Machine Learning meets Econometrics: The British Macroeconomy as a Random Forest

Term paper for the

Seminar *S510/S520 – Master Seminar on Econometrics*

of Prof. Dr. Martin Biewen and Prof. Dr. Joachim Grammig

Supervisor: Dr. Jantje Sönksen

Nico Schwarzer

Lange Gasse 3

72070 Tuebingen

Student-ID: 5632297

M.Sc. Data Science in Business and Economics

Table of Contents

1 Introduction **ii**

2 Related Work **iii**

3 Methodology **vi**

3.1 The Macroeconomic Random Forest vi

3.1.1 Trees in MRF explained vi

3.1.2 From Trees to a Forest viii

3.1.3 Engineering MRF Input Data ix

3.2 The Dataset x

4 Analysis **xi**

4.1 Using MRF for Forecasting xi

4.2 Buidling explainable surrogate trees xiii

4.2.1 Unemployment Rate xv

4.2.2 CPI Inflation xviii

4.2.3 1 Year GVT Bond Yield xx

4.2.4 SPREAD xxii

4.2.5 GDP xxiii

5 Conclusions **xxv**

Appendices **xxx**

A Further Parameter Comparison Graphs **xxx**

B Bayesian Block Bootstrap and Bayesian Inference **xxxii**

C The Parameter Stability test **xxxiii**

D The Dataset used **xxxv**

D.1 ADF-tests for new variables xxxv

D.2 Overview over all used series xxxv

1 Introduction

The rise of (supervised) Machine Learning (ML) methods in quantitative research has lead many econometricians to adapt the latter as a part of their statistical toolbox (Athey, 2019). However, the focus of such methods differs to that of classical econometric models. Supervised ML algorithms aim at predicting outcomes based on input data. This is achieved through learning a complex, yet generalize-able pattern that relates the potentially high dimensional input data x to an outcome variable y (Mullainathan & Spiess, 2017). Such patterns often do not disclose a meaningful relationship of the input variables and the outcome variables. This phenomenon of a loss of explainability is at times alluded to as the *black box* character of ML methods (Buckmann et al., 2021). Classical econometric models, on the other hand, aim at the estimation of parameters β that (linearly) relate input variables x to dependent variables y (Athey, 2019). Such parameters can be interpreted in a straightforward fashion and can reveal insights about the direct and sometimes causal relationship of the variables (Athey, 2019). However, the structural form of models used for parameter estimation restricts the models' flexibility and can thus hurt the models ability to capture the potentially complex and highly non-linear relationship of x and y. This often also leads to worse predictive performances on unseen data when compared with ML methods (Buckmann et al., 2021). This problem is especially pronounced when analysing Macroeconomic Time-Series data as the relationships under consideration may evolve over time or be heavily influenced by other economic variables (e.g. due to Economic crisis, external shocks etc.) (Buckmann et al., 2021). This drawback can be tackled to some degree by time-varying parameters β_t *TVPs*, i.e. the ability to allow parameters in model to change over time depending on the time itself or other external factors (Coulombe, 2020).

Against this background of the all mentioned above, analytical frameworks which incorporate ML algorithms and nonetheless allow for interpretable inference began to surface (Athey, 2019). One of the just mentioned approaches with a specific focus on Time Series settings is the *Macroeconomic Random Forest* (MRF) which was introduced by Coulombe (2020). MRF allows for modelling varying linear parameters very flexibly leading to an overall non-linear relation of the input features to the output variables. Hence, the associated parameters can be deemed *Generalized Time-Varying Parameters* (GTVPs). Upon the introduction of this novel framework, Coulombe (2020) argues that MRF can provide gains over classical methods from either Time Series econometrics and classical ML as they allow for more flexible modelling and forecasting while retaining explainable, linear relationship between economic variables depending on the economic state at hand. This thesis is demonstrated by contrasting the predictive performance of different versions of the MRF against that of classical Time Series models and classical Random Forests using Macroeconomic data from the US. In a second step, Coulombe (2020) illustrates how MRF can give rise to an explainable framework for economic analysis: Based on the most important economic variables driving the changing nature of the GTVPs (derived from MRF), simpler surrogate trees are built using only those variables. The latter are shown to serve as directly interpretable tools which enable econometric inference. Coulombe (2020) finds significant forecasting gains of MRF compared to plain Random Forests and Classical Time Series models for different Macroeconomic variables. The surrogate trees relying on few variables result in well fitting and explainable models - showcasing that some variables are vastly more important that others concerning the influence on the mutating behaviour of the parameters. Upon analyzing the nodes of these trees, Coulombe (2020) is able to deduce those variables and conduct econometric inference from the trees.

In order to acquire a more profound insight into the alleged advantages of the MRF method, the experiments laid out above shall be replicated for the British Macroeconomy in this paper. Firstly, the MRF method is described in greater detail and an overview of the data utilized is given. Following Coulombe (2020), both a comparison of the predictive performance of the MRF against other methods and an analysis of the changing parameters as well as the economic setting influencing the parameters shall be conducted. A summary and future research opportunities are presented it the end.

2 Related Work

The evolving nature of (Macro-) economic relationships have created a need for time series models with *TVP* (D'Agostino et al., 2013). Based on the capability of the latter to better capture economic relationships with time dependencies, such models also commonly deliver more accurate forecasts (Durlauf & Blume, 2016, p.175). One type of non-linear time series approaches model the linear relationships differently depending on pre-determined joint points, which are also referred to as *structural breaks* (Tucci, 1995). *Switching regime models with a known join point* fit different linear models on subsets of the data. For all data up to joint point t^* and for all data with $t > t^*$, separate linear models are estimated. Researchers applying the related concept of *treshold autoregressions* decide on certain values of pre-determined exogenous variables to attain subsets (Durlauf & Blume, 2016, p.172). *Markov switch models* fit different linear equations depending on different latent states. Based on a learned transition matrix, switches between the states are possible (Durlauf & Blume, 2016, p. 172). However, the structural form of models relying on set joint points restrict the ways in which β_t might evolve: When setting a certain joint point t^*, only two regimes can be learned (Coulombe, 2020). Also, the economic interpretation of the model is tied to the type of joint point chosen. Markov switch models do not require set joint points, but the number of states and associated initial probabilities per state must be defined (Neusser, 2016, p. 366f.). Also, the fact that the states are latent introduces difficulties in the economic interpretation.

Another approach to model varying parameters are tree-based models, such as the trees used in MRF. Indeed, several precursors to MRF can be identified. These build on the idea of modelling the parameters in linear equations differently depending on the values on exogenous features, which a tree-based structure allows for (Potts & Sammut, 2005). Economically speaking, the parameters in linear relationships might depend on the economic setting, which in turn is characterized by the values that economically relevant variables take on (Coulombe, 2020). Many of such models have been proposed with a newer generation not merely having a parametric models in final leaves but also letting a model decide on optimal values of the variables to split by based on linear models in each coordinate split (Zeileis et al., 2008). In contrast to *ordinary* regression trees, such *Linear Tree* models do not aim at finding an optimal intercept in each child leave that minimizes the MSE of itself with the values of the outcome variable depending on some split point and variable, but follow the coordinate splitting procedure laid out in equation 1 (Coulombe, 2020).

[1]

[1]Similar splitting rules have been proposed by others researchers, e.g. by Potts & Sammut (2005). However, this notation fits well to the notation used in the MRF trees later. Note that X_t might be (and in the following cases always will be) multivariate. Also note that in a one-dimensional case, if X_t

$$\min_{j \in J^-, c \in \mathbb{R}} [\min_{\beta_1} \sum_{t \in l | S_{j,t} \leq c} (y_t - X_t \beta_1)^2 + \min_{\beta_2} \sum_{t \in l | S_{j,t} > c} (y_t - X_t \beta_2)^2] \tag{1}$$

Y_t and X_t are time-indexed features. J denotes all features to split by, J^- denotes the subset of J which are used as in RF at each split, only a subset is considered. $S_{j,t}$ is the subset of datapoints for which feature j at time t is greater or smaller than c. $l1$ and $l2$ denote the respective child leaves and l all data points assigned to the parent leave by the previous splits. Meek et al. (2002) propose an *Autoregressive Tree* which uses autoregressions and considers the time itself as the only feature to split by, hence performing the splitting as laid out in formula 2.

$$\min_{t \in T} [\min_{\beta_1} \sum_{t \leq t^*} (y_t - X_{t-1} \beta_1)^2 + \min_{\beta_2} \sum_{t > t^*} (y_t - X_{t-1} \beta_2)^2] \tag{2}$$

The relationships of these to the trees in MRF shall be discussed in greater detail in section 3.1.1. Apart from the paper introducing MRF (Coulombe, 2020), Coulombe et al. (2021) conduct an analysis which bears resemblance to the analysis presented here. They deploy MRF (amongst many other models) in a large forecasting study for the British Macroeconomy. While the dataset used in the following analysis is largely identical, there are several important differences between the studies. The authors focus mostly on different forecasting targets and focus on the performance of MRF during the Covid-19 induced recession. More importantly, they do not derive economically meaningful relationships from MRF (as well as all other forecasting tools) but treat them as mere *black boxes*.

is one, the splitting procedure becomes that of an *ordinary* random forest. Also, one could impose l_1 or L_2 regularization in the regression parts.

3 Methodology

3.1 The Macroeconomic Random Forest

3.1.1 Trees in MRF explained

The trees in MRF combine the above mentioned approaches of the *Autoregressive Tree* and the *Linear Tree*. While MRF trees also use multiple features (S_t) to find optimal splitting points, [2] they make use of past realizations of either the dependent variable Y (Y_{t-1}...) or other features X ($X_{t-1,t-2,...}$) in the linear parts. This leads leads to autoregressive models which can also be used for forecasting (Coulombe, 2020). Obviously, in such models, the features used for splitting (S) and the independent variables in the linear parts (X) are lagged compared to the dependent feature (Y_t). To keep things simple and follow the notation of Coulombe (2020), S_t and X_t are assumed to contain the lagged values of the corresponding features which are needed to model Y_t. This idea is explained in further depth in section 3.1.3. Additionally, the time t is added to the set of variables to split by (see chapter 3.1.3). MRF trees also deploy a sophisticated regularization regime, which is explained before the actual splitting procedure is laid out.

