005, Angelini et al., 2006, and Foroni et al., 2019). Specifically, we use cubic spline interpolation to interpolate between monthly revision data points, resulting in daily interpolated data for all macroeconomic variables associated with a given OECD country.[6]

An alternative approach would be to use repeated values of the macroeconomic indicators within a given month to match the daily frequency of the CDS premiums (see, e.g., Alessi et al., 2019). While this approach grounds on the idea that the latest revision release represents the most up-to-date market view for a particular macroeconomic variable, the repeated nature of individual variables for a particular country induces zero daily variance for a given variable between revision releases. In contrast, our cubic spline interpolation approach allows daily fluctuations more consistently aligned with the variation observed in sovereign CDS spreads. In order to apply cubic spline interpolation to our monthly macroeconomic data, we assume revision releases occur on the 15th day of each month. In cases where this date lies on a weekend or public holiday, the closest prior business day is used. Our interpolation only covers business days for which CDS quotes are available; weekends and public holidays are excluded. Further details on the interpolation of the macroeconomic indicators are provided

[6]Cubic spline functions rely on piece-wise polynomials as a means of interpolation. The use of piece-wise polynomials is preferable to polynomial-interpolation, which seeks a high-order polynomial capable of simultaneously fitting all interpolation data points. High-order polynomial interpolation has been shown to display Runge's Phenomenon (Runge, 1901), characterised by large observed oscillations close to the edges of the interpolation interval. Cubic spline interpolation utilises a greater number of smaller-degree polynomials fitted in a piecewise fashion across all points to be interpolated, resulting in a smooth curve that avoids Runge's Phenomenon.

in Appendix B. Appendix B shows that the cubic spline interpolation does not exacerbate the sampling variation of the macroeconomic data. More specifically, Figure B.1 shows that our interpolation schemes matches remarkably the dynamics of the raw macroeconomic data. This holds across macroeconomic indicators and countries. As a result, our daily interpolation represents a smooth approximation of the within-month underlying stochastic process.

Two comments are in order. First, using cubic spline interpolation in the manner described introduces potential look-ahead bias, as all available monthly data points for a given country-variable pair are utilised in the construction of piece-wise daily approximations. In Appendix B we show that the look-ahead bias is negligible as a recursive approximation. More specifically, we show that a cubic spline interpolation based on an expanding window, coherent with the dynamics of the forecasts, would delivery virtually the same results as using the entire data set. Second, when matching daily interpolated macroeconomic variables (i.e. our independent variables) with daily CDS spreads (dependent variables), the revision release assumption introduces a potential endogeneity/overfitting concern. That is, in a scenario where a particular revision release actually occurs on the 30th day of a given month, we risk incorporating future information currently unknown to the market when matching macroeconomic data with CDS spreads. To address this issue, we follow Rapach and Zhou (2021) and lag macroeconomic variables by a month, that is daily interpolated macroeconomic data is matched with 1-month-ahead realised CDS spreads. Figure 4 shows the timeline of the forecasting exercise. Macroeconomic variables are lagged by a month, for instance, the daily interpolated value of a macroeconomic indicator on February 15th is matched with the daily CDS on March 14th so that there is a mitigation of the look-ahead bias which inevitably is inherited by using the full sample for daily interpolation. The split between training/validation and testing is clarified in Section 3.4. In appendix A we provide evidence that a more conservative, and less prone to look-ahead bias, interpolation based on an expanding window instead of full sample does not materially change the results.

Note that our main empirical analysis relies on revised macroeconomic data rather than vintage economic information. The reason is twofold: first, our main objective is, through

the lens of machine learning methods, to provide a better understanding of the dynamics of sovereign CDS pricing and its economic drivers rather than to create a real-time trading strategy, for which vintage economic information may be more useful. In this respect, we rely on an "economic agents know" framework: the econometrician has limited information relative to *economic agents who*, in contrast, *know* the history of prior macroeconomic data and how credit risk spreads react to real quantities (see, e.g., Atanasov et al., 2019 for further discussion). Second, vintage data does not necessarily represents the actual information set used by investors either. This data is still subject to measurement errors and, in fact, investors can use past data that has been revised (see, e.g., Croushore, 2011). In this respect, dubbing vintage data as more "realistic" may make sense from a pure forecasting perspective, but this does not necessarily apply to the measurement of risk premium time-series variation which is the ultimate objective of our paper. In addition, the fact that macroeconomic variables are delayed by one month mitigates the effect of data revision in forecasting, as highlighted by Rapach and Zhou (2021).

3.2 Forecasting methods

3.3 Principal component regressions

In addition to the simple ordinary least squares (OLS), the first class of models we implement are data compression methods. This is a popular strategy which centers on the idea that one can reduce the set of predictors to a few latent components which summarise most of the time series variation in the original data. Undoubtedly, the most used data compression method within the context of forecasting macro and financial variables is the principal component regression (PCR). Although conceptually PCR shares the same goal as penalised regressions (see below), i.e. to reduce model complexity by balancing the bias-variance trade-off, the implementation is different. The main difference lies in the fact that PCRs are based on a set of latent factors, (the "principal components"), which are extracted from the data in an unsupervised manner, without conditioning on the target variable. The literature on forecasting with factor models is enormous and citing all relevant paper would be prohibitive. We leave the interested reader

to Ludvigson and Ng (2009); Stock and Watson (2011) for references.

While the optimal number of principal components can be optimised as a searchable hyperparameter (see, e.g., Bai and Ng, 2002), we err on the side of conservatism and explore a 5-component PCR.[7] Once the desired number of principal components have been extracted, a simple OLS model is fit in order to forecast futures CDS spreads.

3.3.1 Penalized regressions

A second class of models we implement is penalised linear regressions. In its general form, a penalized regression entails adding a penalty term on top of the standard quadratic loss function, i.e., the mean squared error, $\mathcal{L}_{OLS}(\boldsymbol{\beta}) = \frac{1}{t}\sum_{\tau=1}^{t-1}\left(s_{\tau+1} - \boldsymbol{\beta}^{\top}\boldsymbol{x}_{\tau}\right)^2$:[8]

$$\mathcal{L}(\boldsymbol{\beta};\cdot) = \underbrace{\mathcal{L}_{OLS}(\boldsymbol{\beta})}_{\text{Loss Function}} + \underbrace{\phi(\boldsymbol{\beta};\cdot)}_{\text{Penalty Term}} . \tag{3}$$

Depending on the functional form of the penalty term, the regression coefficients can be regularized and shrunk towards zero (as in ridge), exactly set to zero (as in lasso), or a combination of the two (as in elastic net). In Appendix C, we describe each method in detail.

Penalised regressions all involve a linear combination of input features/variables. To highlight the efficacy of non-linear parameter interactions (the hallmark of machine learning techniques), we investigate the possibility that regression trees are simply proxying for nonlinear transformations of the original predictors introduced as additive terms in an otherwise linear model (Gu et al., 2020; Bianchi et al., 2021). More specifically, for a given set of predictors $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ we use a second-order polynomial expansion to derive a new feature set encompassing all original predictors, their squared values and all unique pairwise feature interactions $x_i.x_j$ $(i \neq j)$. For our particular training data set, 13 original explanatory variables imply 105 input features. This new set of features is used to estimate an additional lasso regres-

[7] The first 5 component explain more than 70% of the time series variation in the data. In a set of unreported results we also explored the performance of a 10-component PCR. The results are qualitatively the same.

[8] Notice that for simplicity in the notation we did not include an intercept. However, the model includes a constant term. Penalization on the intercept would make the optimization procedure dependent on the initial values chosen for the CDS premium; that is, adding a fixed constant to the CDS premiums would not simply result in a shift of the prediction by the same amount.

sion. With slight abuse of notation, we dub such extended penalised regression a "generalised linear model" (GLM).

3.3.2 Tree-based regression methods

Both penalised and principal component regressions do not account for non-linear relationships. To address this issue, we consider a third class of models which explicitly consider non-linear transformations of the input features. First suggested by Breiman et al. (1984), regression trees are a non-linear, non-parametric machine learning model constructed via binary recursive partitioning, a process that iteratively splits data into partitions or branches. Regression trees are based on a partition of the input space into a set of "rectangles." Then, a simple linear model is fit to each rectangle. More specifically, start with $\mathcal{R}_1 = \mathbb{R}^d$. For each feature $j = 1, ..., d$, and for each value $v \in \mathbb{R}$ that we can split on, split the dataset:

$$I_< = \{i : x_{ij} < v\} \qquad I_> = \{i : x_{ij} \geq v\} \tag{4}$$

Estimate the parameters $\beta_<$ and $\beta_>$ of a given split:

$$\beta_< = \frac{\sum_{i \in I_<} s_i}{\mid I_< \mid} \qquad \beta_> = \frac{\sum_{i \in I_>} s_i}{\mid I_> \mid} \tag{5}$$

Assess the split quality via MSE:[9]

$$MSE = \sum_{i \in I_<} (s_i - \beta_<)^2 + \sum_{i \in I_>} (s_i - \beta_>)^2 \tag{7}$$

The split point with the lowest loss is chosen. As a final step, the algorithm is recursed on each child node. This process is repeated until the desired stopping criterion is met. Figure C.1 displays an example of a binary partition (Panel (a)) and the corresponding regression tree

[9] An alternative to the mean squared error to measure the split quality is the mean absolute error:

$$MAE = \sum_{i \in I_<} \mid s_i - \beta_< \mid + \sum_{i \in I_>} \mid s_i - \beta_> \mid \tag{6}$$

(Panel (b)). To mitigate concerns of overfitting, we implement a "cost complexity pruning" method which penalises the number of leaves in the tree, that is penalises the depth of the tree structure. We opt to regularise the complexity of our trees by optimising (via cross-validation) the minimum number of samples required at each terminal node subsequent to splitting, as well as the minimum of samples at each internal node required prior to evaluating further split points.

While pruning may mitigate concerns on ovefitting, regression trees suffer from high variance. This means that if we split the training data into two parts at random and fit a regression tree to both halves, the results can be quite different. In contrast, a procedure with low variance will yield similar results if applied repeatedly to distinct data sets; models with low variance tend to generalise well out-of-sample.

To address this concern, Breiman (2001) suggests a bootstrap aggregation ("random forest") extension encompassing multiple decision trees. The random forest model is an ensemble algorithm utilising a forest of individual regression trees to arrive at a final prediction. Each individual decision tree is trained in a similar manner to Section 3.3.2, with two key exceptions. Firstly, smaller bootstrap samples are chosen from the main training sample for model calibration. Secondly, a subset of randomly chosen explanatory variables is used to evaluate split points and minimise the chosen loss function. These exceptions have the benefit of greatly reducing the correlation between individual decision trees in a forest, giving additional power to an ensemble approach. For each regression tree in the forest, a model-implied prediction is calculated. Predictions from all individual decision trees are averaged in order to generate a final random forest model prediction.

3.4 Estimation Strategy

Following common practice in machine learning, we split the data into three sub-samples: a training set used to train the model, a validation set used to calibrate model hyper-parameters and a testing set which represents the out-of-sample period in a typical forecasting exercise.

The existing literature often implements a typical time-series split whereby as the training/validation set increases, the testing set remains the same size throughout the out-of-sample period (see, e.g., Gu et al., 2020; Bianchi et al., 2021). Differently from the previous literature, in this paper we opt for a rolling 1-month daily observation window for both training/validation (in-sample) and testing (out-of-sample). Figure 4 explains this visually. The daily approximation of monthly macroeconomic variables is potentially based on information from month $t-2$ to t, depending on data releases. To mitigate concerns of overlapping information between the predictors and the target CDS spreads, we lag macroeconomic variables by one month - i.e. to predict CDS spreads on December 13th 2019, we use daily macroeconomic observations until November 15th 2019. The out-of-sample evaluation then continues by rolling over both the training/validation and the testing sample by one month.

While it is common practice to use an expanding window approach, a rolling window approach is more robust towards structural breaks and regime changes (see, e.g., Pesaran and Timmermann, 2007; Clark and McCracken, 2009; Pesaran and Pick, 2011). The choice of a one month window of daily observations for both the in-sample and out-of-sample periods is dictated by the fact that we aspire to use a comparable amount of information for both the training/validation and testing periods.

As far as the cross-validation of model hyper-parameters is concerned, we provide all specifics and computational details in Appendix C.

3.5 Statistical performance

We compare the forecasts obtained from each methodology to a naive prediction based on the historical mean CDS spread. In particular, we calculate the out-of-sample predictive performance at each month t as $R^2_{t,oos}$, suggested by Campbell and Thompson (2008):

$$R^2_{t,OOS} = 1 - \frac{\sum_n^N \sum_m^M \sum_\tau^{\mathcal{T}} (s_{n,m,\tau} - \hat{s}_{n,m,\tau}(\mathcal{M}_s))^2}{\sum_n^N \sum_m^M \sum_\tau^{\mathcal{T}} (s_{n,m,\tau} - \bar{s})^2} \tag{8}$$

Here $s_{n,m,\tau}$ represents the realised CDS spread for a given country on a particular day within a month t, $\hat{s}_{n,m,\tau}(\mathcal{M}_s)$ represents the associated model-implied CDS spread forecast from a given model $\mathcal{M}_s$, and $\bar{s}$ represents the mean realised CDS spread for the 1-month train/validation period occurring directly prior to our 1-month out-of-sample test period. Evaluating $R^2_{t,OOS}$ is tantamount to evaluating whether model-implied CDS spreads have lower mean squared predictive errors in a given month t, relative to the 1-month historical average CDS spread in our pooled data set.

4 Sovereign CDS spreads and the macroeconomy

We start by forecasting the daily CDS spreads across countries and maturities assuming a 1-month training/validation and 1-month testing sample split. Both the train/validation and the test samples are rolled forward by a month – i.e. the prior 1-month test window now becomes the 1-month train/validation window in our new rolled-forward iteration (cfr. Figure 4). In such an iterative fashion, we end up with 100 concurrent months of out-of-sample performance figures $R^2_{t,OOS}$ for each model class.

As far as linear models are concerned, we follow Alessi et al. (2019) and consider two alternative specifications. That is, in addition to a simple linear specification, we also adopt a "level + slope" approach whereby the set of macroeconomic indicators is interacted with the maturity m of the associated CDS contract for a given country n, i.e.,

$$s_{nmt} = \beta_1^\top x_{nt} + \beta_2^\top (x_{nt} \times m) + \varepsilon_{nmt} \tag{9}$$

Such an approach allows us to explicitly incorporate slope effects into each linear model, and ensures model predictions are unique to a particular maturity rather than representing an average spread over all maturities. Given our "level-only" model has 13 explanatory variables, the "level + slope" variant has 26 variables; 13 original macroeconomic variables alongside 13 macroeconomic variable-maturity interaction terms.

Note that to incorporate slope effects into a PCA framework, we adopt a slightly different

approach. Rather than interacting each macroeconomic variable with the associated CDS maturity for a given country/day and then applying PCA, we initially apply PCA to the "level-only" framework as before. Once these principal components are obtained, we interact the *principal components* with the associated CDS maturities rather than the macroeconomic variables themselves. Given that principal components are a distilled representation of our original macroeconomic variables, this approach has the benefit of allowing us to capitalize on both dimensionality reduction and model slope effects concurrently.

4.1 Out-of-sample R^2_{oos}

We first report the unconditional out-of-sample performance of each forecasting model, that is, we report the average $R^2_{t,OOS}$ over the recursive testing sample $t = t_0, \ldots, T$. We look at three distinct groups of countries: the entire set of 29 countries in our sample (Global), the European economies (EU), and non-peripheral European economies (core-EU).[10] Among these three clusters we look at three alternative scenarios: the whole sample, periods of high economic uncertainty and periods of low economic uncertainty.

The level of economic uncertainty within a given time period is identified through one of the Economic Policy Uncertainty (EPU) indices introduced by Baker et al. (2016). Our particular focus is on the Sovereign Debt & Currency Crises "categorical" sub-index, which determines monthly uncertainty index values based on sovereign debt crisis and currency crises news mentions across the Access World News database of 2,000 U.S. newspapers.[11] We define a given month as "high-uncertainty" if the associated monthly EPU Sovereign Debt & Currency Crises index value exceeds a threshold equal to one standard deviation above the long-run index mean. Of the 100 out-of-sample test months available within our data set, 13 months are classified as high-uncertainty while the residuals are classified as low-uncertainty periods.

[10] The reason why we focus on core-EU instead of perhipheral EU for the main empirical analysis is due to data limitations. Greek sovereign CDS were not tradable for the vast majority of our sample, so by excluding Greece we would have only been left with Italy, Spain and Portugal as peripheral economies.

[11] Each categorical sub-index incorporates general economic, uncertainty, and policy terms alongside specific "categorical" policy terms and is multiplicatively normalized to have a mean of 100 over the 1985-2010 period. In our case, articles that fulfil the requirement to be classified as EPU as well as containing any sovereign debt/currency crisis-related terminology would be included in the Sovereign Debt & Currency Crises sub-index.

Table 3 reports the results. Three facts emerge: first, by interacting macroeconomic variables with the maturity of each CDS contract the performance of linear models, both dense and sparse, significantly increases. With the exception of PCA, the average $R^2_{t,OOS}$ of all models substantially increase when macroeconomic indicators are interacted with the CDS maturity. This result is in line with Alessi et al. (2019). Second, and perhaps more importantly, while simple regression trees perform as well as penalised regressions, random forest significantly outperform all competing predictive strategies with an average $R^2_{t,OOS}$ that is almost 20% higher across clusters of countries. Notably, the spread in performances tend to be higher for Global premiums relative to Core EU economies. Thirdly, we note only a modest improvement when using penalised regression methods relative to simple OLS models when both level and slope factors are used as predictors. Similar modest outperformance is observed across both the linear model with squared terms and interactions (GLM) as well as the principal component regression with five latent components.

The second and third panels of Table 3 report the performance each predictive strategy across different regimes of economic policy uncertainty, regimes which are identified as described above. Again, two interesting facts emerge: first, the performance of penalised regressions tend to improve (worsen) during periods of low (high) policy uncertainty in relation to sovereign debt/currency crisis. For instance, the average $R^2_{t,OOS}$ during low-uncertainty periods is around 6% higher than during periods of high policy uncertainty. In fact, during low-uncertainty periods the performance of linear methods are somewhat comparable in magnitude to the random forest – e.g, 63% for elastic net in Core EU vs 71% for the random forest over the same cluster of countries. However, such equivalence is not generalised and it is only confined to Core EU economies.

Second, the performance of the random forests seems to increase during periods of high policy uncertainty. For instance, the average $R^2_{t,OOS}$ for the Global premiums is 88% during periods of high policy uncertainty against a value of 80% during low uncertainty periods, a decrease of 10% in relative terms. The difference in the performance, and in favour of the high-uncertainty periods, is even larger for the regression trees. During low-uncertainty periods,

regression trees underperform "level + slope" penalised regressions by a non-trivial margin. As a result, while regression trees perform very strongly in high-uncertainty periods, overfitting in low-uncertainty periods hampers their overall performance relative to top-performing linear models.

As a whole the results show that non-linear models tend to outperform both sparse and dense linear regression methods in measuring the CDS spreads across different clusters of countries. Even more so during periods of debt/currency distress.

In Section 3.4, we highlight the construction of "in-sample" train/validation windows and "out-of-sample" test windows on a 1-month rolling basis. While the rolling out-of-sample window contains data unseen to the model during training/validation, the model has nonetheless seen earlier data from all countries in the model training/validation stage – if spreads for a given country are persistent, the model will give an excessively high view on out-of-sample performance. To expand the previous results and mitigate concerns about the persistence of economic shocks in a time series sense, we bridge the gap between out-of-sample predictability and "synthetic" CDS spreads by evaluating each forecasting method on truly unseen, non-overlapping, macroeconomic data – i.e. the sovereign CDS spreads of countries who are not present in training/validation data. We estimate four different combination of in-sample (IS) vs out-of-sample (OOS) countries: EU (IS) vs non-EU (OOS), Non-EU (IS) vs EU (OOS), Core EU (IS) vs peripheral EU (OOS), and peripheral EU (IS) vs Core EU (OOS).

The bottom panel of Table 3 reports the results. Two interesting results emerge: first, with the exception of random forests, none of the models generalise well to non-EU economies when trained on EU data and vice-versa. Even for random forests, despite the fact that R^2_{oos} is positive and much larger than competing strategies, it is substantially lower than the typical out-of-sample recursive calculation shown in the top three panels of Table 3. This suggests that breaking down the time series dependence of the train/validation vs testing sample substantially decreases the out-of-sample performance of each forecasting model. Related to this, the second result shows that, instead, alternative models tend to generalise better when the in-sample vs out-of-sample data are more coherent, as it is the case for EU economies. More

specifically, when models are trained on Core-EU countries, the performance for the peripheral economies tend to be quite remarkable. One possible reason for such remarkable performance is that core and peripheral European countries share sources of cross-sectional contagion and commonalities in economic fundamentals that makes the train/validate sample and the test sample similar in terms of data generating processes.

Although instructive, Table 3 does not allow us to understand how different models perform dynamically throughout the whole out-of-sample period. As the $R^2_{t,oos}$ is calculated for each testing period monthly, we can look at the time-varying performance of each class of models in turn. Figure 5 reports the differential of the $R^2_{t,oos}$ of each model against the OLS, calculated for each of $t = 1, \ldots, 100$ months that are available as testing samples. A positive value means that $R^2_{t,oos} > R^2_{t,oos}(OLS)$, with $R^2_{t,oos}(OLS)$ the out-of-sample fit obtained from the OLS predictive regression.

For the ease of exposition we report the results for four representative class of models, namely the PCA with 5 components, the elastic net, a standard regression tree and a random forest regression. As far as linear models are concerned, we focus on the "level + slope" specification. The left panel shows the results for global CDS spreads. Two results emerge: first, the performance of elastic net is comparable over time to OLS. This confirms the results of Table 3. While pockets of volatility and uncertainty are present at the national and European level, especially in the early part of the sample, these are likely to be offset by the restrictive nature of the linear forecasts. As a result, localised shocks appear not to have a large impact on linear predictive strategies over time.

Second, while regression trees show a relatively volatile out-of-sample performance – in fact, regression trees underperform OLS during the 2017/2018 period – random forests consistently outperform the benchmark linear predictive regression. This means that, critically, outperformance is maintained across both high-uncertainty and low-uncertainty periods – while regression trees significantly outperform "level + slope" penalised regression models in high-uncertainty periods, our Random Forest model records an *additional* 20/30%, on average, out-of-sample R^2 at the global level. A similar picture emerges for the European countries

sample (mid panel of Figure 5): the flexibility of random forest regressions allow them to consistently outperform linear competing predictive strategies. The right panel of Figure 5 reports the results for the Core EU countries. Here, results are slightly different for the period across 2017/2018, whereby random forest regressions under-perform linear predictive regressions. This is offset by a large and positive $R^2_{t,oos}$ differential over the European sovereign debt crisis of 2011/2012, which characterised a highly volatile period for sovereign CDS (cfr. Fig 1). On the whole, Figure 5 suggests that the lack of notable out-performance from standard regression trees can be attributed to sporadic periods of incredibly poor performance where $R^2_{t,oos}$ figures become deeply negative.

Figure 6 shows the predicted CDS spreads obtained from a random forest regression model. The left panel shows the actual vs predicted CDS for peripheral EU economies, where the forecasts are produced based on a model trained on the macroeconomic indicators of Core-EU countries (cfr. bottom panel of Table 3). Notably, the dynamics of the realised vs predicted CDS spreads are rather consistent, despite the train/validation data for the random forest being based on a different set of countries. The right panel shows the results for a flipped exercise where peripheral EU countries are now used to estimate the random forest regression and then based on the trained/validate model we feed in macroeconomic indicators for Core-EU countries to produce forecasts of the corresponding CDS spreads. Again, despite some difference in the magnitude of the spreads, the dynamics of the actual and the predicted sovereign CDS are largely aligned.

The results in Table 3 and Figures 5-6 extend some of the existing literature on the ability of non-linear machine learning to forecast risk premiums (see, e.g., Gu et al., 2020, and Bianchi et al., 2021). In particular, similar to Bianchi et al. (2021) we find the predictability of sovereign CDS spreads implied by our top-performing machine learning model is not exclusive to "bad times" - this is in contrast with contrasts with evidence for equities (Rapach et al., 2010; Dangl and Halling, 2012) and treasury bonds (Gargano et al., 2019). While regression trees overfit during low-uncertainty periods, our random forest model maintains its strong performance; $R^2_{t,oos}$ outperformance equates to more than 20% over "level + slope" penalised regressions. As

a result, random forests record significant outperformance relative to all "level-only","level + slope" and regression tree models across both high-uncertainty and low-uncertainty periods, with the best performance observed during high-uncertainty periods.

4.2 Pairwise comparison of predictive accuracy

We follow Gu et al. (2020) and implement a pairwise test as proposed by Diebold and Mariano (2002) (DM) to compare the predictions from different models. Diebold and Mariano (2002) show that the asymptotic normal distribution can be a very poor approximation of the test's finite-sample null distribution. In fact, the DM test can reject the null too often, depending on the sample size and the degree of serial correlation among the forecast errors. To address this issue, we adjust the DM test by making a bias correction to the test statistic as proposed by Harvey et al. (1997).

We compare daily cross-sectional average out-of-sample prediction errors from each model. Results are displayed in Table 6. We color-code the statistical significance of the test, with bold values suggesting a rejection of the null – H_0: the pairs have the same performance – at the 5% level. Our null hypothesis assumes identical forecast accuracy across each model, with positive test statistics indicating outperformance of the column model relative to the row model. The top panel reports the results for the whole cross section of countries. We note the statistical outperformance of random forests relative to all other "level-only" and "level + slope" model categories, both linear and non-linear. Also apparent is the poor performance of the GLM model, with DM-test statistics indicating superior predictive power for all "level-only" and "level + slope" OLS/penalised regression models relative to our GLM framework.

Panel B shows the performance concerning EU countries. Looking at our Diebold-Mariano test statistics, we again note the statistical outperformance of random forest forecasts relative to virtually all other model categories at the European level – with test statistics for outperformance significant at the 5% threshold for seven out of twelve pairwise comparisons and very close to being significant for the remaining five cases. Panel C shows that similar results are obtained for Core-EU countries.

5 Understanding the model-implied CDS spreads

In this section we delve further into the dynamics of the CDS premiums implied by the random forest predictive regression, which is unequivocally the best performing strategy when measuring time-series variation in CDS spreads. More specifically, we first look into the importance of each macroeconomic variables within the context of random forests across time and for global, EU and Core-EU economies. Second, we look at the economic content of the model-implied CDS spreads exploring correlations with aggregate measures of economic uncertainty and risk aversion. This is done for the 2011-2019 period in which sovereign CDS data is available, as well as for the pre-2001 period by producing "shadow" CDS spreads based on model estimates using macroeconomic data available over this period.

5.1 Macroeconomic variables importance

We follow in the spirit of Gu et al. (2020) and assess the marginal relevance of each macroeconomic indicator on the forecasting performance by evaluating the effect on the $R^2_{t,oos}$ by dropping individually each explanatory variable from our train/validation and test data set.

More specifically, each of the 13 macroeconomic variables is omitted in turn and our random forest model is cross-validated on the remaining 12 variables. Once hyper-parameters have been tuned, the calibrated model is evaluated on the associated 1-month test window and $R^2_{t,oos}$ is recorded. When repeated for all the 13 variables, we are left with corresponding $R^2_{t,oos}$ figures for 13 different models, each with a different variable dropped. Each of those 13 out-of-sample performance is then compared with the model incorporating all explanatory variables (i.e. the full 13-variable model). Variables are then given a rank based on the drop in $R^2_{t,oos}$ associated with their individual omission from the data set. Our data set is rolled-forward 1-month as highlighted in Figure 4, and the ranking process repeats. The final output is a 1-13 ranking for all explanatory variables during each out-of-sample test month. As a last step, we average the rank for each explanatory variable over all out-of-sample months and normalise the resultant average ranks to lie between 0 and 1. We execute this methodology individually for Global

data, Europe data and Core Europe data. Table 5 reports the results.

As far as the global risk premiums are concerned, Panel A shows that harmonised unemployment rate contributes the most for the vast majority of the sample period, holding the highest ranking from 2014 to 2018. The consumer price index and the composite leading indicator, which measures fluctuations of economic activity around the long term trend, also rank as top marginal forecasting variables throughout the sample. Perhaps surprisingly, variables related to international trade and country-specific industrial production rank at the bottom of the scale when it comes to their marginal forecasting importance. A similar result applies for imports of goods and services.

Except for a few nuances, Panel B shows that a similar picture emerges for Europe. Harmonised unemployment rate and the year-on-year growth rate of the composite leading indicator rank on top, whereas international trade variables rank at the bottom in terms of marginal forecasting importance. As far as Core-EU countries are concerned, there is more heterogeneity and a clear picture is more difficult to ascertain. In particular, consumption expenditures and total GDP also carry a high forecast contribution towards the end of the sample period.