To prevent the trees in a MRF from learning idiosyncrasies, two regularization methods are introduced. Firstly, for each coordinate split, an L_2-Shrinkage is imposed on $\beta_{1/2}$. However, as λ increases, some βs might be shrunk towards zero which is undesirable in a time series setting with a low dimensional X_t. Hence, Coulombe (2020) suggests using a mild L_2 regularization and to combine it with a *Random Walk Regularization*. The latter is based on the idea of smoothing β_t to be close to $\beta_{t\pm1/2}$ through solving a weighted regression problem in each coordinate split which assigns weights below one to time-wise close data-points. The weight kernel utilised for this is a *5-Step Olympic Podium* (see Equation 3) with a smoothing parameter $\xi < 1$.

[2] Technically speaking, as is the case in Random Forests, not all features are considered at each splitting point. The latter statement merely explains the relationship of MRF trees to those introduced in equation 1.

$$\omega(t,\xi) = \begin{cases} 1, & if \;\; t \in l \\ \xi, & if \;\; t \in (l_{+1} \cup l_{-1})/l \\ \xi^2, & if \;\; t \in (l_{+2} \cup l_{-2})/(l \cup (l_{+1} \cup l_{-1})) \\ 0, & otherwise \end{cases} \tag{3}$$

$$l_1(j,c) = \{t \in l | S_{j,t} \leq c\} \; ; \; l_2(j,c) = \{t \in l | S_{j,t} > c\} \tag{4}$$

$$l_i^{RW} = l_i(j,c) + l_i(j,c)_{\pm 1} + l_i(j,c)_{\pm 2} \qquad \forall i \in \{1,2\} \tag{5}$$

The data-points in $t_{\pm 1/2}$ that would be assigned to l via the previous splits of the tree at hand thus influence the splitting procedure of a tree fitted for time t. Again, l_1 and l_2 are defined as the data that fall in the respective leaves for a time t with values of a certain features S lower/higher than the splitting point c (see equation 4). Taking all previous explanations into account, one obtains l_i^{RW} according to equation 5. Finally, the splitting rule of MRF trees looks as follows (Coulombe, 2020). [3]

$$\min_{j \in J^-, c \in \mathbb{R}} [\min_{\beta_1} \sum_{t \in l_1^{RW}(j,c)} \omega(t,\xi)(y_t - X_t\beta_1)^2 + \lambda \|\beta_1\|_2 + \min_{\beta_2} \sum_{t \in l_2^{RW}(j,c)} \omega(t,\xi)(y_t - X_t\beta_2)^2 + \lambda \|\beta_2\|_2] \tag{6}$$

Obviously, for each coordinate split, only those datapoints that are assigned to a leaf through all proceeding coordinate splits are considered. This also applies to the data in l_1 and l_2. The β-parameter values in the final leaves finally constitute the GTVP. When a trained tree is confronted with data, the parameters in the corresponding final leave will be used for the prediction and (ideally) capture the relationship of X and Y against the backdrop of the current realizations of the variables used in the spitting procedure. Choices of the variables in X and S can differ depending on the modelled economic relationship as will be discussed in the following chapters.

[3] Note again that when using auto-regressive setting as will be used in the following analysis, X_t in equation 6 become X_{t-h} and $S_{t,j}$ in equation 4 become $S_{t-1,j}$.

3.1.2 From Trees to a Forest

As its name indicates, a RF ensemble learner based on the trees outlined above is set-up in MRF (Coulombe, 2020). Bootstrap-aggregation and de-correlation (using J^{-} instead of J in the splitting procedure - see formula 6) are being deployed. The Bootstrapping occurs though the sampling of triplets of the form $Z_t = [Y_t,\ X_t,\ S_t]$. [4] Importantly, merely applying this procedure violates the *iid*-assumption needed for bootstrap aggregation is violated as there is temporal dependence (e.g. Y_t depends on Y_{t-1}) (MacKinnon, 2006). Block samples in the form of $Z_b = [Y_{b:B},\ X_{b:B},\ S_{b:B}]$ with $T/BlockSize$ blocks per bootstrap sample have been shown to be *iid* should the input data be weakly stationary (MacKinnon, 2006). [5]

Predictive modelling based on MRF occurs as follows: After each tree learns on its bootstrap sample, the trees are fitted on all data. For those blocks, which were not used in the training set of the respective trees, the parameters of the final node that the data points belonging to t fall into are retrieved. This is referred to as *Out-Of-Bag - OOB* inference. Importantly, Coulombe (2020) shows that this procedure gives rise to a Bayesian Interpretation of the GTVP: One obtains an *OOB* distribution for the parameter values per t, which can be interpreted as a posterior distribution. [6] In this framework, the posterior means constitute the GVTP and credible intervals can be calculated. The above-mentioned bootstrapping mechanism is hence also known as the *Block Bayesian Bootstrap (BBB)*. Should the trees only be fitted using data up to a certain point in time t^*, *Out-Of-Sample (OOS)* inference is possible via fitting all trained trees on the data after t^*. Again, the parameters in the associated final leaves $\forall\ t > t^*$ form a distribution that can be interpreted as laid out above. Ideally, even if $S_{t|t\leq t*}$ differs substantially from $S_{t|t>t*}$, F, the MRF functional that relates the economic input features to the GTVP, can adapt to this and find the corresponding parameters (see formula 7).

$$\beta_t = F(S_t) \tag{7}$$

[4] Note again that X_t and S_t contain lagged values

[5] Technically, $vec(Z_b)$ is *iid*. Also, all transformed data were checked for stationarity - see chapter 3.2

[6] More details on the assumed prior and likelihood functions consult the appendix of Coulombe (2020) and this papers appendix.

Conveniently, each tree in MRF is observeable. The parameters evolve based on clearly detectable structures in the form of economic variables describing economic settings. This does not hold for MRF models, yet the relevance of these splitting variables can be determined using the *Variable Important (VI)* approach (Brieman, 2001). Coulombe (2020) suggests using the *Out-of Samples (OOS)* permutation-accuracy-based VI metric. Firstly, the RMSE for OOS data are computed. Then, per used feature to split by, the values of the datapoints are shuffled at random (keeping all other variables the same) and the model is fitted again with the permuted data. The increase in RMSE serves as a measure of variable importance (Brieman, 2001). [7] Once a few important variables are defined, surrogate trees using only these variables can be set up. These might approximate the learned GTVP well and allow for clear economic interpretation of both the parameters and the economic setting influencing these, both of which are of great interest to economists and policy makers (Coulombe, 2020).

3.1.3 Engineering MRF Input Data

A crucial hyperparameter to consider when making use of MRF is the set-up of the data used in the RF part S_t as the node-splitting and thereby the flexibility of modelling different economic scenarios is influenced by the latter (Coulombe et al., 2021). At first glance, it might seem as if a simple lag of all series except the outcome series (outcome series: Y_t, all other lagged variables: X_{t-1}) might do the trick. Importantly, following Coulombe (2020), in this sub-chapter, X_t does not merely refer to the features included in the linear parts, but to all available series except the output series. However, it may well be that Lag polynomials of the dependent variable Y $(Y_{t-1}, ..., Y_{t-j})$ or further lags of X $(X_{t-2}, ..., X_{t-p})$ constitute important variables for the characterization of economic scenarios. To capture more general trends in single variables, using dense representations in the form of the first cross-sectional K_1 principal components of all X_{t-i} for $i = 1, 2, ...$ and the first K_2 principal components of j Lag Polynomials of Y is suggested by Coulombe (2020). The latter are referred to as *Moving Average Factors* or MAFs. Coulombe (2020) also proposes to include a *Trend* variable as a simple time range to capture external

[7] Brieman (2001) originally suggested to use *Out-Of Bag* VI, for which the just explained procedure is conducted on datapoints which were not used to train the respective trees. Also, the method was originally called 'Accuracy-based' since (Brieman, 2001) uses this measure for classification problems.

structural breaks. For the following analysis S_t, the input for the MRF tree part and X_t, the data set used for the classical methods, were set up as explained in table 1. [8] Table 2 below shows in greater detail for which models X_t and S_t are used, respectively.

Data set	4 lags of Y_t	Trend T	X_t lagged once	X_t lagged twice	2 lags of 5 Factors of X_t	2 MAFs on 8 lags
S_t	✓	✓	✓	✓	✓	✓
X_t	-	-	✓	-	-	-

Table 1: Arrangement of Series for different Input data sets

3.2 The Dataset

Coulombe (2020) applies MRF to US-Macro data, both on a quarterly and a monthly level. In total, he makes use of 248 quarterly Series and 119 monthly series. The gap between these numbers can easily be explained by the fact that many Macroeconomic series (e.g. GDP, Investment data) are not available at a monthly frequency. All variables are assigned to 14 overarching groups. To ensure the high dimensionality of the input data for MRF, the *UK MD* dataset, which is provided by Coulombe et al. (2021) shall be used here. It is compromised of 113 series at a monthly frequency. The only monthly variable added to the latter is the 1 Year Yield of British Government Securities, obtained from Fusion Media Limited (2022). All variables in *UK MD* are real, seasonally adjusted and at the same *level* (e.g. Monetary Figures in 1M Pound). Also, the authors provide stationary counterparts as well as procedures to attain the stationary series. [9]

For data at a quarterly frequency, aggregates of the raw, non-stationary monthly variables were calculated with either their sum or their mean (see *Aggr.Type* in Tables 11 to 14). Furthermore, a total of 13 additional quarterly available series are obtained from the FRED Database *FRED* and the National Institute of Statistics *ONS*. [10] The latter have been chosen in such a way that the 14 data categories also used in Coulombe (2020) have multiple series series assigned to them (see *Group* in Tables 11 to 14). Next, some added

[8] Notably, Coulombe (2020) uses more lags for some series in the sets and includes MAFs for each series. This was not replicated 1:1 because of the shorter series at hand. For more details, see Table 1.

[9] An Augmented-Dickey-Fuller Test has been conducted on the Raw 1 Year Government Bond Yield and its I(1) counterpart, the latter of which is stationary and thus used in the monthly dataset. More on this procedure in the text below.