As a whole, the contribution of unemployment and the deviations from the trend of output growth seem to carry the highest predictive content for CDS spreads across countries and over time.

5.2 Expected CDS spreads and sovereign debt crisis

We next investigate whether the expected CDS spreads forecasts are linked to sovereign debt and currency crises. More specifically, we regress the CDS spread forecast averaged across maturities and countries obtained from the best-performing random forest on a Economic Policy Uncertainty (EPU) sub-index as introduced by Baker et al. (2016). We focus in particular on the Sovereign Debt & Currency Crises "categorical" sub-index.[12]

[12]The "Sovereign Debt & Currency Crises" focuses on several keywords from media, such as sovereign debt, currency crisis, currency crash, currency devaluation, currency revaluation, currency manipulation, euro crisis, Eurozone crisis, european financial crisis, european debt, asian financial crisis, asian crisis, Russian financial crisis, Russian crisis, exchange rate.

In addition to uncertainty related to sovereign debt crises we also consider a set of potentially key drivers of macroeconomic uncertainty and risk aversion as suggested by asset pricing theory and previous evidence (see, e.g., Alessi et al., 2019). In particular, we consider the Geopolitical Risk Index (GRI) proposed by Caldara and Iacoviello (2018). The index reflects media coverage of geopolitical risk for each month, normalized to average a value of 100 in the 2000-2009 decade. The second variable included in our set of additional variables of interest is the Global Risk Aversion Index (GRAI) proposed by Bekaert et al. (2019). This index represents a utility-based measure of time-varying risk aversion calculated from observable financial information at high frequencies.[13] Thirdly, we also consider the Global Uncertainty Index (GUI) proposed by Bekaert et al. (2019). Their measure is 81% correlated with the Jurado et al. (2015) measure, extracted from macro data, and 34% correlated with the Economic Political Uncertainty index constructed in Baker et al. (2016).

Figure 7 shows some preliminary visual correlation. Each time series has been standardised for ease of exposition. The top-left panel shows the average forecast of the global CDS spreads against the GUI. The correlation with global uncertainty as measured by Bekaert et al. (2019) is quite notable for the first part of the testing sample and for the 2017/2019 period. On the other hand, the spike in global uncertainty over the 2015/2017 is not followed by the model-implied expected global CDS spreads. When considering the EPU Sov index, the top-right panel shows that there is indeed a quite remarkable correlation with the expected CDS spreads throughout the testing sample. That is, the (standardised) trajectory of expected CDS spreads follows closely media attention and coverage of sovereign debt and currency crises.

The bottom panels implement a "backward" looking graphical representation. More specifically, we train/validate the random forest regression model over the 2011-2019 period and produce forecasts for the 1992-2001 period. This is a full out-of-sample exercise in the spirit of the bottom panel of Table 3. The bottom-left panel of Figure 7 compares the expected CDS spread against the GUI. Similar to the sample from 2011 to 2019, although far from being

[13]The instrument set includes a detrended earnings yield, corporate return spread (Baa-Aaa), term spread (10yr-3mth), equity return realized variance, corporate bond return realized variance and equity risk-neutral variance.

perfect, there is significant correlation between our "synthetic" CDS spreads and the index of global uncertainty, particularly over the period from 1994 to 1998. On the other hand, the bottom-right panel shows that the correlation between the expected global CDS spread and the EPU Sov index is relatively lower in the pre-2001 period. The main reason lies in the spike in the index during the Russian crisis towards the end of 1997/early 1998. In this respect, lower correlation between the synthetic CDS and the EPU for sovereign and currency crisis is not entirely unexpected; Russia is not within the set of countries used to generate our out-of-sample forecasts.

Table 6 expands the visual impression of Figure 7 and explores more formally the determinants of the expected global CDS spreads. More specifically, we estimate time series regressions where the dependent variable is the expected CDS spread averaged across countries and maturities and the independent variables are the set of uncertainty and risk aversion indexes outlined above. In the spirit of Figure 7, we estimate the regressions both for the sample from 2011 to 2019, for which sovereign CDS and macroeconomic data are available, and for the sample from 1992 to 2001, where sovereign CDS contracts are not available for the OECD countries in our sample.

Two comments are in order. First, while non-OECD nations lacking tradable CDS contracts exist, detailed monthly macroeconomic data for those countries are mostly unavailable to publicly accessible sources. As a result, in order to explore the concept of sovereign "shadow CDS pricing", we generate sovereign "synthetic" CDS spreads for *historical periods* where CDS prices were unavailable for *existing* OECD nations. This approach grants us rich monthly macroeconomic data for periods prior to the introduction of CDS contracts as a financial instrument in 2001, and allows us to see how machine learning model-implied CDS spreads correlate with existing measures of sovereign risk present during the historical periods in question.

Second, most of the indexes used as covariates for the regression analysis are only available at the monthly frequency and are represented on different scales. As a result, we take the monthly average of the daily forecasts as a measure of expected CDS spreads for the post-2011 sample. For the pre-2001 sample, we forgo daily interpolation and instead utilise monthly

uncertainty index values alongside monthly "shadow" CDS forecasts. In addition, we rescale both the expected CDS spreads and the uncertainty/risk aversion indexes so that the regression coefficients can be interpreted as % sensitivity.

The left panel of Table 6 reports regression results for the sample in which both CDS and macroeconomic data are available. The results show that a 1% increase in the economic policy uncertainty related to sovereign debt and currency crisis translates to a 0.77% higher, on average, CDS spread across maturities and countries. The EPU index itself explains a great deal of the time series variation present within predicted CDS spreads with an in-sample R^2 as high as 82%, both when the EPU is considered in isolation or jointly with the GRAI index. The latter is the only alternative index possessing significant correlation with our expected CDS spreads.

Interestingly, similar results hold when considering "shadow" CDS spreads. Although by a lower magnitude, there is still a significant positive correlation between both the economic policy uncertainty index and the global risk aversion index and expected global CDS spreads.

6 Conclusions

We provide evidence in favour of a significant non-linear, time-varying dependence between sovereign credit default swap (CDS) spreads on country-specific macroeconomic indicators for the majority of OECD economies. Macroeconomic fundamentals explain more than 80% of the out-of-sample variation in CDS spreads when robust regression tree methods are used. This is more than 25% higher than benchmark sparse and dense linear machine learning methods.

We test the consistency of the model-implied CDS spreads across different purely out-of-sample scenarios, e.g., training a random forest regression on EU countries and predicting the CDS spreads of non-EU economies. Our predicted CDS premiums correlate with the uncertainty on sovereign debt economic policies, and are primarily driven by unemployment rates and the fluctuation of economic activity around its long term level.

In addition, we provide evidence that "shadow" sovereign CDS spreads, based on macroe-

conomic fundamentals during historical periods for which sovereign CDS contract were unavailable, highly correlate with economic policy uncertainty measures related to both sovereign debt/currency crises and global risk aversion.

References

Aizenman, J., Hutchison, M., and Jinjarak, Y. (2013). 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(C):37–59.

Alessi, L., Balduzzi, P., and Savona, R. (2019). Anatomy of a sovereign debt crisis: CDS spreads and real-time macroeconomic data.

Ang, A. and Longstaff, F. A. (2013). Systemic sovereign credit risk: Lessons from the us and europe. *Journal of Monetary Economics*, 60(5):493–510.

Angelini, E., Henry, J., and Marcellino, M. (2006). Interpolation and backdating with a large information set. *Journal of Economic Dynamics and Control*, 30(12):2693–2724.

Atanasov, V., Moller, S. V., and Priestley, R. (2019). Consumption fluctuations and expected returns. *The Journal of Finance*, n/a(n/a).

Augustin, P., Sokolovski, V., Subrahmanyam, M. G., and Tomio, D. (2021). In sickness and in debt: The covid-19 impact on sovereign credit risk. *Journal of Financial Economics*.

Augustin, P. and Tédongap, R. (2016). Real economic shocks and sovereign credit risk. *Journal of Financial and Quantitative Analysis*, pages 541–587.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty*. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Basu, S. S. and Perrelli, R. (2019). External crisis prediction using machine learning: Evidence from three decades of crises around the world.

Bekaert, G., Engstrom, E. C., and Xu, N. R. (2019). The time variation in risk appetite and uncertainty. Technical report, National Bureau of Economic Research.

Bernoth, K. and Erdogan, B. (2012). Sovereign bond yield spreads: A time-varying coefficient approach. *Journal of International Money and Finance*, 31(3):639–656.

Bianchi, D., Büchner, M., and Tamoni, A. (2021). Bond risk premiums with machine learning. *The Review of Financial Studies*, 34(2):1046–1089.

Blommestein, H., Eijffinger, S., and Qian, Z. (2016). Regime-dependent determinants of euro area sovereign cds spreads. *Journal of Financial Stability*, 22:10–21.

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

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and Regression Trees*. Taylor & Francis.

Bühlmann, P., Rütimann, P., van de Geer, S., and Zhang, C.-H. (2013). Correlated variables in regression: clustering and sparse estimation. *Journal of Statistical Planning and Inference*, 143(11):1835–1858.

Caldara, D. and Iacoviello, M. (2018). Measuring geopolitical risk. *FRB International Finance Discussion Paper*, (1222).

Campbell, J. Y. and Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Chen, L., Pelger, M., and Zhu, J. (2019). Deep learning in asset pricing. *Working Paper*.

Chernov, M., Schmid, L., and Schneider, A. (2020). A macrofinance view of us sovereign cds premiums. *The Journal of Finance*, 75(5):2809–2844.

Clark, T. E. and McCracken, M. W. (2009). Improving forecast accuracy by combining recursive and rolling forecasts. *International Economic Review*, 50(2):363–395.

Croushore, D. (2011). Frontiers of real-time data analysis. *Journal of economic literature*, 49(1):72–100.

Dangl, T. and Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*, 106(1):157–181.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134–144.

Doshi, H., Ericsson, J., Jacobs, K., and Turnbull, S. M. (2013). Pricing credit default swaps with observable covariates. *The Review of Financial Studies*, 26(8):2049–2094.

Doshi, H., Jacobs, K., and Zurita, V. (2017). Economic and financial determinants of credit risk premiums in the sovereign cds market. *The Review of Asset Pricing Studies*, 7(1):43–80.

English, W., Tsatsaronis, K., Zoli, E., et al. (2005). Assessing the predictive power of measures of financial conditions for macroeconomic variables. *BIS Papers*, 22:228–252.

Feng, G., Giglio, S., and Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, 75(3):1327–1370.

Fioramanti, M. (2008). Predicting sovereign debt crises using artificial neural networks: a comparative approach. *Journal of Financial Stability*, 4(2):149–164.

Foroni, C., Ravazzolo, F., and Rossini, L. (2019). Forecasting daily electricity prices with monthly macroeconomic variables.

Freyberger, J., Neuhierl, A., and Weber, M. (2020). Dissecting characteristics nonparametrically. *The Review of Financial Studies*, 33(5):2326–2377.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1.

Galil, K., Shapir, O. M., Amiram, D., and Ben-Zion, U. (2014). The determinants of cds spreads. *Journal of Banking & Finance*, 41:271–282.

Gargano, A., Pettenuzzo, D., and Timmermann, A. (2019). Bond return predictability: Economic value and links to the macroeconomy. *Management Science*, 65(2):508–540.

Gourieroux, C., Monfort, A., and Polimenis, V. (2006). Affine models for credit risk analysis. *Journal of Financial Econometrics*, 4(3):494–530.

Gu, S., Kelly, B., and Xiu, D. (2019). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273.

Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, 13(2):281–291.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Hsiang, T. (1975). A bayesian view on ridge regression. *Journal of the Royal Statistical Society: Series D*, 24(4):267–268.

Jeanneret, A. (2015). The dynamics of sovereign credit risk. *Journal of Financial and Quantitative Analysis*, pages 963–985.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

Kelly, B. and Pruitt, S. (2013). Market expectations in the cross-section of present values. *The Journal of Finance*, 68(5):1721–1756.

Kelly, B. and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2):294–316.

Kelly, B. T., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524.

Kim, T. S., Park, J. W., and Park, Y. J. (2017). Macroeconomic conditions and credit default swap spread changes. *Journal of Futures Markets*, 37(8):766–802.

Kozak, S., Nagel, S., and Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292.

Lahiani, A., Hammoudeh, S., and Gupta, R. (2016). Linkages between financial sector cds spreads and macroeconomic influence in a nonlinear setting. *International Review of Economics & Finance*, 43:443–456.

Longstaff, F. A., Pan, J., Pedersen, L. H., and Singleton, K. J. (2011). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*, 3(2):75–103.

Ludvigson, S. C. and Ng, S. (2009). Macro Factors in Bond Risk Premia. *Review of Financial Studies*, 22(12):5027–5067.

Manasse, P., Schimmelpfennig, A., and Roubini, N. (2003). Predicting sovereign debt crises. *IMF Working Papers*, 03.

Marcellino, M. (2004). Forecasting emu macroeconomic variables. *International Journal of Forecasting*, 20(2):359–372.

Mitchell, T. M. et al. (1997). Machine learning.

Oehmke, M. and Zawadowski, A. (2017). The anatomy of the cds market. *The Review of Financial Studies*, 30(1):80–119.

Pasquariello, P. (2014). Financial market dislocations. *The Review of Financial Studies*, 27(6):1868–1914.

Pesaran, M. H. and Pick, A. (2011). Forecast combination across estimation windows. *Journal of Business & Economic Statistics*, 29(2):307–318.

Pesaran, M. H., Schuermann, T., Treutler, B.-J., and Weiner, S. M. (2006). Macroeconomic dynamics and credit risk: a global perspective. *Journal of Money, Credit and Banking*, pages 1211–1261.

Pesaran, M. H. and Timmermann, A. (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics*, 137(1):134–161.

Rapach, D. and Zhou, G. (2021). Sparse macro factors. *SSRN Scholarly Paper*, (3259447).

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2):821–862.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2013). International stock return predictability: What is the role of the united states? *Journal of Finance*, 68(4):1633–1662.

Rossi, A. G. and Timmermann, A. (2015). Modeling covariance risk in merton's icapm. *The Review of Financial Studies*, 28(5):1428–1461.

Runge, C. (1901). Uber empirische funktionen und die interpolation zwischen aquidistanten ordinaten, zeitung fur math.

Silva, D. R. B. d., Rego, T. G. d., and Frascaroli, B. (2019). Sovereign risk ratings' country classification using machine learning.