[10] See all citations for the added series in the references

variables were transformed to be real and at the same *level* (see *Pre-transf.* Tables 11 to 14). As a final step, transformations resulting in stationary counterparts were put in place. Coulombe et al. (2021) provide methods per series to achieve the latter for all data in the UK MD dataset. For the further series, Augmented-Dickey-Fuller (ADF) tests were conducted to determine stationarity and in the case of the latter not holding, specific transformations were applied to the series and second ADF-tests validating that stationarity has been achieved were conducted. [11]

4 Analysis

4.1 Using MRF for Forecasting

MRF's alleged ability to forecast flexibly depending on the current economic state leads Coulombe (2020) to contrast the predictive performance of the latter (with different linear parts) against classical parametric models and *plain* RF. He finds that MRF models have an edge over classical models when confronted with unseen data. Even though the economic settings as characterized by the S_t variables might be different from those present in the data used for learning, MRFs seem to be able adapt correctly and well-fitting parameters depending on the new economic scenarios can be retrieved. [12] Mathematically speaking, S_t might change, but F is stable and can deal with the change at hand (see equation 7). To assess if this finding also holds up for the British Macro data, forecasting studies using both the monthly and the quarterly data sets for different respective forecasting horizons H were conducted. Following Coulombe (2020), the study shall focus on 5 different Macro-economically relevant dependent variables: Unemployment rate, CPI inflation, 1 Year Government Bond Yield, SPREAD (10 Year GVT Bond Yield - Bank rate) [13] and GDP - the latter of which is only available at a quarterly frequency.

As stated before, a multitude of MRF models with differing linear parts can be set up. Obviously, the linear parts should include reasonable regressors describing a potential

[11]The results of the ADF-tests can be found in Table 10, while the corresponding Transformation Codes can be found in the *Stat.-Transf.* Column of Tables 11 to 14.

[12]The parameters for forecasting will in such cases be those of the final leaves in which fewer observations in the training data fell into

[13]Coulombe (2020) uses the Federal Funds Rate as a dominant indicator of short term interest rates.

relationship of X and Y. For the inquiry at hand, MRF models with auto-regressive and vector-auto-regressive linear parts are chosen. These are contrasted against an AR(4) Model and a plain RF and a RF based on the enlarged dataset S. The specifics of the models are disclosed in table 2 are being used. [14] Importantly, for $H > 1$, the data in the linear part and the RF part is lagged further. For example, with $H = 2$ the forecasting equation in the ARRF (leaves) is laid out in equation 8.

Name	Acronym	Linear Part	RF Part
AR(4)-Model	AR(4)	[1, $Y_{t-[1:4]}$]	-
Plain RF	RF	-	X_t
RF with S_t	MAF-RF	-	S_t
Autoregressive MRF	ARRF	[1,$Y_{t-[1:2]}$]	S_t
Vector-Autoregressive MRF	VARRF	[1,$Y_{t-[1:2]}$, 1 Lag of other dependent variables]	S_t

Table 2: Different forecasting Models. Note that X_t and S_t refer to the data needed to forecast Y_t and therefore contain values corresponding to previous periods (see table 1).

$$Y_{t+2} = \alpha_t + \beta_{1,t} Y_t + \beta_{2,t} Y_{t-1} + u_{t+2} \tag{8}$$

Tables 3 and 4 show the results of the study with the *RMSE* scores depicted in the columns. The *OOS*-sets compromise the last quarter of the data. [15] To obtain stable forecasting results, 100 trees were train per MRF model. The correct choice of Hyper-parameters (*HP*) differs greatly for both Y and H. Thus, for all MRF models, an HP-grid search was conducted and the model with the best *OOS* performance was chosen. [16] To keep the comparison fair, small HP-searches were also conducted for the *RF* and *MAF-RF* models.

Tables 3 and 4 clearly show that both the *ARRF* and the *VARRF* approaches provide forecasting gains over all methods analyzed here while the *RF* and *MAF-RF* models tend to under-perform the plain AR(4) model. Broadly speaking, the autoregressive linear

[14]To keep things simple, fewer models than in Coulombe (2020) shall be used here. The models deployed have been found to work well for the dataset at hand

[15]The *OOS*-sample thus starts at 1.10.2015 for the monthly data and Q2 2016 for the quarterly data.

[16]E.g.: For a series with less spikes, a lower value of ξ in the random walk regularization might be feasible. The other HP considered were λ - see formula 6, the fraction of all features to consider per split and the block size

H	AR (4)	RF	MAF-RF	ARRF	VARRF	AR (4)	RF	MAF-RF	ARRF	VARRF
	Unemployment Rate					CPI Inflation				
1	0.0854	0.0986	0.0955	0.0858	**0.0832**	0.0022	0.002	0.0019	**0.0018**	0.002
3	0.0952	0.0843	0.083	0.0861	**0.0818**	0.0021	0.0021	0.0021	0.0021	**0.0019**
6	0.1004	0.0998	0.0967	**0.0916****	0.0919	0.0021	0.0021	0.0021	**0.0021**	0.0021
9	0.1014	0.1113	0.1087	**0.0971**	0.0973	0.0022	0.0022	0.0023	**0.0020**	0.0021
12	0.0963	0.1001	0.0988	0.093	**0.0904**	0.0022	0.0021	0.0022	0.0021	**0.0021***
	1 Year GVT Bond Yield					SPREAD				
1	0.1124	0.1542	0.1646	**0.1083**	0.1121	0.1318	0.1349	0.1352	**0.1264**	0.1280
3	0.1114	0.1571	0.1537	**0.1042***	0.1077*	0.1427	0.1451	0.1439	**0.1348**	0.1351
6	0.1071	0.1649	0.1482	**0.1058**	0.1058	**0.1318**	0.1629**	0.1616**	0.1323	0.1342
9	0.103	0.1314	0.1291	**0.0968**	0.1008	0.1372	0.1527	0.1466	**0.1335**	0.1337
12	0.1058	0.1279*	0.1202*	**0.1001**	0.1015	0.1276	0.1207	0.1283	**0.1203**	0.1238

Table 3: Monthly forecasting results in the form of respective rounded $RMSE$ values. Best performing method per Y and H is in bold. The stars indicate significance values of two-sided Diebold-Mariano tests against AR(4) [* for $\alpha = 0.1$, ** for $\alpha = 0.05$, *** for $\alpha = 0.01$].

parts seem to provide useful information that a simple RF cannot capture. On the other hand, GTVP, which the AR(4) model does not allow for, also enhance the predictive performance. Depending on Y and H, these forecasting gains are statistically significant. Following Coulombe (2020), two-sided Diebold-Mariano tests were conducted. These are commonly used statistical tests to determine significances in the difference of forecasting performance. [17] All in all, these findings (for all Y) are in line with the results of the US-study of Coulombe (2020). The forecasting study thus provides more evidence for MRFs ability to capture changing Macro-economic relationships. Indeed, for the British Macroeconomic Random Forests at hand, F seems too be stable regardless of S_t in the input sample.

4.2 Buidling explainable surrogate trees

While MRFs can evidently bring about forecasting gains, the economic relationships modelled are hard to explain: As was mentioned above, the single trees in an MRF are directly explainable models, yet the aggregation into an MRF distorts this direct interpretation. What is more, especially trees with few splits are of greater utility for economists and policy-makers as they indicate economic relationships which are easily understandable and interpretable (Coulombe, 2020). This gives rise to the idea of fitting linear, surro-

[17]Technically, the loss differentials in the tests were based on squared residuals. Also, the respective forecasting horizon was accounted for. For more info, consult Diebold & Mariano (1995).

H	AR (4)	RF	MAF-RF	ARRF	VARRF	AR (4)	RF	MAF-RF	ARRF	VARRF
	Unemployment Rate					CPI Inflation				
1	0.1178	0.0987	0.0999	0.0995	**0.0942***	0.0023	0.0026	0.0024	0.0023	**0.002**
2	0.1186	0.1299	0.1202	0.1057	**0.0982**	0.0025	0.0027	0.0026	**0.0024**	0.0025
4	0.1325	0.1507	0.1499	**0.1129**	0.1158	0.0029	0.0026	0.0026	0.0031	**0.0025**
	1 Year GVT Bond Yield					SPREAD				
1	0.1723	0.1561	0.1744	**0.1541**	0.1708	0.2479	0.7818**	0.8014**	**0.2085**	0.2561
2	0.1633	0.1766	0.1821	**0.1304**	0.1493	0.3608	0.8747*	0.891*	0.3209	**0.2892**
4	0.173	0.1994	0.2302	0.163	**0.156**	0.5267	0.9493*	0.9276*	**0.3287****	0.4193
	GDP									
1	0.0807	0.0621	**0.0621**	0.0697	0.0836					
2	0.0655	0.0623	**0.063**	0.0641	0.0635					
4	0.0665	0.0672	0.067	**0.0664**	0.0664					

Table 4: Quarterly forecasting results in the form of respective rounded $RMSE$ values. Best performing method per Y and H is in bold. The stars indicate significance values of two-sided Diebold-Mariano tests against AR(4) [* for $\alpha = 0.1$, ** for $\alpha = 0.05$, *** for $\alpha = 0.01$].

gate trees based on only a handful of relevant variables. Based on those variables, which were found to be most relevant based on the VI approach for MRF (see Methodology), Coulombe (2020) is able to set up simple trees disclosing understandable economic relationships. These findings further elevate the alleged practicality of MRFs and tree-based modelling.

As was done in the forecasting study, these findings are put to the test for the British Macro-data at hand. The relevant variables were derived from the ARRF models because of their superior predictive performance. Also, the corresponding AR(2) models were used for the coordinate splits and the final leaves of the linear trees. To keep things simple, one tree was set up per dependent series. A forecasting horizon of one quarter was chosen to ensure comparability between monthly and quarterly frequencies in the datasets. Hence, $H = 1$ holds for GDP and $H = 3$ holds for all other Y. Since the VI approach is computationally expensive, no variable importances were derived in the forecasting study. Thus, using the optimal Hyper-parameters, the corresponding ARRF-MRF models were refit. As is suggested by Coulombe (2020), 200 trees were now used per MRF to achieve more stable results for the variable importances and the OOB and OOS parameter distributions. [18] For each Y, the four variables deemed most relevant for the splitting process were retrieved and subsequently used as the only variables to split by for the surrogate

[18] While the parameter distributions are not directly needed for the VI approach or the set-up of the linear trees, these shall be displayed visually in plots comparing the resulting parameters.

trees. [19] As linear models with few splits are of greatest value, regularization methods restricting the size of the trees come in handy. Zeileis et al. (2008) suggest a procedure in which a further split is only conducted if one of the parameters can be said to be varying systematically by at least one of the variables to split by. Should that not be the case, further splits cannot be justified based on the idea of GTVP. This procedure is based on a *Parameter Stability Test*, which assumes no systematic difference in the parameters by any of the splitting variables (Zeileis et al., 2008). If this H_0 cannot be rejected for any of the splitting variables, the splitting process to an end. This method was deployed with a significance level of 0.1. Further information on the Stability Test can be found in the appendix. Note that as was the case in the forecasting study, both the MRF and the single trees were only fit on roughly three quarters on the data to allow for measuring the *OOS* fit. Only if the *OOS*-RMSE score for a single tree is substantially similar to that of the corresponding MRF, the alleged benefit of the linear trees as explainable models, which also capture the underlying economic relationship well, can be validated. The resulting linear trees shall be explained and their parameters contrasted to those of the MRF and simply OLS models. In the following, the findings per dependent variable are laid out. As Coulombe (2020) uses a different linear equations in the trees (based on an MRF model which was not included in the forecasting study) and only reports on trees for the unemployment rate and CPI inflation, comparing the findings per dependent variable is not possible. Note that like in the forecasting study, the stationary counterparts are used for all variables. Those plots showing the surrogate parameters in contrast to the MRF and OLS parameters which are not included in the following section can be found in the appendix.