Sirignano, J., Sadhwani, A., and Giesecke, K. (2016). Deep learning for mortgage risk. *arXiv preprint arXiv:1607.02470*.

Stock, J. H. and Watson, M. (2011). Dynamic factor models. *Oxford Handbooks Online*.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288.

Wu, T. T., Lange, K., et al. (2008). Coordinate descent algorithms for lasso penalized regression. *Annals of Applied Statistics*, 2(1):224–244.

Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49–67.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

Table 1: **Macroeconomic variables**

This table summarises the set of macroeconomic variables used in the main empirical analysis. The data are sampled from the OECD Revision Analysis Data set, an online facility documenting monthly revisions to OECD Main Economic Indicators (MEI) across major economies. The scale and the transformation of each variable is reported in parenthesis.

Variable	Description	Category
bop	Balance Of Payments - Current Account Balance (% of GDP)	Fiscal
unemp	Harmonised Unemployment Rate (%)	Labour Market
comp	Composite Leading Indicator: (% Growth)	Output
exp	Gdp: Exports Of Goods And Services (% Growth)	Output
gov	Gdp: Government Consumption Expenditure (% Growth)	Output
cap	Gdp: Gross Fixed Capital Formation (% Growth)	Output
imp	Gdp: Imports Of Goods And Services (% Growth)	Output
cons	Gdp: Private Consumption Expenditure (% Growth)	Output
gdp	Gdp: Total (% Growth)	Output
indus	Index Of Industrial Production (% Growth)	Output
cpi	Consumer Price Index (% Growth)	Prices
int_ex	International Trade In Goods - Exports (% Growth)	Trade
int_imp	International Trade In Goods - Imports (% Growth)	Trade

Table 2: **Descriptive statistics**

This table reports a set of descriptive statistics for all macroeconomic indicators used in the main empirical analysis and for each country in our sample. In addition to the sample mean and standard deviation of each macroeconomic indicator we report the mean and standard deviation of the 5-year CDS spread for each country. The sample period is from February 2011 to November 2019, monthly.

		Macro indicators													
		bop	comp	cpi	gdp_ex	gdp_gov	gdp_cap	gdp_imp	gdp_cons	gdp_tot	unemp	indus	int_ex	int_imp	CDS 5-year
Australia	Mean	-2.62%	2.72%	2.16%	4.76%	3.27%	1.40%	3.95%	2.74%	2.57%	5.51%	2.24%	8.01%	4.75%	0.39%
	Std	1.31%	1.69%	0.67%	3.03%	1.57%	4.40%	6.10%	0.60%	0.69%	0.39%	2.31%	13.66%	7.82%	0.18%
Austria	Mean	2.39%	1.79%	1.87%	3.17%	1.02%	1.38%	2.72%	0.87%	1.43%	4.92%	2.74%	4.26%	4.11%	0.49%
	Std	3.61%	2.56%	0.74%	4.32%	0.86%	3.04%	3.68%	0.62%	1.40%	0.61%	3.61%	6.03%	6.75%	0.47%
Belgium	Mean	-0.26%	1.44%	1.83%	3.26%	0.58%	1.26%	2.82%	0.83%	1.09%	7.60%	1.98%	2.62%	1.75%	0.79%
	Std	2.61%	2.10%	1.06%	3.81%	0.56%	3.47%	3.81%	0.87%	0.96%	0.98%	4.73%	6.85%	7.79%	0.77%
Canada	Mean	-2.89%	1.45%	1.72%	2.31%	1.55%	1.11%	2.45%	2.33%	1.90%	6.85%	2.03%	5.06%	4.43%	0.33%
	Std	0.67%	2.56%	0.67%	3.77%	1.18%	4.43%	4.27%	0.71%	1.08%	0.72%	2.93%	7.70%	5.26%	0.09%
Chile	Mean	0.00%	2.48%	3.12%	2.52%	3.73%	4.50%	4.46%	4.38%	3.47%	6.57%	1.62%	5.45%	6.61%	0.79%
	Std	0.01%	2.67%	1.13%	4.47%	1.96%	8.41%	7.98%	2.86%	2.12%	0.43%	4.73%	10.41%	9.21%	0.26%
Czech Rep	Mean	-0.78%	1.72%	1.68%	6.01%	1.55%	1.58%	5.49%	1.32%	1.81%	5.22%	4.04%	7.54%	6.39%	0.64%
	Std	3.30%	2.07%	1.01%	4.56%	1.97%	5.41%	4.95%	2.10%	2.24%	2.00%	4.74%	6.77%	8.05%	0.31%
Denmark	Mean	6.33%	0.49%	1.27%	1.89%	0.68%	0.28%	1.68%	0.99%	0.94%	6.49%	0.53%	4.28%	3.61%	0.33%
	Std	1.46%	2.28%	0.89%	3.91%	1.05%	5.61%	3.98%	1.36%	1.52%	0.97%	4.74%	5.82%	6.93%	0.32%
Estonia	Mean	1.72%	2.19%	2.34%	6.92%	1.90%	8.80%	8.41%	3.73%	3.18%	7.82%	4.87%	8.87%	8.97%	0.72%
	Std	3.91%	3.63%	1.93%	10.30%	4.92%	14.40%	10.23%	2.35%	2.35%	2.73%	8.66%	16.70%	14.93%	0.24%
Finland	Mean	0.03%	-0.16%	1.31%	0.56%	0.52%	0.48%	-0.64%	1.45%	0.92%	8.26%	0.53%	3.17%	1.72%	0.28%
	Std	2.53%	3.89%	1.11%	5.95%	1.18%	6.15%	5.79%	1.50%	2.35%	0.75%	4.63%	10.38%	10.12%	0.17%
France	Mean	-1.34%	0.31%	1.15%	3.46%	1.45%	0.49%	3.26%	0.89%	0.99%	9.95%	0.86%	4.04%	3.96%	0.63%
	Std	0.85%	2.62%	0.73%	3.53%	0.43%	2.78%	3.77%	0.68%	0.82%	0.65%	2.70%	5.49%	6.34%	0.49%
Germany	Mean	6.98%	2.12%	1.39%	4.67%	1.74%	2.22%	4.72%	1.12%	1.56%	4.88%	2.46%	5.50%	5.60%	0.30%
	Std	1.19%	4.13%	0.69%	5.12%	0.92%	3.73%	4.59%	0.80%	1.46%	1.19%	5.20%	6.62%	8.12%	0.24%
Hungary	Mean	2.54%	1.39%	2.09%	5.71%	0.43%	6.02%	5.71%	2.51%	2.44%	6.83%	3.67%	6.32%	6.76%	1.88%
	Std	1.87%	2.82%	1.96%	2.73%	2.15%	11.93%	3.26%	2.43%	2.12%	2.81%	3.73%	5.23%	5.16%	1.24%
Ireland	Mean	4.00%	4.52%	0.48%	6.66%	-0.23%	6.48%	5.89%	0.29%	3.01%	10.45%	3.97%	5.37%	0.63%	2.00%
	Std	7.99%	3.84%	1.32%	5.21%	3.94%	44.21%	12.32%	2.60%	4.03%	3.60%	10.73%	10.26%	11.04%	2.47%
Israel	Mean	0.59%	3.45%	0.95%	2.45%	3.52%	4.31%	4.68%	4.01%	3.35%	5.32%	1.94%	-2.00%	3.77%	0.96%
	Std	0.50%	0.80%	1.28%	3.91%	1.64%	7.95%	5.72%	1.44%	0.96%	1.02%	3.97%	9.10%	12.55%	0.41%
Italy	Mean	0.44%	-1.03%	0.99%	2.70%	-0.07%	-0.82%	1.36%	-0.20%	-0.08%	10.76%	-0.15%	5.37%	4.40%	2.02%
	Std	2.43%	3.49%	2.64%	4.40%	0.72%	4.70%	5.85%	1.73%	1.45%	1.52%	3.91%	6.36%	10.04%	1.08%
Japan	Mean	2.18%	1.27%	0.49%	4.44%	1.37%	1.04%	3.30%	0.66%	0.99%	3.63%	1.80%	4.77%	5.24%	0.54%
	Std	1.52%	1.99%	1.07%	9.10%	0.83%	3.96%	5.72%	1.64%	1.60%	0.89%	6.87%	10.81%	13.77%	0.30%
Korea	Mean	0.00%	3.96%	1.80%	4.78%	3.61%	2.48%	4.87%	2.45%	3.22%	3.56%	3.17%	9.19%	8.35%	0.72%
	Std	0.00%	2.89%	1.17%	4.70%	1.35%	4.88%	5.34%	1.06%	1.38%	0.35%	7.07%	15.40%	15.30%	0.33%
Mexico	Mean	-0.10%	2.35%	4.01%	6.81%	1.32%	1.65%	5.77%	2.68%	2.53%	4.44%	1.43%	11.87%	11.96%	1.21%
	Std	0.06%	2.50%	1.00%	7.43%	1.52%	4.45%	7.20%	1.86%	1.79%	0.78%	2.44%	10.25%	10.18%	0.27%
Netherlands	Mean	9.47%	1.41%	1.62%	3.88%	0.69%	1.63%	3.53%	0.28%	1.16%	5.27%	-0.14%	6.09%	5.56%	0.37%
	Std	1.99%	2.88%	0.85%	3.33%	1.01%	6.93%	3.74%	1.72%	1.70%	1.24%	4.34%	7.78%	8.47%	0.28%
New Zealand	Mean	-3.25%	2.37%	1.61%	2.98%	2.45%	3.80%	5.09%	2.97%	2.54%	5.71%	1.04%	3.55%	4.28%	0.42%
	Std	1.59%	2.56%	1.17%	2.51%	1.55%	4.59%	4.94%	1.14%	0.96%	0.96%	1.94%	10.78%	9.16%	0.23%
Norway	Mean	9.77%	1.63%	2.07%	0.05%	2.13%	1.75%	2.60%	2.33%	1.45%	3.74%	-1.22%	2.48%	4.43%	0.18%
	Std	3.46%	1.04%	0.85%	3.16%	0.90%	6.22%	4.35%	0.79%	1.37%	0.51%	5.01%	12.06%	10.53%	0.08%
Poland	Mean	-0.48%	4.14%	1.67%	6.12%	2.57%	3.23%	5.16%	2.90%	3.44%	7.64%	5.04%	5.39%	5.09%	0.99%
	Std	0.53%	2.41%	1.67%	3.50%	1.66%	5.53%	4.84%	1.50%	1.25%	2.52%	3.31%	5.92%	7.61%	0.55%
Portugal	Mean	-2.53%	0.23%	1.20%	4.99%	-0.48%	-1.86%	2.70%	0.22%	0.39%	11.78%	-0.13%	6.87%	4.10%	3.46%
	Std	4.55%	3.39%	1.26%	3.67%	2.46%	8.16%	6.04%	3.00%	2.02%	3.25%	3.34%	6.39%	7.36%	3.23%
Slovenia	Mean	6.29%	0.64%	1.21%	5.83%	0.11%	-0.41%	4.21%	0.49%	1.89%	7.95%	3.98%	7.49%	5.78%	1.63%
	Std	4.36%	2.10%	1.06%	3.30%	3.64%	8.12%	4.56%	2.61%	2.75%	1.83%	3.49%	6.22%	7.50%	1.12%
Spain	Mean	-0.57%	0.19%	1.29%	4.81%	0.00%	-1.26%	2.01%	0.94%	1.08%	20.81%	0.15%	4.69%	2.00%	1.67%
	Std	2.61%	2.83%	1.28%	3.60%	2.32%	6.40%	5.23%	2.29%	1.95%	3.94%	3.25%	7.00%	7.40%	1.32%
Sweden	Mean	5.39%	2.39%	1.11%	3.53%	1.52%	3.72%	3.50%	2.05%	2.66%	7.41%	2.30%	4.46%	4.73%	0.23%
	Std	1.66%	2.12%	1.05%	4.74%	0.76%	5.33%	5.16%	0.98%	1.96%	0.76%	5.37%	8.66%	8.11%	0.15%
Swiss	Mean	11.70%	2.73%	0.03%	2.13%	1.65%	1.99%	0.64%	1.45%	1.56%	4.40%	1.76%	4.51%	4.67%	0.29%
	Std	2.94%	4.10%	0.70%	4.22%	1.18%	1.96%	4.21%	0.61%	0.86%	0.38%	4.36%	4.69%	9.02%	0.16%
Turkey	Mean	-2.71%	4.05%	9.89%	5.47%	5.25%	4.51%	3.62%	3.64%	4.26%	10.46%	4.79%	17.45%	15.08%	2.32%
	Std	1.69%	3.59%	3.97%	6.37%	5.46%	13.47%	11.68%	4.65%	3.51%	1.65%	5.91%	18.28%	19.61%	0.82%
UK	Mean	-3.73%	0.58%	2.28%	1.99%	1.49%	1.15%	2.50%	1.29%	1.41%	6.16%	0.44%	5.10%	4.20%	0.40%
	Std	1.26%	2.81%	1.25%	3.62%	0.84%	4.59%	4.06%	1.47%	1.28%	1.63%	2.04%	7.87%	7.35%	0.22%
USA	Mean	-2.58%	2.40%	1.80%	3.85%	0.04%	3.05%	4.09%	2.38%	2.21%	6.38%	2.87%	5.86%	6.00%	0.35%
	Std	0.48%	2.26%	0.88%	3.51%	1.23%	3.06%	4.22%	0.52%	0.68%	2.08%	2.10%	8.54%	8.80%	0.18%

Table 3: **Out-of-sample R^2_{oos}**

This table reports the out-of-sample R^2_{oos} for each model class and each group of countries wihtin our sample. Daily forecasts are generated for each month based on a rolling one-month train/validation window. The R^2_{oos} each month is calculated by comparing daily, within-month, realised CDS spreads against daily model-implied CDS spread predictions. The top panel considers the whole sample from February 2011 to November 2019. The second and the third panels split the sample in two based on the economic policy uncertainty indicator proposed by Baker et al. (2016). The bottom panel breaks the temporal relationship between in-sample and out-of-sample on a given month by considering a given set of countries as train/validation and a different, non-overlapping set of countries as a testing sample. We color-code the out-of-sample performance, with green numbers suggesting a higher R^2_{oos} and red numbers a lower R^2_{oos}.