4.2.1 Unemployment Rate

Table 5 shows the most important variables to split by in the MRF for the unemployment rate as the dependent variable. While these variables surely seem like features influencing the unemployment rate, it is important to bear in mind that these features were found to influence the one-quarter auto-regressive relationship of the latter. E.g. a decrease in the unemployment rate in t paired with an increase in the vacancies t or greater imports

[19]Coulombe (2020) uses more variables, yet four variables were found to suffice for well-fitted and explainable trees.

of machinery indicating greater production in t might lead to a lower unemployment rate in t+3. This does not seem feasible for a situation in which the vacancies or machinery imports decrease in t. Increases in the vacancies in t-1 with no decrease in unemployment in t-1 might still lead to an decrease in t as employment might lag behind the vacancies. This might carry through to t+3. In that case, the auto-regressive relationship from t-1 to t+3 would be negative, while that from t to t+3 would be positive. Changes in the unemployment might have different effects on changes in the employment rate in t+3 depending on how many people are already tied to work, which AVG-WEEK-HRS-FULL might proxy for. Nonetheless, direct influences, i.e. changes in the important variables leading directly to changes in the unemployment rate, are captured by different intercept values.

Rank	Name	Lag	Short Description
V1	VAC-TOT	4	Total Vacancies (first diff of log)
V2	VAC-TOT	3	Total Vacancies (first diff of log)
V3	AVG-WEEK-HRS-FULL	3	Average weekly hours of work (first diff of log)
V4	IMP-MACH	3	Trade Imports - Machinery (first diff of log)

Table 5: Four most important splitting variables in ARRF for CPI Inflation. *Lag* denotes that time lag compared to forecasted variable. Consult tables 11 to 14 for more details on the variables.

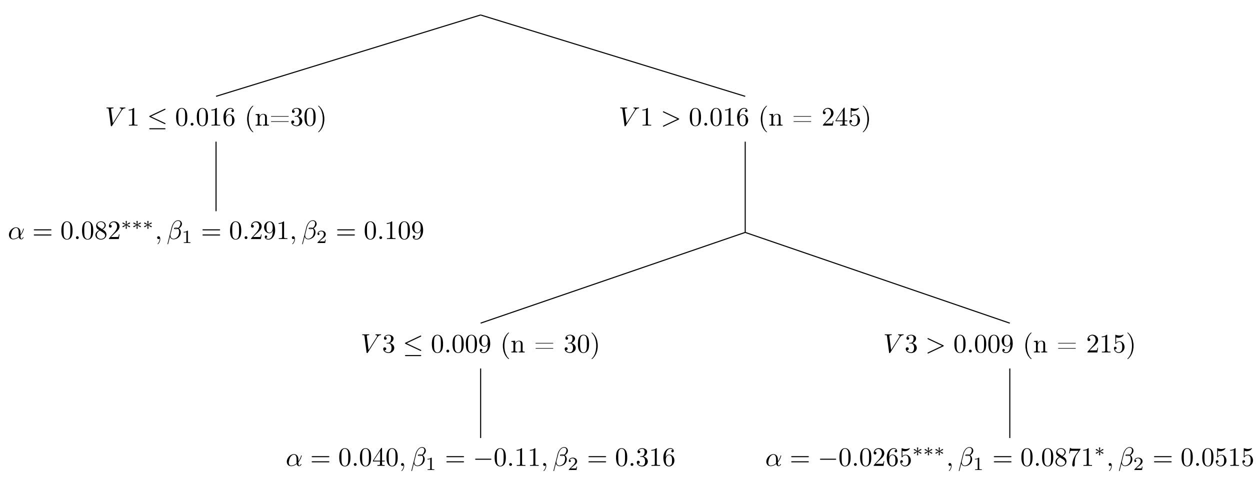

Figure 1: The structure of the surrogate AR(2) Linear Tree for Unemployment rate. Stars indicate significance of OLS-parameters [* for $\alpha = 0.1$, ** for $\alpha = 0.05$, *** for $\alpha = 0.01$].

The resulting surrogate tree is displayed below. Its *OOS*-RSME is 0.0854 and thus close to that of the corresponding ARRF model. Apparently, after splitting by the vacancies in t-1 and the hours worked in t, the AR(2) parameters do not vary systematically by the other variables. Upon analyzing the tree, it is striking that two parameters values of the intercept are highly significant. The high increase in the unemployment rate in t+3 for a decrease in vacancies in t-1 seems feasible (left leave). Despite using a different model, Coulombe (2020) also reports something similar as he finds that increases in private sector employment lead to a drop in the unemployment rate *per se*. Indeed, examining figure 2 reveals that during recessions, the α-parameters tended to be high. Albeit not significant, the high β_1 value seems reasonable as in such a setting as a decrease in employment might be associated with further decreases. [20] In contrast, a relatively strong increase in vacancies (t-1) and relatively high work load increase in t lead to a significant drop in unemployment. Such a setting might characterize an economic expansion. For many months during the past 20 years excluding time of recession, this seems to have been the case (see figure 2). However, noticeable exceptions can be found (e.g. from 2005-2006). The significant and positive β_1 parameter also fits to this narrative as this setting might lead to persistent drops in unemployment. A setting characterized by a rather strong increase in vacancies in t-1 without a strong increase in AVG-WEEK-HRS-FULL in t seems more ambiguous regarding the auto-regressive pattern. All the splits in V1 and V3 characterize economic settings in which the unemployment rate in t+3 *per se* differs from one setting to another. The most important part of the linear parts is hence the intercept. This is in line with the good performance results if the MARF-RF and the RF as these methods merely consider a single value (like an intercept) in the splitting procedure and the final leaves (Coulombe, 2020).

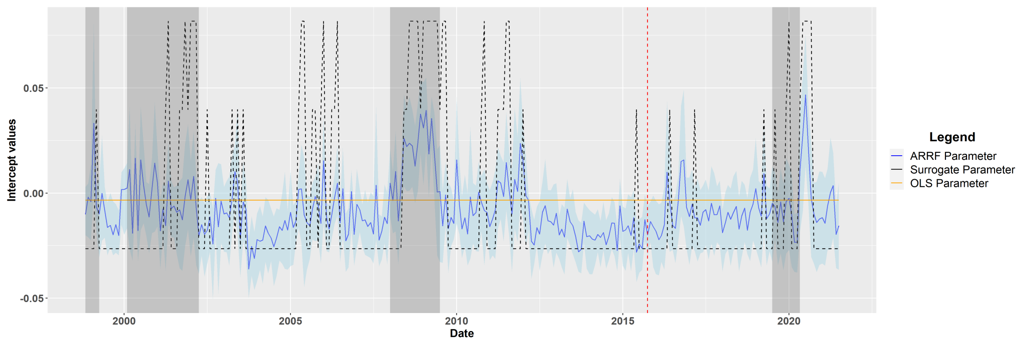

Figure 2: Comparison of intercept parameters for ARRF(2) for Unemployment Rate. Shaded areas indicate recessions. Red line indicates start of OOS-Period. The light blue bands show the 15%- and 75% credible intervals for the MRF parameters. Recession data was obtained from (Federal Reserve Bank of St. Louis, 2022).

[20] The corresponding plot in the appendix reveals that high β_1 values were rather present during recessions

4.2.2 CPI Inflation

The four most relevant variables in the MRF splitting process for the CPI inflation are displayed in table 6. Increases in BCI might correspond to increases in inflation *per se* due economic growth and to a strong auto-regressive relationship of inflation in t and inflation in t+3. The same might hold for trade exports as these lead to a greater inflow of money or proxy economic expansions. Strikingly, the first MAF was found to be an important variable. The nature of the auto-regressive relationship of the CPI inflation thus seems to depend on the development of this very feature across the last 8 months. [21] The finding that the past values determine the auto-regressive relationship is further strengthened by the fact that the fourth most important feature is the CPI inflation value in t-1 itself. Higher values of inflation might lead to stronger auto-regressive relationships from t to t+3.

Rank	Name	Lag	Short Description
V1	BCI	3	Business Confidence Index (first diff)
V2	EXP-CRUDE-MAT	3	Trade Exports - Crude Materials (first diff of log)
V3	MAF-PC 1	3	1^{st} Moving Average Factor
V4	CPI-ALL	4	Dependent Variable: Consumer Price Index (first diff)

Table 6: Four most important splitting variables in ARRF for CPI Inflation. *Lag* denotes that time lag compared to forecasted variable. Consult tables 11 to 14 for more details on the variables.

The tree's *OOS* RMSE of 0.00204 is quite similar to that of the ARRF counterpart. The only statistically significant parameter in the left leave is the intercept. In the case of a rather notable drop in BCI, the inflation still rises *per se*. The fact that the intercept values for the two right leaves are also highly significant with comparable magnitudes gives rise to the hypothesis that the inflation rises somewhat regardless of exogenous variables. Noticeably, the β parameters in the left leave are slightly negative and insignificant. In the described economic setting, when controlling for the general trend in inflation (measured by α), no significant auto-regressive dependencies can be determined. On the other hand, the β parameters in the two right leaves are stronger and significant. Apparently, when the drop in BCI is less pronounced, changes in inflation seem to be persistent.

In the case of rather low inflation in the preceding months (measured by MAF), the

[21] Recall that two Moving Average Factors were calculated as the first two principal components of the last 8 lags for the respective variable.

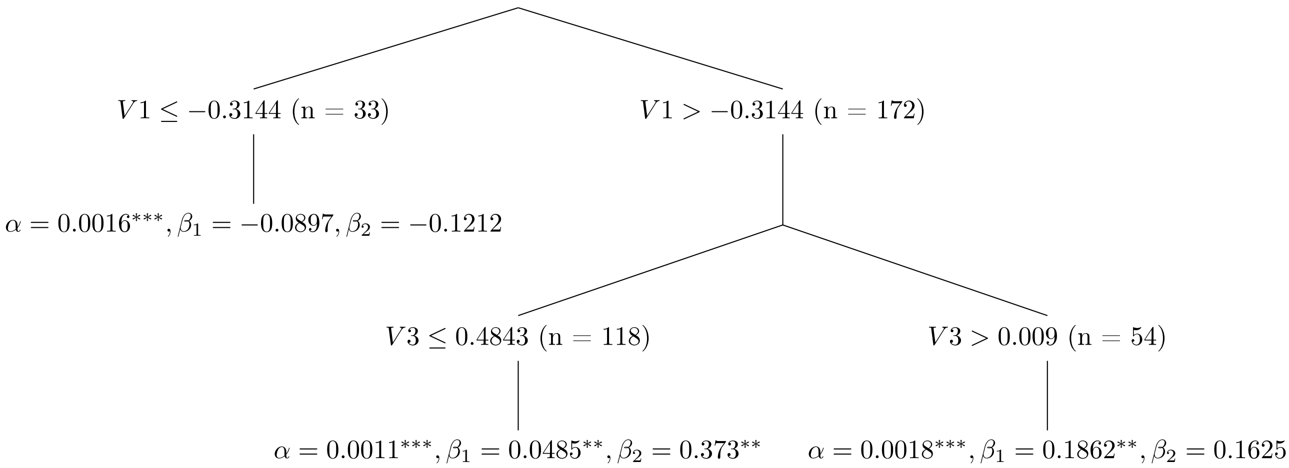

Figure 3: The structure of the surrogate AR(2) Linear Tree for CPI Inflation. Stars indicate significance of OLS-parameters [* for $\alpha = 0.1$, ** for $\alpha = 0.05$, *** for $\alpha = 0.01$].

inflation appears the most persistent as increases in t and t-1 carry through to the t+3 period. In the case of rather high preceding inflation, the inflation appears too be less persistent. These might be economic settings in which the inflation starts to cool off. Interestingly, figure 4 as well as the other corresponding pots in the appendix reveal that it is not uncommon for the economic setting characterized by the left leave to be followed by the economic setting characterized in the very right leave. It might be that after a drop in BCI, the inflation itself remains rather unaffected so that when the BCI increases again, the past inflation can still be deemed rather persistent. Unlike the tree at hand, Coulombe (2020) finds variables describing the housing market to be of great importance in the splitting process.

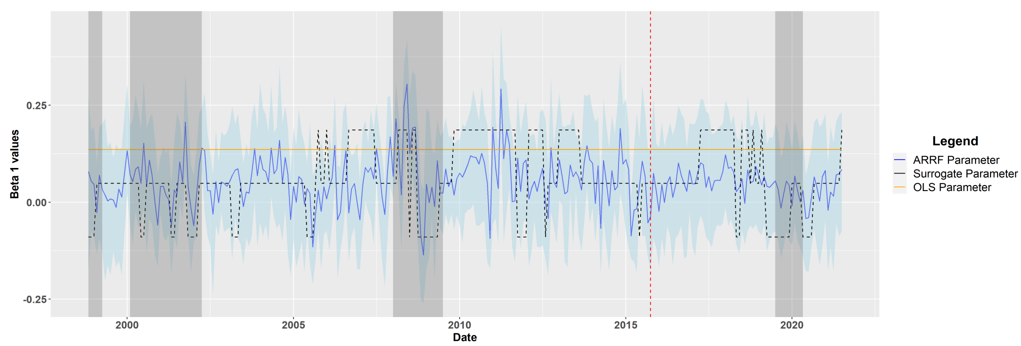

Figure 4: Comparison of intercept parameters for ARRF(2) for CPI inflation. Shaded areas indicate recessions. Red line indicates start of OOS-Period. The light blue bands show the 15%- and 75% credible intervals for the MRF parameters. Recession data was obtained from (Federal Reserve Bank of St. Louis, 2022).

Rank	Name	Lag	Short Description
V1	CPI-All	3	Consumer Price Index (first diff)
V2	BCI	3	Business Confidence Index (first diff)
V3	M2	4	M2 monetary aggregate (first diff of log)
V4	Bank-rate	3	Official Bank Rate (first diff)

Table 7: Four most important splitting variables in ARRF for 1 Year GVT Bond Yield. *Lag* denotes that time lag compared to forecasted variable. Consult tables 11 to 14 for more details on the variables.

4.2.3 1 Year GVT Bond Yield

The most important variables for the 1 Year Government Bond Yield, especially the first two, rather seem to be features which might have an effect on the yield development *per se*. Since the British Government Bonds pay a fixed coupon, inflation lowers the net worth of the latter coupon paid relative to the face value price of a bond, which decreases the yield (United Kingdom Debt Management Office, 2022). A drop in the BCI might trigger a flight into safer assets such as Government bonds and thereby lead to lower yields. Treating BCI as an indicator for future economic development, changes in BCI might also cause changes in the yields in the future. Thus, BCI might also influence the auto-regressive relationship. The other two variables mirror the Bank of England's monetary instruments: QE and setting the bank rate (Bank of England, 2022). As the bank rate is lowered, the coupons immediately become more attractive, which increases the demand for bonds and thereby lower their yields. An increase in M2 in t-1 might go along with greater bond purchasing. However, these two changes should mainly effect bond yields straight away. It might be that changes in both the bank rate and M2 indicate rather long term shifts in monetary policy and thereby indirectly affect yields in one quarter. Under this assumption, policy shifts might also affect the auto-regressive relationships: While yields might have risen throughout month t-1, if a drop in the bank rate at the beginning of month t may cause yields to fall in t - starting a persistent trend. In such a setting, a negative (positive) auto-correlation between the yields in t-1 (t) and t+3 could be explained.

Like in the preceding cases, the *OOS* RMSE of 0.1117 is close to that of the ARRF model. After splitting by the CPI inflation in t, no further systematic variations in any of the parameters can be explained. Strikingly, both intercept values are insignificant, which might indicate that no *per se* effect of a change in BCI on the yields exists. Given

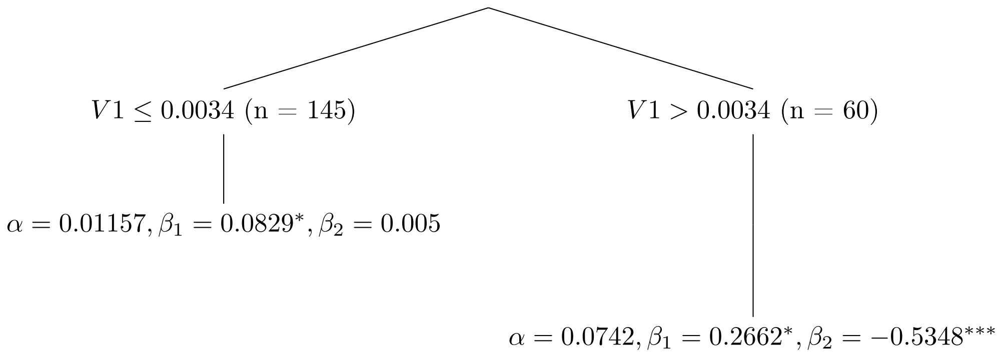

Figure 5: The structure of the surrogate AR(2) Linear Tree for 1 Year GVT Bond Yield. Stars indicate significance of OLS-parameters [* for $\alpha = 0.1$, ** for $\alpha = 0.05$, *** for $\alpha = 0.01$].

a rather low increase in BCI in t (left leave), a weak effect of a change in yields to that in one quarter seems to exist. This finding becomes more interesting when contrasting it to a higher increase in the BCI (right leave). The auto-regressive effect from t to t+3 is more than twice as strong. On the other hand, a negative, yet even stronger effect from the change in t-1 to t+3 exists. A plausible narrative might be that a yield decrease in t-1 followed by a strong increase in BCI (serving as a proxy for future economic expansions), might go along with a flight from safe assets which persists into t+3. However, with the exception of the Covid crisis and the subsequent economic recovery, figure 6 does not necessarily support this narrative as economic settings characterized by the right leave are also present in recessionary periods.

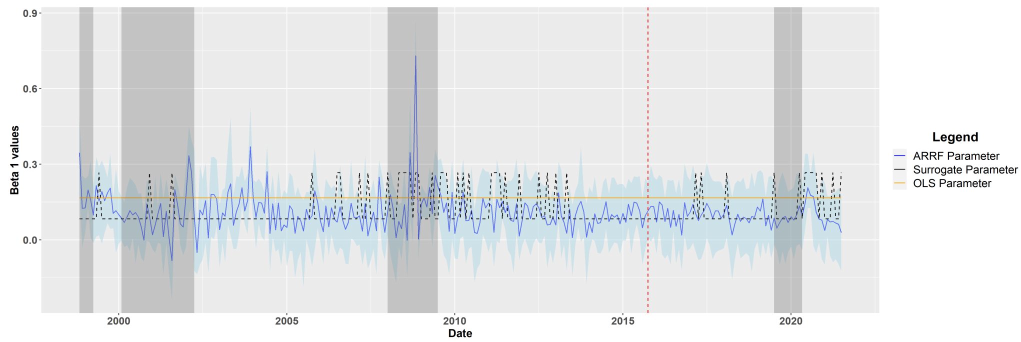

Figure 6: Comparison of intercept parameters for ARRF(2) for 1 Year GVT Bond Yield. Shaded areas indicate recessions. Red line indicates start of OOS-Period. The light blue bands show the 15%- and 75% credible intervals for the MRF parameters. Recession data was obtained from (Federal Reserve Bank of St. Louis, 2022).

Noticeably, the low predictive value of the intercepts coincides with the superior predictive performance of the ARRF and VARRF models compared to the RF and MAF-RF models.

4.2.4 SPREAD

Recall that the SPREAD under consideration here is the 10 Year GVT Bond Yield minus the bank rate. The first important variable appears relevant when considering the idea of interest rate parities across different currency zones. Assuming no changes in future interest rates, investors could exploit the drop in the spot rate by investing in safe long term bonds, like a 10 Year Government bond. This would decrease the yields of the latter and close the investment opportunity (Levich, 2012).