	Level only					Level + Slope					Non-linear models		
	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	GLM	Regression Tree	Random Forest
Global	0.505	0.507	0.508	0.454	0.427	0.584	0.587	0.589	0.558	0.436	0.517	0.590	0.813
Europe	0.380	0.388	0.401	0.455	0.423	0.469	0.481	0.493	0.561	0.435	0.632	0.563	0.791
Core Europe	0.486	0.498	0.520	0.521	0.506	0.614	0.614	0.617	0.619	0.528	0.346	0.559	0.721
High Risk Periods													
	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	GLM	Regression Tree	Random Forest
Global	0.520	0.522	0.522	0.486	0.491	0.537	0.539	0.539	0.507	0.497	0.451	0.660	0.885
Europe	0.365	0.377	0.380	0.584	0.551	0.379	0.397	0.398	0.599	0.558	0.604	0.825	0.890
Core Europe	0.322	0.338	0.349	0.466	0.461	0.362	0.381	0.392	0.550	0.465	0.109	0.669	0.774
Low Risk Periods													
	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	GLM	Regression Tree	Random Forest
Global	0.503	0.504	0.506	0.449	0.417	0.591	0.594	0.596	0.566	0.427	0.527	0.580	0.802
Europe	0.382	0.390	0.404	0.435	0.404	0.482	0.494	0.507	0.556	0.416	0.636	0.523	0.776
Core Europe	0.511	0.522	0.546	0.529	0.512	0.650	0.624	0.612	0.636	0.537	0.638	0.542	0.713
Pure OOS													
	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	OLS	Ridge	Lasso	ElasticNet	PCR (5 comp)	GLM	Regression Tree	Random Forest
EU (IS) vs Non-EU (OOS)	-0.924	-0.923	-0.907	-0.068	-0.335	-0.873	-0.872	-0.846	-0.114	-0.370	-0.573	-0.176	0.409
Non-EU (IS) vs EU (OOS)	-0.134	-0.134	-0.132	0.060	0.106	-0.137	-0.137	-0.135	0.039	0.104	0.051	0.161	0.341
Core EU (IS) vs Periphery EU (OOS)	0.328	0.318	0.319	0.281	0.327	0.318	0.324	0.305	0.314	0.343	0.080	0.266	0.460
Periphery EU (IS) vs Core EU (OOS)	-0.021	-0.019	-0.008	0.299	0.264	-0.053	-0.052	-0.037	0.318	0.310	-0.093	-0.378	0.434

Table 4: **Diebold-Mariano tests**

This table reports the results of a pairwise test of predictive accuracy as proposed by Diebold and Mariano (2002) (DM) to compare predictions from different models. We adjust the DM test by making a bias correction to the test statistic as proposed by Harvey et al. (1997). The table reports the significance of the performance gaps. We color-code the statistical significance of the test, with bold numbers suggesting a rejection of the null (H_0: the pairs have the same performance) at the standard 5% confidence level. The sample period is from February 2011 to November 2019.

Panel A: Global

		Level-only					Level + Slope					Non-linear		
	Models	OLS	Ridge	Lasso	E-Net	PCA (5)	OLS	Ridge	Lasso	E-Net	PCA (5)	GLM	Reg tree	Rand forest
	OLS		1.91	**2.38**	**2.09**	1.16	-1.03	1.83	**2.39**	1.99	1.16	**-2.05**	-0.17	**2.46**
	Ridge			0.51	**2.07**	1.13	**-2.07**	-0.06	1.06	1.97	1.13	**-2.07**	-0.20	**2.44**
Level-only	Lasso				**2.06**	1.12	**-2.50**	-0.51	0.90	1.95	1.12	**-2.07**	-0.21	**2.44**
	E-Net					-1.44	**-2.09**	**-2.05**	-2.03	**-2.14**	-1.44	**-2.69**	-1.83	1.93
	PCA (5)						1.17	1.12	1.10	-0.75	0.37	**-2.35**	-1.24	**2.39**
	OLS							2.11	**2.58**	1.99	1.17	**-2.04**	-0.16	**2.47**
	Ridge								1.53	1.95	1.12	**-2.06**	-0.20	**2.43**
Level+Slope	Lasso									1.92	1.10	**-2.07**	-0.22	**2.42**
	E-Net										-0.75	**-2.60**	-1.56	**2.04**
	PCA (5)											**-2.35**	-1.24	**2.39**
	GLM												1.67	**3.21**
Non-linear	Regr Tree													**4.41**
	Rand Forest													

Panel B: Europe

		Level-only					Level + Slope					Non-linear		
	Models	OLS	Ridge	Lasso	E-Net	PCA (5)	OLS	Ridge	Lasso	E-Net	PCA (5)	GLM	Reg tree	Rand forest
	OLS		0.97	**2.21**	**2.01**	1.82	-0.35	1.17	**2.48**	1.97	1.82	0.67	1.77	**2.08**
	Ridge			1.47	**2.05**	1.85	-0.98	**2.87**	1.92	**2.01**	1.85	0.65	1.79	**2.12**
Level-only	Lasso				1.99	1.79	**-2.22**	-0.93	0.94	1.95	1.79	0.61	1.74	**2.07**
	E-Net					-1.87	**-2.01**	**-2.04**	-1.98	**-2.50**	-1.86	-1.68	0.07	1.57
	PCA (5)						1.82	1.84	1.78	-0.78	1.13	-1.47	0.62	1.86
	OLS							1.18	**2.49**	1.97	1.82	0.67	1.77	**2.08**
	Ridge								1.34	**2.00**	1.84	0.64	1.78	**2.11**
Level+Slope	Lasso									1.93	1.78	0.60	1.73	**2.05**
	E-Net										-0.78	-1.54	0.41	1.83
	PCA (5)											-1.47	0.62	1.85
	GLM												1.46	1.96
Non-linear	Reg Tree													**2.22**
	Rand Forest													

Panel C: Core-Europe

		Level-only					Level + Slope					Non-linear		
	Models	OLS	Ridge	Lasso	E-Net	PCA (5)	OLS	Ridge	Lasso	E-Net	PCA (5)	GLM	Reg tree	Rand forest
	OLS		-0.37	1.90	**2.09**	1.64	-0.39	0.13	1.78	**2.01**	1.64	-0.38	1.33	**2.16**
	Ridge			1.51	**2.02**	1.59	0.36	0.94	1.52	1.94	1.59	-0.38	1.31	**2.10**
Level-only	Lasso				**2.04**	1.59	-1.91	-1.78	0.53	1.96	1.59	-0.43	1.28	**2.11**
	E-Net					-1.48	**-2.09**	**-2.04**	**-2.04**	-3.13	-1.48	-1.10	-0.72	**2.00**
	PCA (5)						1.64	1.60	1.59	-0.54	**2.06**	-0.94	-0.15	**2.61**
	OLS							0.15	1.80	**2.01**	1.64	-0.38	1.33	**2.16**
	Ridge								1.74	1.96	1.61	-0.39	1.31	**2.12**
Level+Slope	Lasso									1.96	1.59	-0.43	1.28	**2.12**
	E-Net										-0.54	-1.04	-0.35	**2.23**
	PCA (5)											-0.94	-0.15	**2.61**
	GLM												0.89	1.24
Non-linear	Reg Tree													**2.63**
	Rand Forest													

Table 5: **Importance of macroeconomic variables for non-linear forecasts**

This table reports the ranking of each macroeconomic variable in terms of its marginal relevance to the forecasting performance within the context of a random forest framework. We color-code the ranking of each macroeconomic variable, with green numbers suggesting a higher ranking and red numbers lower marginal relevance.

Panel A: Global

	Global - Average Rank (Normalised)								
Variable	2011	2012	2013	2014	2015	2016	2017	2018	2019
Harmonised Unemployment Rate	0.55	0.65	0.76	1.00	1.00	1.00	1.00	1.00	0.84
Composite Leading Indicator: Year On Year Growth Rate	0.86	0.61	0.55	0.39	0.58	0.77	0.65	0.00	0.45
Gdp: Government Consumption Expenditure, Constant Prices	0.48	0.00	1.00	0.13	0.39	0.58	0.23	0.62	0.14
Consumer Price Index	0.54	0.73	0.79	0.89	0.91	0.50	0.88	0.89	1.00
Gdp: Gross Fixed Capital Formation, Constant Prices	0.94	1.00	0.48	0.00	0.19	0.42	0.53	0.30	0.00
Gdp: Total, Constant Prices	0.51	0.30	0.20	0.28	0.21	0.35	0.36	0.56	0.97
Gdp: Exports Of Goods And Services, Constant Prices	0.46	0.22	0.48	0.21	0.41	0.34	0.39	0.30	0.28
Gdp: Private Consumption Expenditure, Constant Prices	0.61	0.26	0.49	0.09	0.00	0.29	0.44	0.25	0.42
Balance Of Payments - Current Account Balance (% of GDP)	1.00	0.39	0.39	0.59	0.46	0.23	0.36	0.36	0.17
Gdp: Imports Of Goods And Services, Constant Prices	0.20	0.07	0.32	0.28	0.08	0.21	0.48	0.37	0.17
International Trade In Goods - Exports	0.01	0.08	0.15	0.31	0.13	0.08	0.36	0.55	0.49
Index Of Industrial Production	0.43	0.03	0.00	0.52	0.04	0.05	0.19	0.31	0.32
International Trade In Goods - Imports	0.00	0.07	0.19	0.11	0.00	0.00	0.00	0.32	0.20

Panel B: Europe

	Europe - Average Rank (Normalised)								
Variable	2011	2012	2013	2014	2015	2016	2017	2018	2019
Harmonised Unemployment Rate	0.72	0.60	0.79	1.00	1.00	1.00	1.00	1.00	1.00
Composite Leading Indicator: Year On Year Growth Rate	0.51	0.32	0.79	0.61	0.70	0.94	0.89	0.25	0.21
Gdp: Private Consumption Expenditure, Constant Prices	0.38	0.30	0.91	0.26	0.40	0.68	0.40	0.16	0.01
Gdp: Total, Constant Prices	0.60	0.46	0.23	0.18	0.60	0.64	0.18	0.57	0.91
Consumer Price Index	0.72	0.56	0.76	0.49	0.21	0.48	0.29	0.11	0.22
Gdp: Exports Of Goods And Services, Constant Prices	0.43	0.04	0.55	0.38	0.50	0.44	0.49	0.20	0.21
Gdp: Imports Of Goods And Services, Constant Prices	0.00	0.32	0.51	0.64	0.21	0.38	0.78	0.35	0.04
Gdp: Gross Fixed Capital Formation, Constant Prices	0.91	1.00	0.53	0.36	0.00	0.34	0.48	0.30	0.06
Gdp: Government Consumption Expenditure, Constant Prices	0.64	0.00	1.00	0.26	0.80	0.33	0.40	0.42	0.25
Balance Of Payments - Current Account Balance (% of GDP)	1.00	0.14	0.44	0.30	0.28	0.30	0.28	0.24	0.40
International Trade In Goods - Imports	0.17	0.14	0.12	0.12	0.13	0.14	0.27	0.12	0.35
Index Of Industrial Production	0.74	0.00	0.00	0.16	0.31	0.05	0.24	0.00	0.00
International Trade In Goods - Exports	0.30	0.25	0.32	0.00	0.23	0.00	0.00	0.08	0.20

Panel C: Core-Europe

	Core Europe - Average Rank (Normalised)								
Variable	2011	2012	2013	2014	2015	2016	2017	2018	2019
Harmonised Unemployment Rate	0.35	0.81	1.00	1.00	0.74	0.49	0.31	0.43	0.61
Gdp: Imports Of Goods And Services, Constant Prices	0.29	0.41	0.31	0.78	0.36	0.37	0.50	0.54	0.56
Gdp: Gross Fixed Capital Formation, Constant Prices	0.71	1.00	0.30	0.76	0.50	0.49	0.91	0.30	0.63
Gdp: Total, Constant Prices	0.42	0.89	0.31	0.53	0.52	0.45	0.19	1.00	1.00
Gdp: Private Consumption Expenditure, Constant Prices	0.42	0.46	0.74	0.47	0.24	1.00	1.00	0.24	0.58
Gdp: Exports Of Goods And Services, Constant Prices	0.42	0.27	0.56	0.47	0.66	0.98	0.22	0.18	0.26
Gdp: Government Consumption Expenditure, Constant Prices	0.34	0.58	0.84	0.45	0.28	1.00	0.48	0.05	0.21
Composite Leading Indicator: Year On Year Growth Rate	0.34	0.15	1.00	0.45	1.00	0.57	0.48	0.26	0.16
Consumer Price Index	0.66	0.93	0.66	0.38	0.00	0.53	0.50	0.09	0.12
Index Of Industrial Production	0.43	0.00	0.00	0.31	0.96	0.24	0.00	0.25	0.11
International Trade In Goods - Imports	0.03	0.16	0.13	0.20	0.50	0.00	0.09	0.09	0.00
Balance Of Payments - Current Account Balance (% of GDP)	1.00	0.76	0.34	0.09	0.46	0.65	0.43	0.00	0.39
International Trade In Goods - Exports	0.00	0.26	0.22	0.00	0.02	0.12	0.09	0.16	0.16

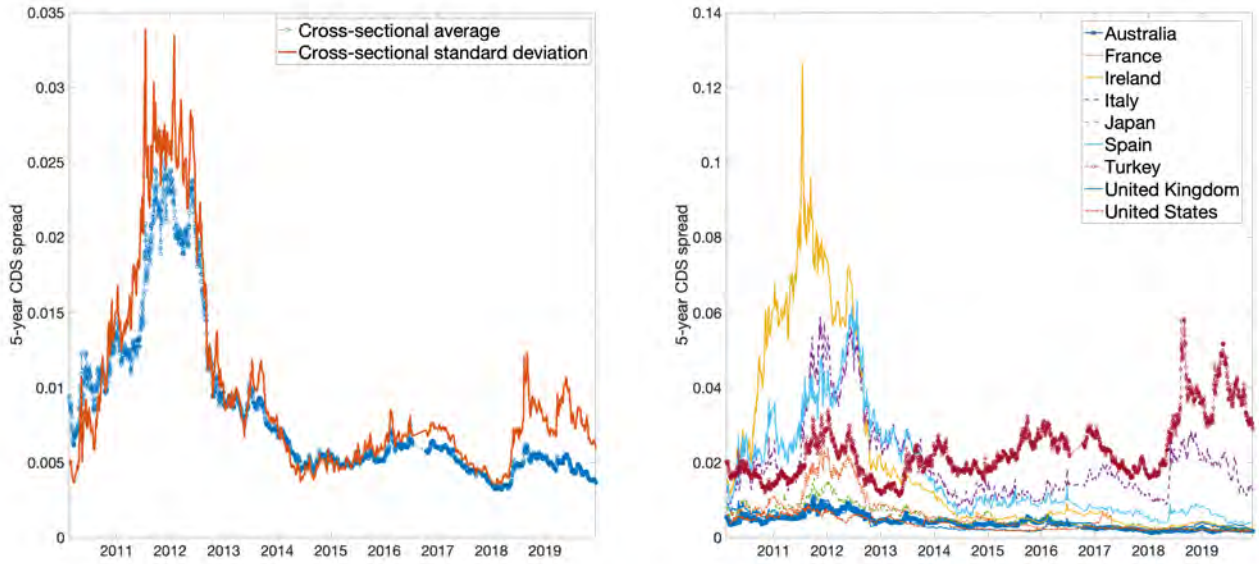
Table 6: **Expected CDS spreads, risk aversion and economic uncertainty.**

This table reports the results of a set of regressions in which the dependent variable is the monthly expected CDS spread averaged across countries and maturities and the independent variables are constituted by a set of economic policy and risk aversion indicators. We consider the Economic Policy Uncertainty (EPU) index as introduced by Baker et al. (2016). We focus in particular on the Sovereign Debt & Currency Crises "categorical" sub-index. We also consider the Geopolitical Risk Index (GRI) proposed by Caldara and Iacoviello (2018) and the Global Risk Aversion Index (GRAI) and the Global Uncertainty Index (GUI), both proposed by Bekaert et al. (2019). The left panel shows the results for the 2011-2019 sample in which tradable CDS contracts were available. The right panel shows the results for the 1992-2001 sample in which CDS contracts were unavailable. In this respect, shadow CDS spreads are used to check the external validity of our results.

Sample 2011-2019					Sample 1992-2001				
EPUI	GRI	GRAI	GUI	R^2	EPUI	GRI	GRAI	GUI	R^2
0.776***				0.837	0.189***				0.112
(0.099)					(0.065)				
	-0.635			0.485		-0.079			0.078
	(0.169)					(0.109)			
		0.481***		0.767			0.412**		0.153
		(0.117)					(0.219)		
			0.465	0.181				-0.078	0.074
			(0.269)					(0.156)	
0.695***		0.201***		0.842	0.187***		0.421**		0.175
(0.085)		(0.049)			(0.063)		(0.221)		

Figure 1: **A snapshot of CDS spreads**

This figure reports the cross-sectional average and standard deviation of 5-year CDS spreads (left panel) and the time series of 5-year CDS spreads (right panel) for a selection of OECD countries in our sample . Our daily sample spans the February 2011 to November 2019 period.

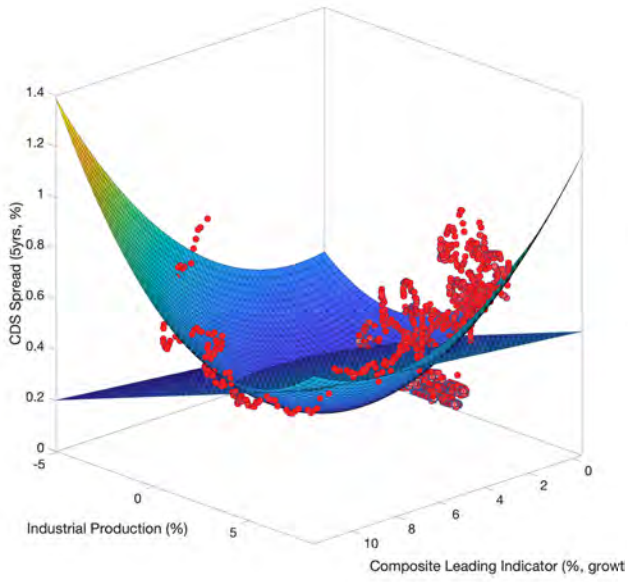


(a) Cross-sectional mean and volatility

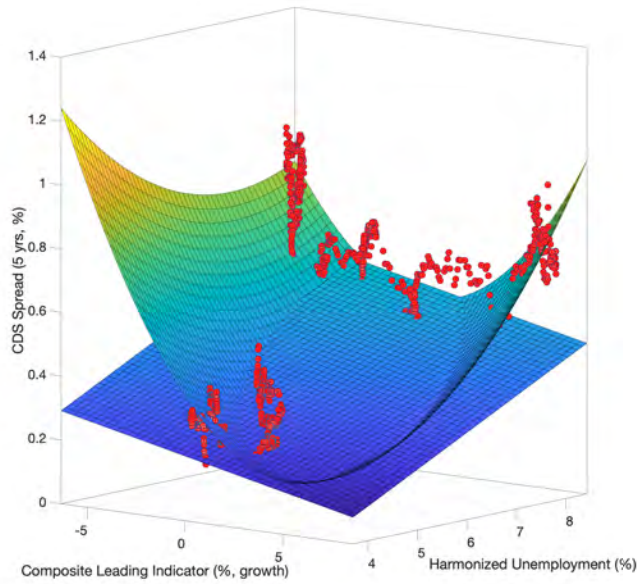
(b) Time-series dynamics

Figure 2: **A first look at non-linearity between sovereign CDS and macro variables**

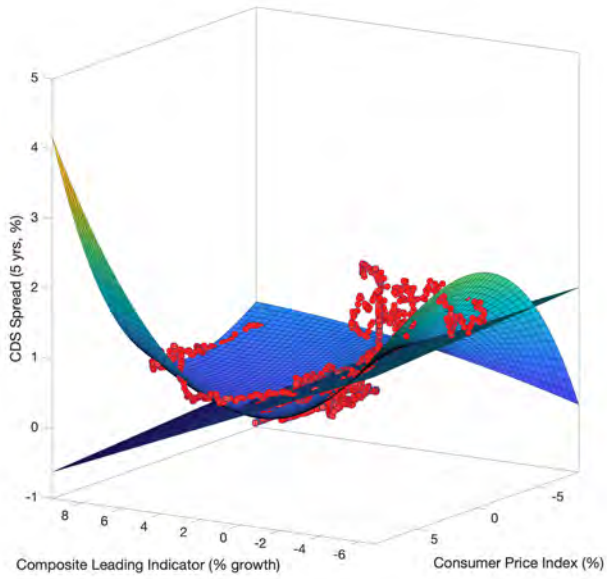
This figure reports the relationship between pairs of macroeconomic variables for a selection of countries and the associated 5-year CDS spreads. The flat surface represents the fitted values of a multiple linear regression whereas the curved surface represents the fitted values of a higher-order polynomial.



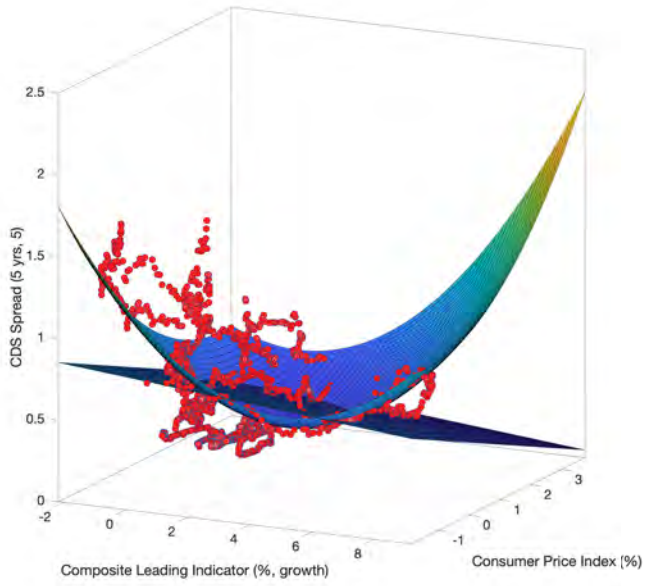
(a) United States



(b) United Kingdom



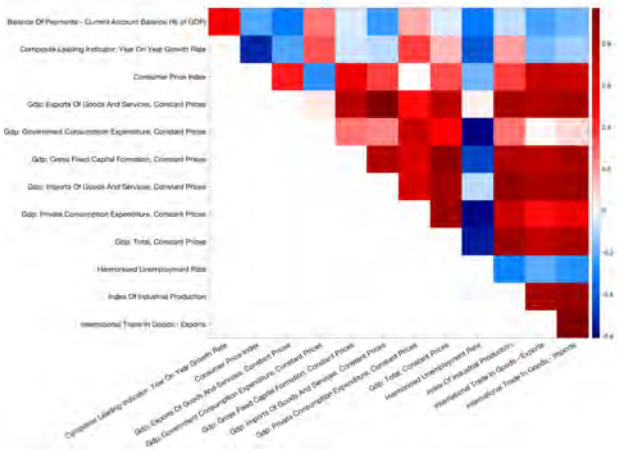
(c) France



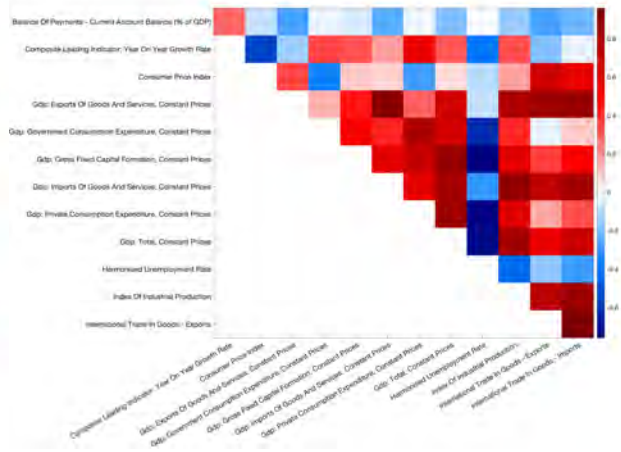
(d) Japan

Figure 3: **Correlation structure of the macroeconomic variables**

This figure reports cross-sectional correlation of macroeconomic variables for the average global country in our sample (left panel) and the average European country (right panel). The sample is from February 2011 to November 2019. We color-code the correlation coefficients, with a darker red (blue) color indicating more positive (negative) correlation.



(a) Global



(b) Europe

Figure 4: **Timeline of the forecasting framework**

This figure provides a sketch of the timeline for the train/validation and test split considering the daily cubic spline interpolation of monthly macroeconomic variables. The one-month train/validate, one-month test splitting is applied recursively so that the out-of-sample evaluation is performed on 100 monthly periods.

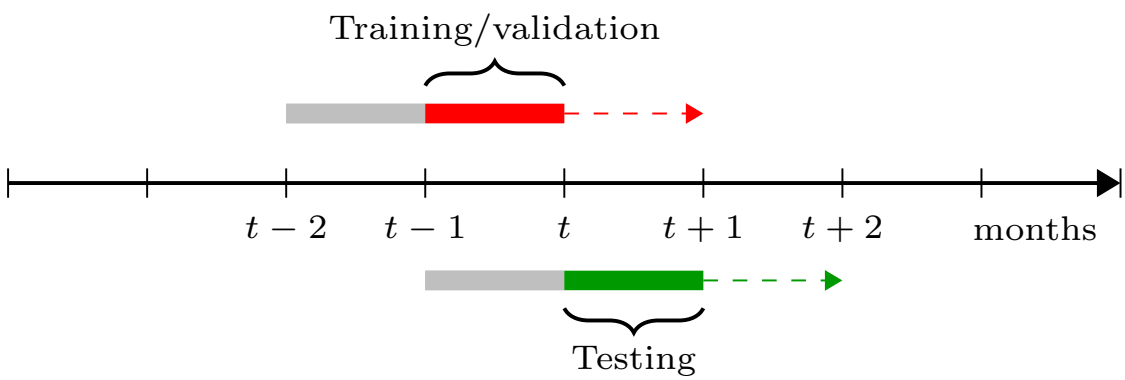
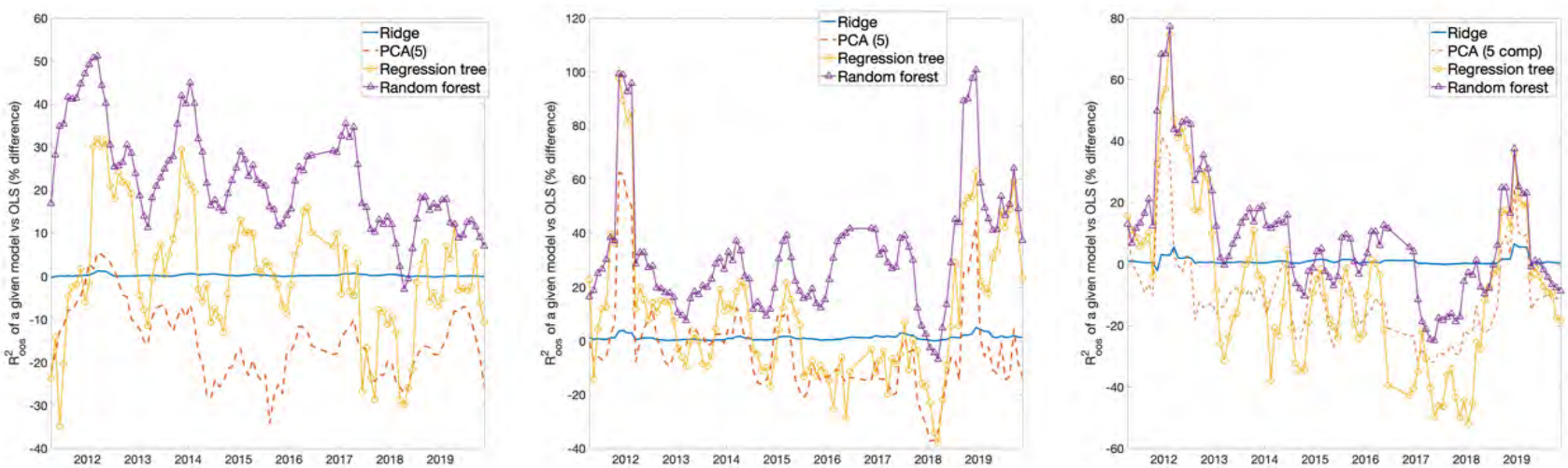