Rank	Name	Lag	Short Description
V1	GBP-US	3	Average Spot exchange rate, US-Dollar into Sterling (first diff)
V2	LIBOR-3mth	4	3-Month London Interbank Offered Rate (first diff)
V3	CCI	4	Consumer confidence index (first diff)
V4	TOT-WEEK-HRS	3	Total actual weekly hours worked (first diff of log)

Table 8: Four most important splitting variables in ARRF for SPREAD. *Lag* denotes that time lag compared to forecasted variable. Consult tables 11 to 14 for more details on the variables.

As the LIBOR is greatly affected by key interest rates like the bank rate, decreases in the latter might be associated with two effects at the same time. Firstly, a decrease in the LIBOR might proxy a drop in the bank rate, which directly increases the SPREAD. However, this effect would come into action right away. Secondly, a drop in LIBOR might serve as a proxy for strong future economic development and a subsequent flight from save long-term assets. However, under the assumption that market participants can fully interpret the latter, this piece of information should already be priced in for the bonds in t (Reserve Bank of Australia, 2022). An increase in the CCI in t-1 as well as a lagging increase in the amount of hours worked in t might indicate positive economic developments. However, this should directly influence the yield curve since these expectations are commonly being priced in (Reserve Bank of Australia, 2022). Theoretically, a decrease in SPREAD in t-1 paired with an increase in CCI, which translates into an expansion in t, which persists into t+3, might explain negative (positive) auto-regressive relations of the SPREAD values from t-1 (t) to t+3.

Once again, the *OOS* RMSE of 0.1443 is close to that of the ARRF model and the linear tree is rather shallow. As was hypothesized above, in the case of a pronounced decrease of the LIBOR in t-1, the SPREAD increases *per se*. Noticeably, no other statistically significant relationships can be deduced. Surprisingly, many of such economic

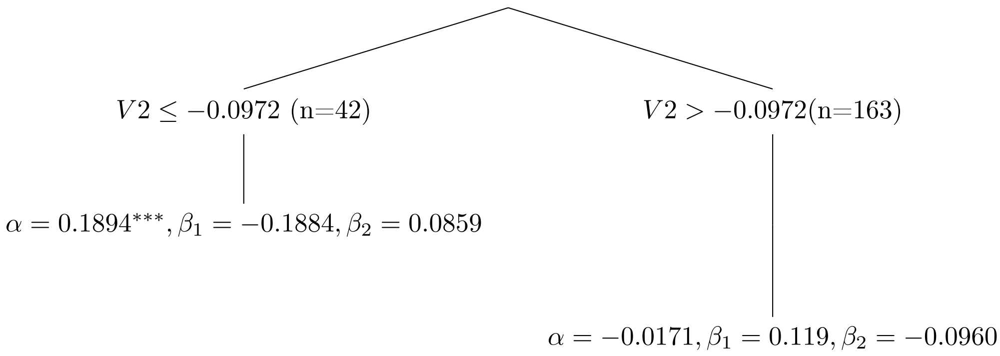

Figure 7: The structure of the surrogate AR(2) Linear Tree for SPREAD. Stars indicate significance of OLS-parameters [* for $\alpha = 0.1$, ** for $\alpha = 0.05$, *** for $\alpha = 0.01$].

settings (left leave) fell into recession periods (see figure 8). In such cases, long-term yields might already be low. Further decreases in the bank rate could drive LIBOR rates further down and open the SPREAD again. It does not come as a surprise that the predictive performance of the ARRF was strikingly similar to that of the MAF-RF and the RF, which only use an intercept.

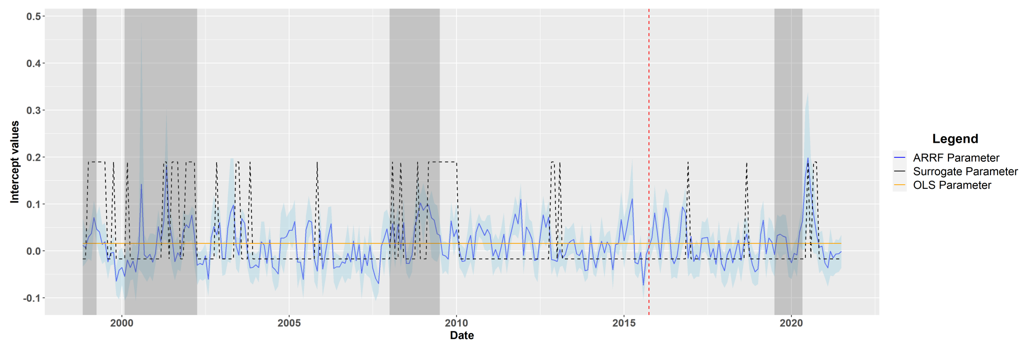

Figure 8: Comparison of intercept parameters for ARRF(2) for SPREAD. Shaded areas indicate recessions. Red line indicates start of OOS-Period. The light blue bands show the 15%- and 75% credible intervals for the MRF parameters. Recession data was obtained from (Federal Reserve Bank of St. Louis, 2022).

4.2.5 GDP

As mentioned before, for the GDP as Y, the forecasting horizon becomes 1 and the AR(2) function becomes: $Y_{t+1} = \alpha + \beta_1 Y_t + \beta_2 Y_{t-1} + u_{t+1}$. Again, the *important variables* seem reasonable. If the LIBOR increased (decreased) in t-1, a rising (decreasing) GDP in t might turn into a slower increasing or decreasing (less decreasing or rising) one in t+1 in case of a lagged effect. The same might hold for the BCI in t-1. LIBOR and BCI might also be directly linked to the GDP due to lower values in key interest rates such as LIBOR and higher values in BCI fueling GDP growth. Decreases (increases) in unemployment might go hand in hand with GDP increases which might carry through to t+3. The same is economically feasible for increases or decreases in the production of durables.

Rank	Name	Lag	Short Description
V1	LIBOR-3mth	2	3-Month London Interbank Offered Rate (first Diff)
V2	BCI	2	Business confidence index (first Diff)
V3	UNEMP-DURA-12mth	1	Unemployment over 12 months (Log of first Diff)
V4	IOP-DUR	1	Index of Production - Consumer Durables (Log of first Diff)

Table 9: Four most important splitting variables in ARRF for GDP. *Lag* denotes that time lag compared to forecasted variable. Consult tables 11 to 14 for more details on the variables.

The tree, which is laid out below is rather shallow. Nonetheless, its *OOS*-RMSE of 0.0806, which again is fairly close to that of the ARRF counterpart and validates its ability to correctly model the underlying economic relationship.

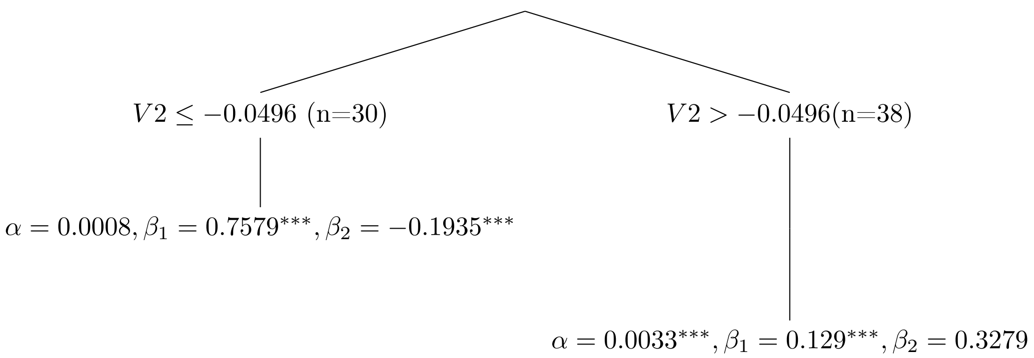

Figure 9: The structure of the surrogate AR(2) Linear Tree for GDP. Stars indicate significance of OLS-parameters [* for $\alpha = 0.1$, ** for $\alpha = 0.05$, *** for $\alpha = 0.01$].

When controlling for a division by rather strong (left leave) and rather weak (right leave) decreases in the BCI in t-1, the other features do not explain further systematic variation in the AR(2) parameters. A look at the intercepts gives rise to the feasible narrative that a less pronounced decrease in BCI *per se* leads to a higher GDP growth. In the case of a comparatively strong decrease of BCI in t-1, a negative/positive GDP change in t seemingly carries though to t+1 (see left β_1). Strikingly, β_2 is negative though. Treating BCI as a predictive indicator for GDP development, one might hypothesise that when GDP development is still positive in t-1 and BCI decreases in t-1, a GDP drop will follow in t which carries through to t+1. The narrative is partially supported by the fact that in recessionary times, β_1 (β_2) was high (low) as can be deduced from figures 10 and 11. On the other hand, the β-parameters took on those values during many non-recession periods as well. In the case of a less pronounced drop in BCI, both β-parameters are positive with the first being statistically significant. In such a more ordinary setting, the changes in GDP expectedly seem to be more persistent, albeit with a weaker effect from t to t+1.

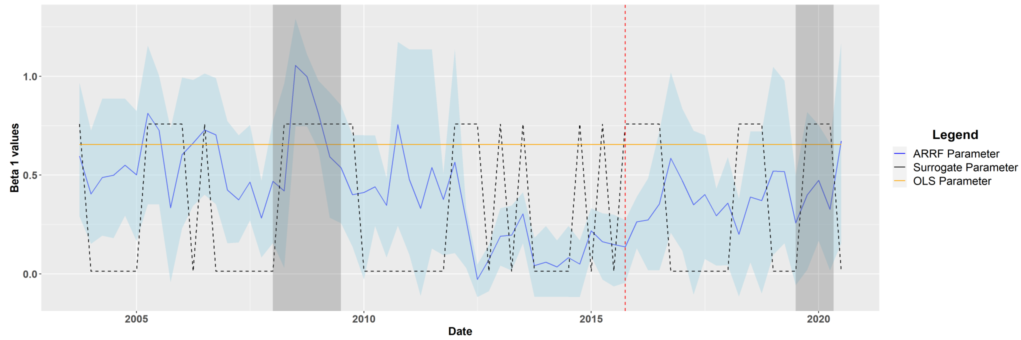

Figure 10: Comparison of β_1 parameters for ARRF(2) for GDP. See figure 11 for more info.