Figure 5: **Out-of-sample forecasting performance over time**

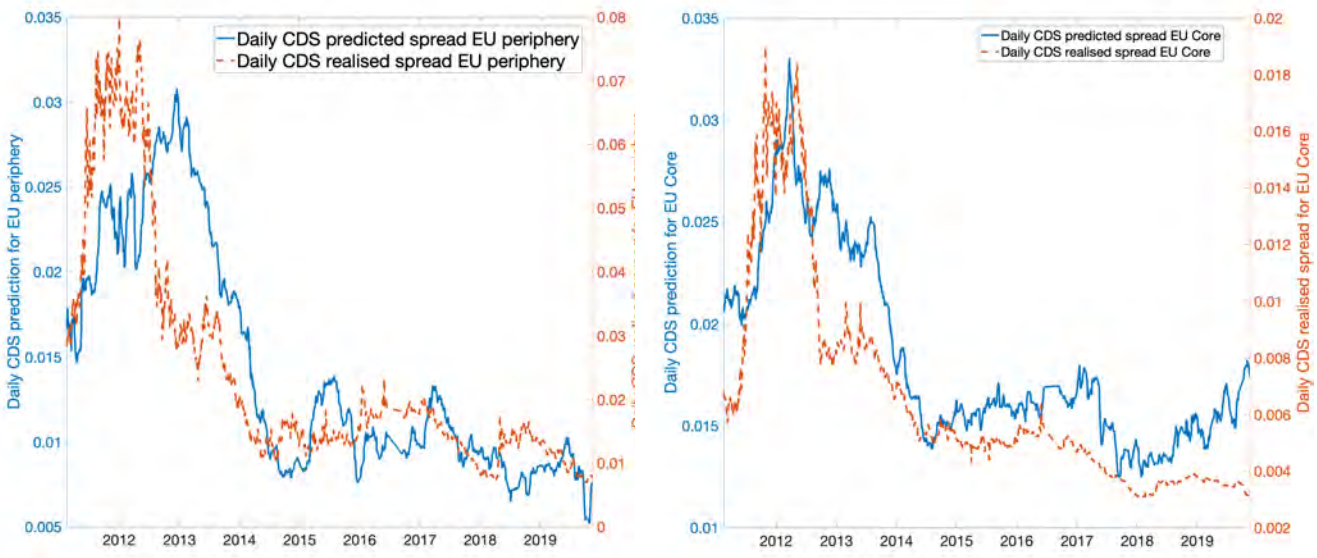
This figure reports the recursive out-of-sample differential of the $R^2_{t,oos}$ of each model against a simple OLS model, calculated for each of $t = 1, \ldots, 100$ months that are available as test samples. A positive value indicates that $R^2_{t,oos} > R^2_{t,oos}(OLS)$, with $R^2_{t,oos}(OLS)$ the out-of-sample fit obtained from the OLS predictive regression. For ease of exposition we report the results for the ridge regression, PCA (5 components), regression trees, and random forests.



(a) Global (b) Europe (c) Core Europe

Figure 6: **Core-EU vs peripheral EU CDS spreads**

This figure reports the actual vs predicted CDS spreads obtained from our random forest regressions. The left panel shows the forecast for peripheral EU countries obtained via a model trained and validated on Core-EU economies. The right panel shows the forecast for Core-EU economies obtained via a model trained/validated on peripheral EU countries. The forecasts $R^2_{t,oos}$ are calculated for each of $t = 1, \ldots, 100$ months that are available as test samples.

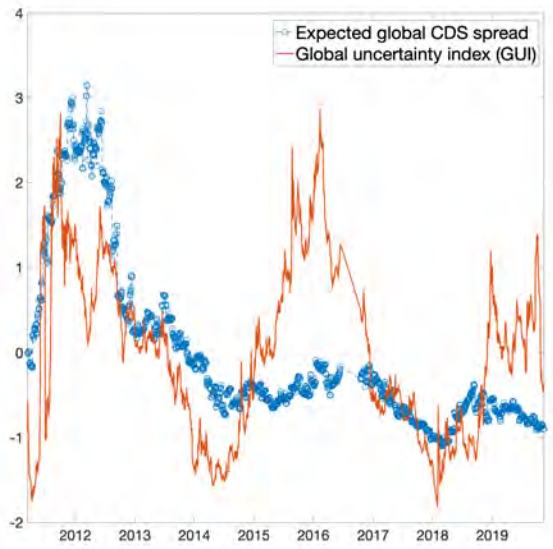


(a) Core (In-sample) vs Periphery (OOS)

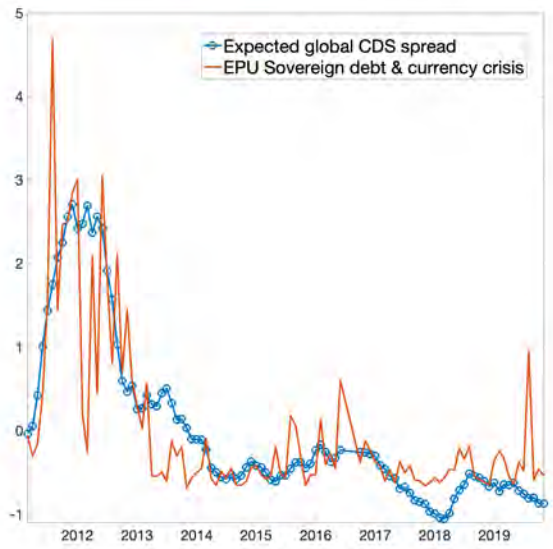
(b) Periphery (In-sample) vs Core (OOS)

Figure 7: **Expected CDS spreads and economic uncertainty**

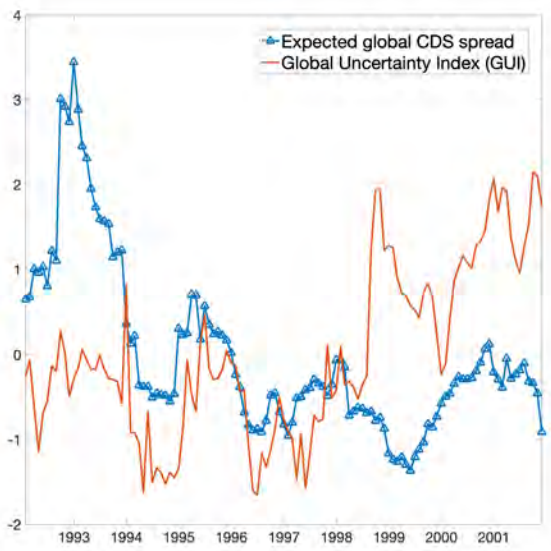
This figure reports the expected CDS spreads averaged across maturities and countries against the Economic Policy Uncertainty: Sovereign Debt & Currency Crises sub-index from Baker et al. (2016) and the Global Uncertainty Index (GUI) proposed by Bekaert et al. (2019). The top panels report the results for the sample from 2011 to 2019 where CDS contracts are available. The bottom panels report the results for "shadow" CDS spreads, that is by producing model-implied CDS spreads for the 1992-2001 period in which CDS contracts were not tradable.



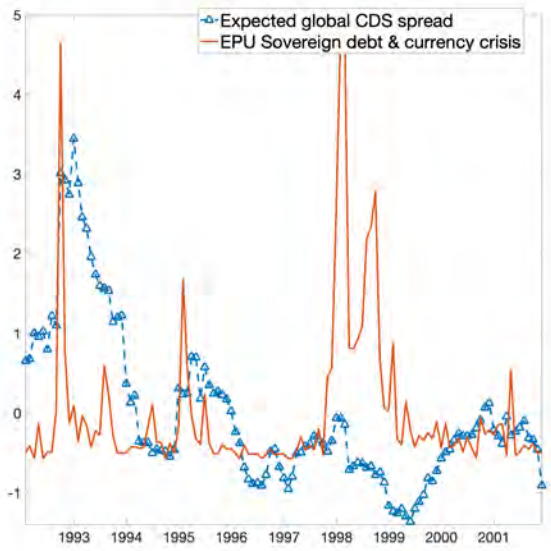
(a) Expected CDS vs GUI



(b) Expected CDS vs EPU Sovereign crisis



(c) Shadow CDS vs GUI



(d) Shadow CDS vs EPU sovereign crisis

A A simple motivating framework

Credit default swaps (CDS) are financial contracts between two parties over a specific time interval $[s, s+T]$ to exchange cashflows. The *protection buyer* of a CDS usually seeks compensation for the amount $L = 1 - R$, the expected loss as a fraction of the notional based on the recovery rate $R \in [0, 1)$. On the other hand, the *protection seller* charges the buyer a premium that, without loss of generality, can be thought of as being a continuously paid $S := \{S_t^*\}_{t \geq 0}$. Hence at the conclusion of a CDS contract at time $t \geq 0$, the two parties agree on both the date of expiry $t + T$, and the compensation payment $L \in (0, 1]$ when the reference issuer fails to meet her commitment.

Under these conditions, the *premium leg* of the CDS represents the positive cash flow for the seller and is defined as

$$PL\left(S_t^*; t, t+T\right) = \mathbb{I}_{\{\tau > t\}} \int_t^{t+T} E\left[S_t^* e^{-\int_t^u r_s + \lambda_s ds} | \mathcal{F}_s\right] du \tag{A.1}$$

where $\mathbb{I}_{\{\tau > t\}}$ is the default indicator process, $\mathcal{F}_s$ is the filtration generated by the history of the macroeconomic indicators $\mathcal{F}_s = \{\boldsymbol{x}_{s-i}\}_{i \geq 0}$, and r_s and λ_s are the risk-free discount rate and the default intensity rate at time s, respectively. Similarly, the *protection leg* of the CDS represents the cash flow obtained by the protection buyer in case of default and can be defined as

$$DL\left(S_t^*; t, t+T\right) = L\mathbb{I}_{\{\tau > t\}} \int_t^{t+T} E\left[e^{-\int_t^u r_s + \lambda_s ds} \lambda_u | \mathcal{F}_s\right] du \tag{A.2}$$

Then, under standard no-arbitrage conditions (see, e.g., Augustin and Tédongap, 2016 and the references therein), the CDS premium is called "fair" if the premium and protection legs are equal in terms of present value, that is $PL\left(S_t^*; t, t+T\right) = DL\left(S_t^*; t, t+T\right)$, i.e.,

$$\mathbb{I}_{\{\tau > t\}} \int_t^{t+T} E\left[S_t^* e^{-\int_t^u r_s + \lambda_s ds} | \mathcal{F}_s\right] du = L\mathbb{I}_{\{\tau > t\}} \int_t^{t+T} E\left[e^{-\int_t^u r_s + \lambda_s ds} \lambda_u | \mathcal{F}_s\right] du \tag{A.3}$$

The equality yields that

$$S_t^* = L\lambda_t, \qquad t \geq 0 \tag{A.4}$$

Assuming the intensity rate λ_t is linear in the parameters such that $\log \lambda_t = \boldsymbol{\beta}' \boldsymbol{x}_{t-1}$ (see, e.g., Gourieroux et al., 2006), the log of the CDS spread s_t can be defined as

$$s_t = \text{constant} + \lambda_t \propto \boldsymbol{\beta}' \boldsymbol{x}_{t-1}, \tag{A.5}$$

as in Eq.1 in the main text. Here the constant is the log of $L = 1 - R$. In the main empirical analysis we depart from Eq.1 and assume that the mapping between macroeconomic indicators and the CDS spread is unknown a priori (cfr. Eq.2).

B Interpolating macroeconomic variables

The main empirical results are based on the assumption that daily macroeconomic interpolated data are extracted using the full sample of observations. The first question that could arise is how good the interpolation may look compared to the original data. Figure B.1 shows this case in point. We report few examples of original monthly (red circles) and daily interpolated (solid blue lines) data. The top panels report respectively the monthly harmonised unemployment rate for the UK (left panel) and the private expenditures for France (right panel). The accuracy of the cubic spline interpolation is quite remarkable. The same is confirmed for the bottom panels, which report the cpi for Spain (left panel) abd the industrial production for Canada (right panel). Again, the accuracy of the approximation seems quite remarkable.

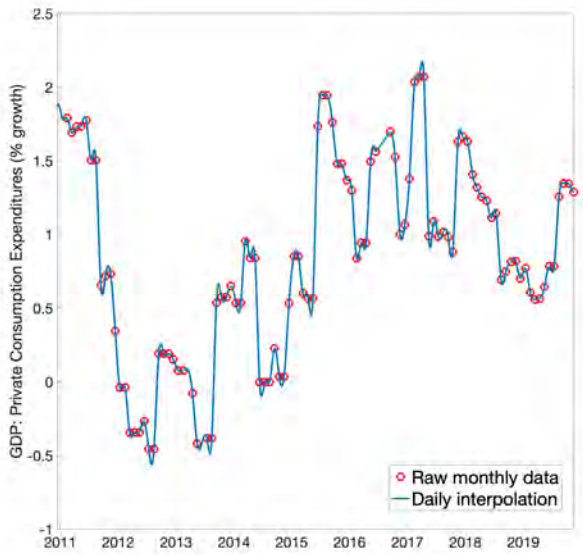
Although there is a trade-off between the length of the time series used to approximate a more finer grid of observations and the precision of such approximation, this potentially creates a form of look-ahead bias. In order to address this issue, we replicate the main forecasting results adopting an "expanding" window approach to cubic spline interpolation. In this approach, interpolation is carried out using data up to the closest future month to a particular day we require interpolated values for. For instance, if we require interpolated daily data for 18th May 2012, we utilise data from the start date of our dataset (Jan 2011) until the closest future month to 18th May 2012 (in this case June 2012). In this fashion, the data set utilised for interpolation expands in line with the time period we require interpolated data for. While this still leaves on the table a potential 15 days look ahead bias, our goal of investigating the robustness of the results when the look-ahead period is significantly reduced remain intact. Table B.1 reports the replication of some of the main results in the paper by using such expanding window approach.

The results suggest that by using an interpolation scheme which is by construction much less prone to a look-ahead bias the results remain largely unaffected. That is, there is only a minimal variation in the out-of-sample performance of our Random Forest model. In fact, we observe marginal performance *improvement* when using such expanding interpolation window approach. We conclude that the overall effects from look-ahead bias via cubic spline interpolation on all available data for a given country-variable pair are minimal. As a result, we opt to use a daily macroeconomic data set generated via cubic spline interpolation on all available data to increase the precision of the approximation (see Figure B.1).

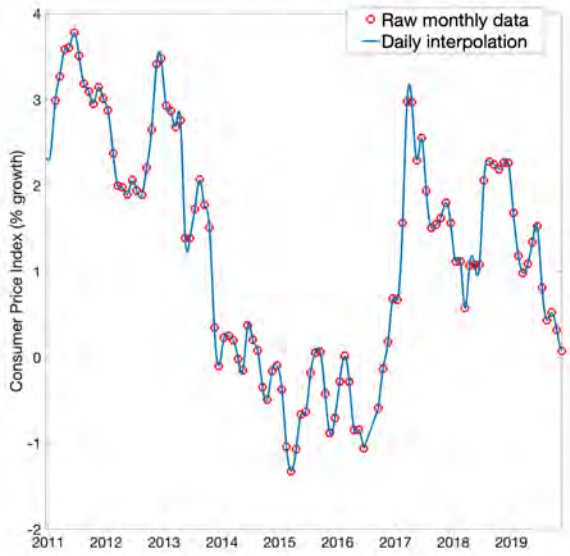
Figure B.1: **Raw monthly vs interpolated daily macroeconomic variables**

This figure reports the original monthly data (red circles) and the daily interpolated macroeconomic data (solid lines), for a variety of countries and macroeconomic indicators. The sample is from February 2011 to November 2019.

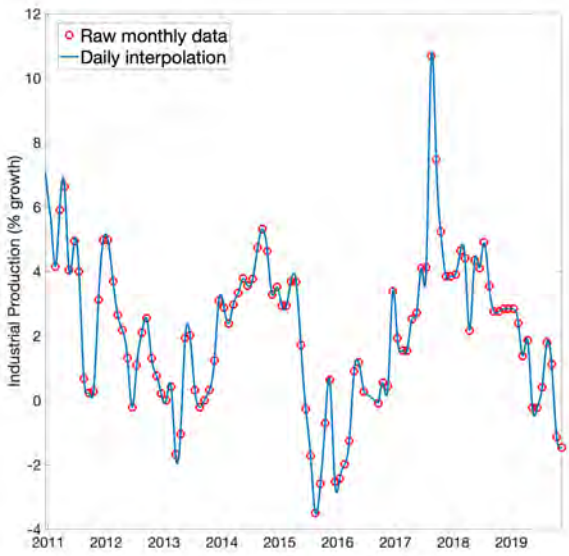
(a) Unemployment UK



(b) GDP: Private Expenditures France



(c) CPI Spain



(d) Industrial Prod Canada

Table B.1: Average R^2_{OOS} - All Data Interpolation vs. Expanding Window Interpolation

	Full Sample		High Risk Period		Low Risk Period	
	All Data Interpolation	Expanding Window Interpolation	All Data Interpolation	Expanding Window Interpolation	All Data Interpolation	Expanding Window Interpolation
Global	81.25%	82.12%	88.55%	89.51%	80.16%	81.02%
Europe	79.11%	79.99%	88.99%	90.18%	77.64%	78.46%
Core Europe	72.11%	73.76%	77.43%	79.75%	71.32%	72.86%

C Algorithmic details

C.1 Penalised Regression

Echoing our earlier suggestion, a simple way of mitigating the effect of statistically insignificant variables is to introduce the concept of sparsity into a simple OLS model via an additive penalty term $\lambda \mathbf{P}(\boldsymbol{\beta})$, where $\mathbf{P}(\boldsymbol{\beta})$ depends on the choice of penalty model. λ captures the degree of shrinkage and is a hyperparameter whose value is optimised via cross-validation during our estimation strategy, while $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ represents our explanatory variable coefficients. This penalty term enters the familiar simple OLS loss function $\mathcal{L}(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^T)$

$$\mathbf{P}(\boldsymbol{\beta}) = \begin{cases} \sum_{j=1}^{p} \beta_j^2 & \text{Ridge} \\ \sum_{j=1}^{p} |\beta_j| & \text{Lasso} \\ \mu \sum_{j=1}^{p} \beta_j^2 + \frac{(1-\mu)}{2} \sum_{j=1}^{p} |\beta_j| & \text{Elastic-net} \end{cases}$$

Apart from the estimation of $\boldsymbol{\theta}$ and the shrinkage parameter λ, the elastic-net requires a regularisation parameter μ - otherwise known as the $\mathcal{L}1$ ratio. Both λ and μ are chosen from a suitable range of values by evaluating the pseudo out-of-sample performance of the model on a validation sample.

Due to the presence of an $\mathcal{L}1$ regularisation term, closed-form solutions for $\boldsymbol{\theta}$ cannot be explicitly obtained for the Lasso and Elastic-net models. As a result, we estimate $\boldsymbol{\theta}$ by means of coordinate descent proposed by Wu et al. (2008) and extended by Friedman et al. (2010).

Algorithm 1: Coordinate Descent

Choose initial estimates for $\hat{\alpha} = \bar{y}$ and $\boldsymbol{\beta}^{(0)}$ for given λ and μ, where $\bar{y}$ is the unconditional mean of y.

Standardize the inputs x_{ij} such that $\sum_{i=1}^{N} x_{ij} = 0$, $\frac{1}{N}\sum_{i=1}^{N} x_{ij}^2 = 1$, for $j = 1, ..., p$.

Set ϵ to desired convergence threshold

while *there is an improvement in the loss function, i.e.* $|\mathcal{L}(\boldsymbol{\theta})^{(k+1)} - \mathcal{L}(\boldsymbol{\theta})^{(k)}| > \epsilon$ **do**

  for *all predictors j = 1,...,p* **do**

    $\hat{y}_i^{(j)} = \hat{\alpha} + \sum_{l \neq j} x_{il}\hat{\beta}_l$, i.e. the fitted value when omitting the covariate x_{ij}

    $\hat{\beta}_j \leftarrow \frac{S(\frac{1}{N}\sum_{i=1}^{N} x_{ij}(y_i - \hat{y}_i^{(j)}), \lambda\mu)}{1+(1-\mu)}$ defines the parameter-wise update, where S, the soft-thresholding operator, is given by $S(a,b) = \begin{cases} a-b, \text{if } a > 0 \vee b < |a| \\ a+b, \text{if } a < 0 \vee b < |a| \\ 0, b \geq a \end{cases}$

  end

end

Result: Estimates $\hat{\beta}$ for given level of λ, μ

Based on it's reliance on a $\mathcal{L}2$ regularisation term, closed-form solutions for $\boldsymbol{\beta}$ exist in a Ridge regression setting and are of the form

$$\hat{\beta}_{Ridge} = (\mathbf{X}'\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}'\mathbf{y} \tag{C.1}$$

where $\mathbf{X}$ is a $N \times p$ matrix of regressors, $\mathbf{I}$ is an $N \times N$ identity matrix and $\mathbf{y}$ is our vector of dependent sovereign CDS spreads. As with Lasso/Elastic-net, the shrinkage parameter λ is chosen by cross-validation.

C.2 Regression Trees

First proposed by Breiman et al. (1984), a regression tree is a hierarchically organized structure with each node splitting the data space into partitions based on value of a particular feature. This is equivalent to a partition of $\mathbb{R}^d$ into K disjoint feature sub-spaces $\{\mathcal{R}_1, ..., \mathcal{R}_k\}$, where each $\mathcal{R}_j \subset \mathbb{R}^d$. On each feature subspace $\mathcal{R}_j$ the same decision/prediction is made for all

$x \in \mathcal{R}_j$.

Algorithm 2: Regression Tree

Initialise tree $T(D)$ where D denotes the depth; denote by $R_l(d)$ the covariates in branch l at depth d.

for d = *1,...,D* **do**

  for $\tilde{R}$ *in* $\{R_l(d), l = 1, ..., 2^{d-1}\}$ **do**

    Given splitting variable j and split point s, define regions

$$R_{left}(j,s) = \{X|X_j \leq s, X_j \cap \tilde{R}\} \text{ and } R_{right}(j,s) = \{X|X_j > s, X_j \cap \tilde{R}\}$$

    In the splitting regions set

$$c_{left}(j,s) \leftarrow \frac{1}{|R_{left(j,s)}|} \sum_{x_i \in R_{left(j,s)}} y_i(x_i) \text{ and } c_{right}(j,s) \leftarrow \frac{1}{|R_{right(j,s)}|} \sum_{x_i \in R_{right(j,s)}} y_i(x_i)$$

    Find j, s that optimize

$$j,\ s = \underset{j,s}{\text{argmin}}[\sum_{x_i \in R_{left(j,s)}} (y_i - c_{left}(j,s)^2) + \sum_{x_i \in R_{right(j,s)}} (y_i - c_{right}(j,s)^2)]$$

    Set the new partitions

$$R_{2l}(d) \leftarrow R_{right}(j,s) \text{ and } R_{2l-1}(d) \leftarrow R_{left}(j,s)$$

  end

end

Result: A fully grown regression tree T of depth D. The output is given by

$$f(x_i) = \sum_{k=1}^{2^L} \text{avg } (y_i|x_i \in R_k(D))\mathbb{1}_{x_i \in R_k(D)}$$

i.e. the average response in each region R_k at depth D.

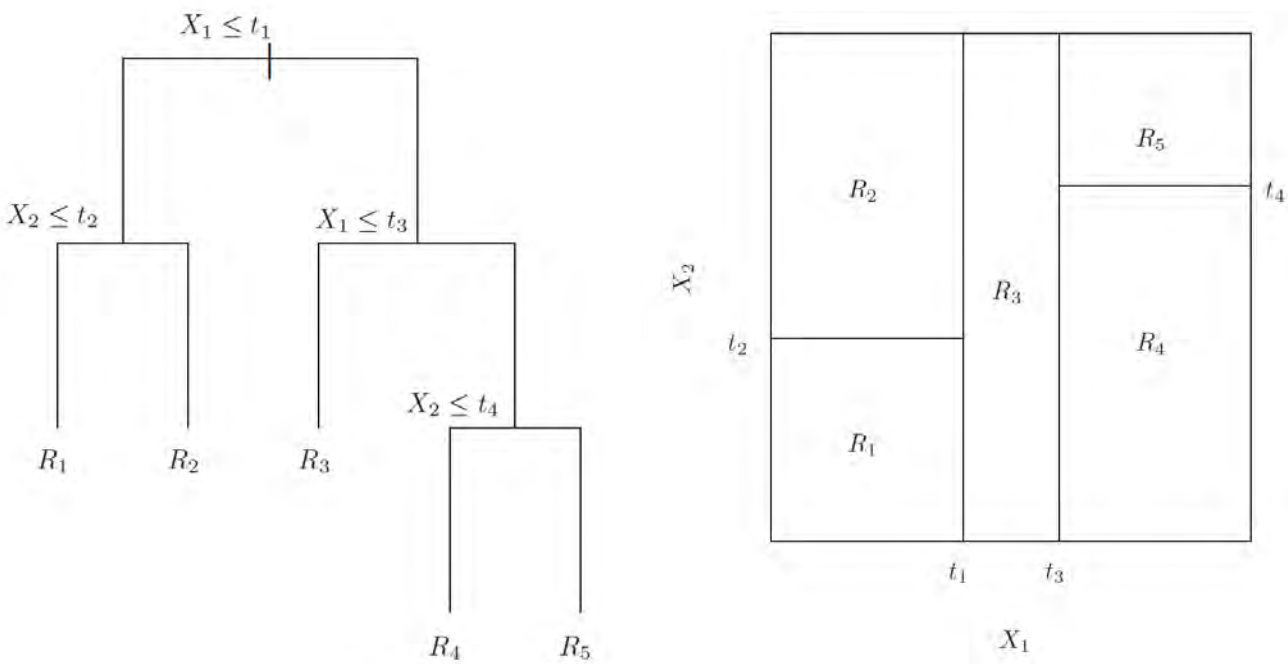
Ideally, would like to find partition that achieves minimal risk, i.e. the lowest mean-squared error for a regression problem. Given the number of potential partitions is too large to search exhaustively, greedy search heuristics must be used to determine the optimal partition - starting at the root node, we evaluate the loss for splitting on all combinations of features j and and split points s. The optimal pair $(j^{\cdot}s^{)}$ determines the members of each child node. Finally, we recurse on all child nodes iteratively until some stopping criterion is met.

Tree complexity needs to be regularised in order to prevent over-fitting, and more generally

find the tree size/structure that delivers optimum predictive performance. As a result, we adopt several pruning rules to manage tree complexity. We initially cross-validate the minimum number of samples required at a particular node in order to evaluate further split points. The smaller this sample number, the greater the tree depth and hence the greater is model complexity. Secondly, we cross-validate the minimum number of samples required at each leaf node. If a split point is determined and the number of samples present within a resultant leaf node is less than the stated minimum, the split is not executed.

Figure C.1: **Regression tree**

This figure reports...



(a) Recursive binary splitting

(b) Partition tree

C.3 Random Forest

While regression trees offer non-parametric, supervised non-linear framework for modelling, they are often prone to overfitting training data - i.e. they record low bias and high variance Mitchell et al. (1997). Random forests utilise an ensemble approach, combining the output of multiple decision trees in a bootstrap-aggregation ("bagging") format. Bagging relies on the theory that larger numbers of weak learners perform better in aggregation relative to small numbers of more complex learners. While the hyperparameters for individual trees are similar in both regression tree and random forest model structures, random forests incorporate additional randomness at the tree-level; rather than searching through all features when evaluating node split points, the algorithm searches for the best feature among a random subset of

features. The resultant individual trees display lower correlation and hence offer more power when used in an ensemble format. As a result, an additional hyperparameter we tune in our random forest model is the number of features to randomly select when evaluating split points for individual trees. Due to the computational load associated with tuning large numbers of hyperparameter, we opt to fix the number of trees in our random forest at 100.

Algorithm 3: Random Forest

Determine forest size F

for t = *1,...,F* **do**

Obtain bootstrap sample Z from original data.

Grow full trees following Algorithm (2) with the following adjustments:

1. Select $\tilde{p}$ variables from the original set of p variables.
2. Choose the best combination (j, s) (c.f. Algorithm (2)) from $\tilde{p}$ variables
3. Create the two daughter nodes

Denote the obtained tree by T_t

end

Result: Ensemble of F many trees. The output is the average over the trees in the forest given as

$$f(x_i) = \frac{1}{F} \sum_{t=1}^{F} T_t(x_i)$$

D Computational Details

Our machine learning library of choice is the popular `scikit-learn` package used within a Python 3 programming framework. We use `pandas` for data manipulation and `numpy` for mathematical operators. Our regression package of choice is the `statsmodels` API. For data pre-processing, we utilise the `StandardScaler` class from the `scikit-learn` package, as well as making use of the `Pipeline` feature to prevent data leakage between test/train datasets. In order to efficiently optimise hyperparameter values we utilise the `GridSearchCV` class within `scikit-learn`.

D.1 Setup

While cubic spline interpolation results in daily sovereign CDS data over a 10-year period for 29 countries across 4 CDS maturities, the small size our of rolling train/validation and

test windows coupled with the relatively low complexity associated with training regression tree/random forest algorithms implies powerful hardware instances (such as the high-performance GPU computing capabilities offered via Amazon Web Services) are not required. All work was carried out on a single 2.60 GHz, 16GB RAM node with 6 cores.