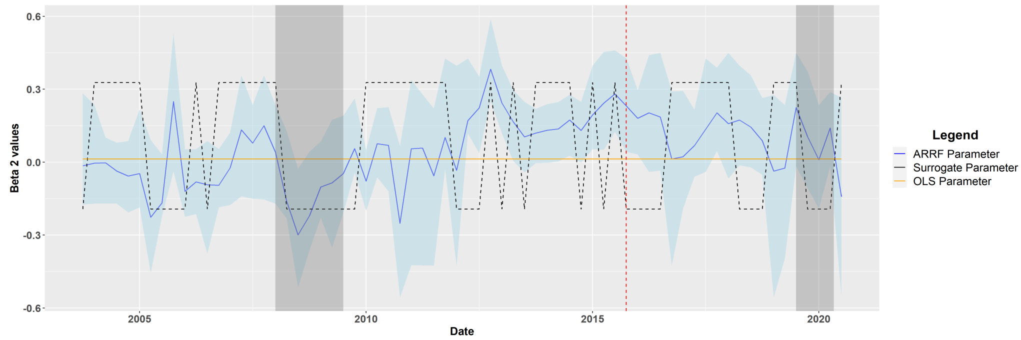

Figure 11: Comparison of intercept parameters for ARRF(2) for GDP. Shaded areas indicate recessions. Red line indicates start of OOS-Period. The light blue bands show the 15%- and 75% credible intervals for the MRF parameters.

5 Conclusions

Key alleged advantages of MRF methods were put to the test using a large British Macroeconomic dataset. After performing the prepossessing steps of making data stationary and setting up different input datasets, the predictive performance of two MRF methods was contrasted to that of a classical Time Series model and an *ordinary* Random Forest. For both frequencies under consideration and for most forecasting horizons, the MRF methods have proven to provide more accurate *OOS* forecasts. Indeed, the great flexibility enables an improved capturing and modelling the changing nature of economic relations depending on exogenous variables. This is highlighted by the ability to forecast well on unseen data. To test if an extension to the MRF, i.e. linear, surrogate trees, also provide explainable models for economic inference, these were build using the four *most important* variables from the ARRF-MRF models using the VI approach. A splitting procedure based on parameter stability testing ensured rather shallow and easily interpretable trees. For all dependent variables considered, the *OOS* predictive performance was close to that of of their ARRF counterparts. This strengthens the thesis that using some *important* variables, the underlying economic relationships can be approximated well. Also, all relevant variables derived can be directly linked to the dependent variable (or their auto-regressive relation) by economic reasoning. Upon analyzing the parameters in the final leaves, one notices that for some dependent variables, the intercept parameters are of greatest importance. The features to split by thus merely characterized economic states in which the future development of the dependent variables differs *per se* - hence regardless of current changes in these variables. As one would expect, in such cases, the predictive performance of the ARRF model was mostly close to that of the RF and MAF-RF model, which only rely on an intercept. For both linear trees constructed in (Coulombe, 2020), this issue does not occur. For other dependent features, the splitting also leads to significant parameters in the linear parts indicating differences in the auto-regressive relationships. Upon analyzing these linear trees, one can derive economically interesting patterns. However, as the linear trees merely reflect such patterns, one can only hypothesise how such patterns are linked to other economic features and fit to specific economic narratives. Based the analysis presented above, MRF methods and the corresponding surrogate trees seem to provide a useful extension to the econometric toolbox of economists and policy makers as forecasting in changing economic settings can be improved and non-linear relationships can be modelled in a flexible, yet interpreteable fashion. Nonetheless, one should not over-interpret the linear trees.

In a next step, more MRF models can be included in the forecasting study. Also, linear trees for different forecasting horizons and and for more dependent variables can be set up. Furthermore, adding specific regressors which are known to be related to the dependent features to the linear parts to construct more feasible VARRF models might provide further forecasting gains. In the linear trees, this might mitigate the issue of only finding statistically significant intercept parameters in some cases. Lastly, as MRF methods are rather new, these results should be validated using Macro-economic data from other countries. Above all, a replication study focused on less developed countries (than the US and the UK) would put the alleged advantages to a test.

References

Athey, S. (2019). The economics of artificial intelligence: An agenda. In A. Agrawal, J. Gans, , & A. Goldfarb (Eds.), (pp. 507–548). Chicago: University of Chicago Press.

Bank of England. (2022). *Monetary policy.* Retrieved 2022-06-15, from https://www.bankofengland.co.uk/monetary-policy

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

Buckmann, M., Andreas, J., & Robertson, H. (2021). Data science for economics and finance. In S. Consoli, D. R. Recupero, & M. Saisana (Eds.), (chap. Opening the Black Box: Machine Learning Interpretability and Inference Tools with an Application to Economic Forecasting). Cham: Springer.

Coulombe, P. G. (2020). The macroeconomy as a random forest. *arXiv preprint arXiv:2006.12724*.

Coulombe, P. G., Marcellino, M., & Stevanovic, D. (2021). Can machine learning catch the covid-19 recession? *arXiv preprint arXiv:2103.01201*.

D'Agostino, A., Gambetti, L., & Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of applied econometrics*, *28*(1), 82–101.

Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of business and Economics Statistics*, *13*, 253–265.

Durlauf, S., & Blume, L. (2016). *Macroeconometrics and time series analysis.* Springer.

Federal Reserve Bank of St. Louis. (2022). *OECD based Recession Indicators for the United Kingdom from the Peak through the Trough (GBRRECDM).* Retrieved 2022-05-31, from https://fred.stlouisfed.org/series/GBRRECDM

Fusion Media Limited. (2022). *Investing.com: United kingdom 1-year bond yield overview.* Retrieved 2022-05-04, from https://www.investing.com/rates-bonds/uk-1-year-bond-yield

Levich, R. (2012). The evidence and impact of financial globalization. academic press. In G. Caprio (Ed.), (chap. Interest Rate Parity). Academic Press.

MacKinnon, J. G. (2006). Bootstrap methods in econometrics. *Economic Record*, *82*, 2–18.

Matthew, T., Chen, C. S., Yu, J., & Wyle, M. (2015). Bayesian and empirical bayesian forests. *International Conference on Machine Learning*, 967–976.

Meek, C., Chickering, D. M., & Heckerman, D. (2002). Autoregressive tree models for time-series analysis. *Proceedings of the 2002 SIAM International Conference on Data Mining*, 229–224.

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

Neusser, K. (2016). *Time Series Econometrics.* Springer.

Potts, D., & Sammut, C. (2005). Incremental learning of linear model trees. *Machine Learning*, *61*(1), 5–48.

Reserve Bank of Australia. (2022). *Bonds and the yield curve.* Retrieved 2022-06-15, from `https://www.rba.gov.au/education/resources/explainers/bonds-and-the-yield-curve.html`

Taddy, M., Gardner, M., Chen, L., & Draper, D. (2016). A nonparametric bayesian analysis of heterogenous treatment effects in digital experimentation. *Journal of Business Economic Statistics*, *34*(4), 661–672.

Tucci, M. P. (1995). Time-varying parameters: a critical introduction. *Structural Change and Economic Dynamics*, *6*(2), 237–260.

United Kingdom Debt Management Office. (2022). *About Gilts.* Retrieved 2022-06-15, from `https://www.dmo.gov.uk/responsibilities/gilt-market/about-gilts/`

Zeileis, A., & Hornik, K. (2007). Generalized m-fluctuation tests for parameter instability. *Statistica Neerlandica*, *61*(4).

Zeileis, A., Hothorn, T., & Hornik, K. (2008). Model-based recursive partitioning. *Journal of Computational and Graphical Statistics*, *17*(2), 492–514.

Zeileis, A., Hothorn, T., & Hornik, K. (2010). party with the mob: Model-based recursive partitioning in r. *R package version 0.9-9999*.

Appendices

A Further Parameter Comparison Graphs

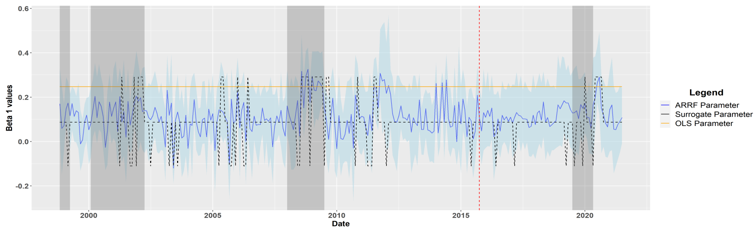

Figure 12: Comparison of β_1 parameters for **Unemployment Rate**. See figure 20 for more info.

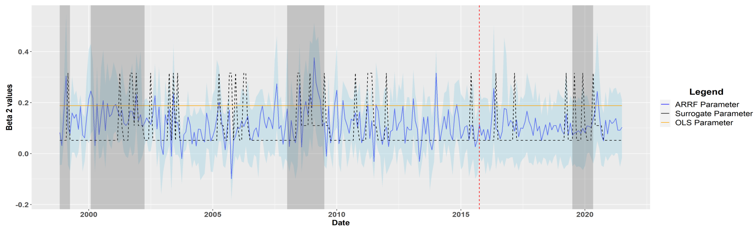

Figure 13: Comparison of β_2 parameters for **Unemployment Rate**. See figure 20 for more info.

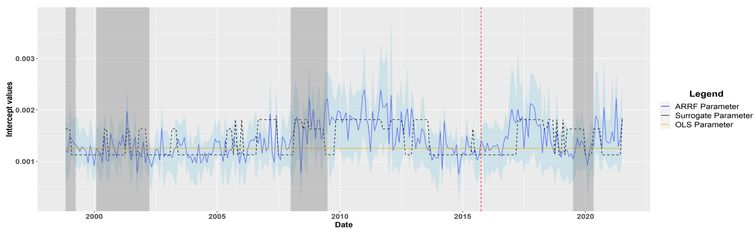

Figure 14: Comparison of α parameters for **CPI Inflation**. See figure 20 for more info.

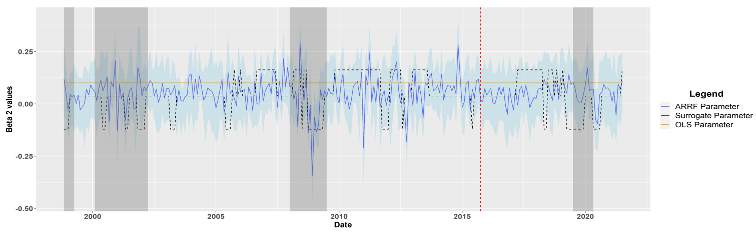

Figure 15: Comparison of β_2 parameters for **CPI Inflation**. See figure 20 for more info.

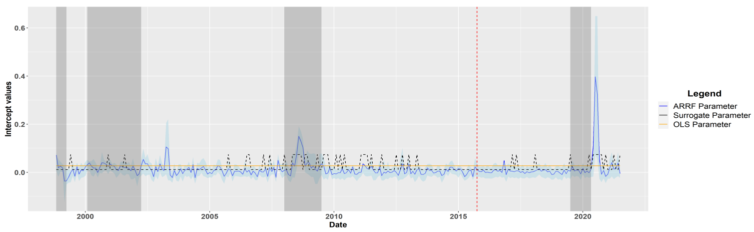

Figure 16: Comparison of α parameters for **1 Year GVT Bond Yield**. See figure 20 for more info.

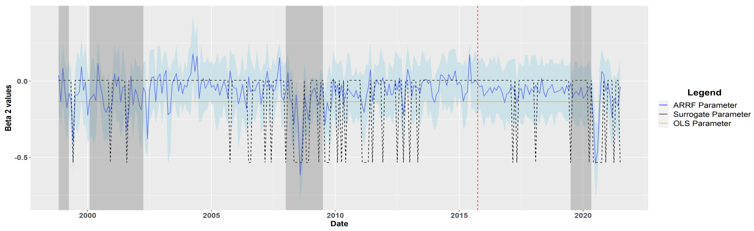

Figure 17: Comparison of β_2 parameters for **1 Year GVT Bond Yield**. See figure 20 for more info.

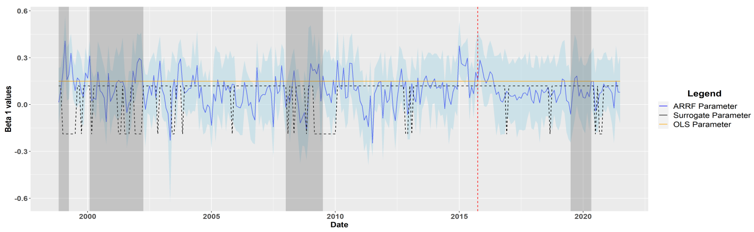

Figure 18: Comparison of β_1 parameters for **SPREAD**. See figure 20 for more info.

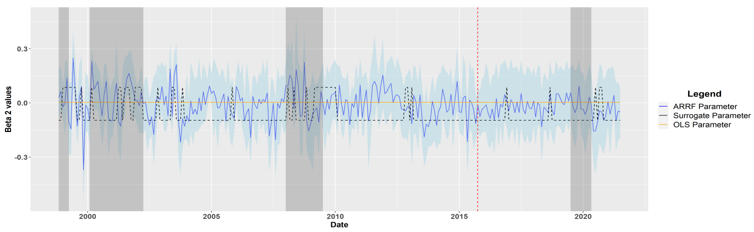

Figure 19: Comparison of β_2 parameters for **SPREAD**. See figure 20 for more info.

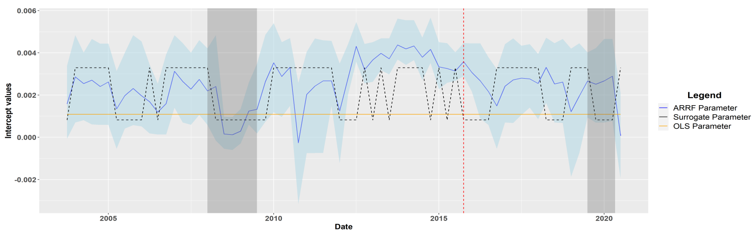

Figure 20: Comparison of α parameters for **GDP**. See figure XX for more info. Shaded areas indicate recessions. Red line indicates start of OOS-Period. The light blue bands show the 15%- and 75% credible intervals for the MRF parameters.

B Bayesian Block Bootstrap and Bayesian Inference

As has been stated in the methodology part, the trees are fitted on the respective bootstrap samples (made up of blocks) and then evaluated on all data (Coulombe, 2020). The GTVP used for inference are those that result from trees which did not use a block to which the evaluated triplet belongs - hence only a subset of all trees. This is called *Out-of-Bag (OOB)* inference (see methodology). In the following explanation, firstly only one tree is considered. Because of the randomness of the (block-)bootstrapping, some blocks may have occurred more often than others in the trees considered for *OOB* inference. Per block-triplet Z_b, the frequency of occurring $N_b = \sum_{b=1}^{B} 1(Z_b = z_b)$ is associated with a relative weight θ_b which describes the importance for the inference (per evaluated data point). The (posterior) distribution of the resulting *weight vector* given the frequencies $\pi(\theta|z)$ can be decomposed into likelihood, prior and evidence. The likelihood distribution of the occurrences of the blocks given the weights $\pi(z|\theta)$ is Multinomial. The conjugate prior $\pi(\theta)$ on the weights and hence the posterior $\pi(\theta|z)$ are Dirichlet distributions. When assuming equal α parameters in the prior Dirichlet distribution and independence of the blocks (which is reasonable in this case), the following holds for the single weight parameters: $\theta_b \sim exp(\alpha)$ (Taddy et al., 2016). Coulombe (2020) further suggests using improper and uninformative priors through choosing $\alpha_b = 0 \; \forall \, b$, which leads to a flat prior Dirichlet distribution with probability 0 for each θ. In that case, the single posterior weight vectors are again exponentially distributed: $\pi(\theta_b|z) \sim exp(\alpha+1)$ (Taddy et al., 2016). With $\alpha = 0$, $\pi(\theta_b|z) \sim exp(1)$ results. However, one is not actually interested in θ, but a functional of the latter $T(\theta)$. For an MRF tree, such a functional are the GTVP per tree (and per data point) (Coulombe, 2020).

Now, multiple trees are present in an MRF. The resulting distribution over all (*OOB*-) parameters is thus the empirical equivalent of repeated sampling from $T(\theta)$ based on $\pi(\theta_b|z) \sim exp(1)$ in each tree (Matthew et al., 2015). The posterior mean are the MRF GTVP and credible intervals can be deduced. Importantly, no actual sampling of θ_b from $exp(1)$ occurs, but this theoretical framework construct allows for a Bayesian interpretation of OOB inference on parameters - hence the name *Block Bayesian Bootstrap.* All findings above also apply to the *OOS* inference with the extension that the parameters of all trees are used since none of them were built on *OOS* data (Coulombe, 2020).

C The Parameter Stability test

The parameter stability test alluded to when explaining the surrogate trees is explained in greater detail here. It is based on the broad class of generalized M-fluctuation test (Zeileis & Hornik, 2007). The exact approach used is that described in Zeileis et al. (2008). Firstly, the test is explained for one splitting variable only and the combination of different splitting variables for the splitting decision is highlighted later on. Per splitting variable and having $i = 1, ...N$ observations, it is tested whether any associated varying parameter might not be constant.

$$H_0 : \beta_i = \beta_0 \;\; \forall \; i = 1, ..., N \tag{9}$$

The alternative hypothesis would state that for at least one parameter $\beta_i \neq \beta_0$ holds. β might be multivariate and in the present case is three-dimensional (Zeileis & Hornik, 2007). The test is based on the distribution of the score function ψ, which is the derivative of the SSR of the associated linear model with respect to the parameters β, where $\beta \in \mathbb{R}$ (see figure 10) (Zeileis et al., 2008). Other functions linked to other estimation procedures such as the negative Log-Likelihood for Maximum Likelihood estimation can also be envisioned, but Zeileis et al. (2010) suggest using the derivatives of the SSR since the underlying function in the splitting procedure is an OLS estimation.

$$\psi(Y, \beta) = \frac{\partial \, SSR(Y, \beta)}{\partial \beta} \tag{10}$$

Let $i = 1, ..., N$ denote the observations ordered by time and $j = 1, .., J$ denote the variables to split by (here $J = 4$). The empirical score function evaluated at the estimated optimal parameters $\hat{\beta}$ then is: $\hat{\psi}_i = \psi(Y_i, \hat{\beta})$. While $\sum_{n=1}^{N} \psi(Y_i, \beta) = 0$ by definition, the more the $\hat{\psi}_i$ are dispersed around 0, the more unstable the corresponding parameter can be deemed. More importantly, one should examine if the fluctuations around 0 might be explained by one of the splitting Variables $Z_1, ..., Z_J$ (Zeileis et al., 2008). The *empirical fluctuation process* $W_j(t)$ encapsulates such deviations. It is (per splitting Variable Z_j) defined as

$$W_j(t) = \hat{J}^{-1/2} n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} \hat{\psi}_{\sigma(Z_{ij})} \quad with \quad \hat{J} = \frac{1}{n} \sum_{i=1}^{n} \psi(Y_i, \hat{\beta}) \psi(Y_i, \hat{\beta})^T \tag{11}$$

$W_j(t)$ is three-dimensional in the ARRF(2) case. Note that $\hat{J}$ is thus the variance-covariance matrix of $\hat{\psi}$. Also, $\sigma(Z_{ij})$ denotes the anti-rank of all $\hat{\psi}$ ordered by the splitting variable under consideration. This ordering of $\hat{\psi}$ by the latter and the cumulative nature of the process enable an analysis of potential differences in the fluctuations by values of Z_j (Zeileis & Hornik, 2007). Should a systematic fluctuation by the exogenous variable be present, then $\hat{\psi}_{\sigma(Z_{ij})}$ would systematically systematically lie above/below for 0 for many subsequent i. Up to a certain t, this would carry through into $W_j(t)$. As $W_j(t)$ is defined for continuous time with t ranging from 0 to 1, $\lfloor nt \rfloor$ denotes the last fraction of the observations under consideration ($\frac{i}{n}$). A central limit theorem applies to $W_j(t)$ in the sense that in the case of H_0 holding, it converges to a (multivariate) Brownian Bridge W^0, which is a Wiener Process for which at times $t = 0$ and $t = T$, the process must take on the same value (Zeileis et al., 2008; Zeileis & Hornik, 2007).

With the help of a scalar mapping $\lambda(\cdot)$, $W_j(t)$ and it's limiting distribution can me reduced to a useful test statistics (per splitting variable) (Zeileis et al., 2008). Should the values of for the first fraction of associated Z_j values be systematically different from 0, then the cumulative nature of $W_j(t)$ and thereby $\lambda(W_j(\cdot))$ would capture this leading to higher/lower values of the latter. Under H_0, no systematic difference in the $\hat{\psi}$ values by any variable could be observeable. Zeileis et al. (2008) propose using the λ_{supLM} approach, which is applicable for numerical variables Z_j, which are present in the analysis at hand.

$$\lambda_{supLM}(W_j) = \underset{i=1,\ldots,n}{max} \left(\frac{i}{n} * \frac{n-i}{i}\right) \left\| W_j\left(\frac{i}{n}\right) \right\|_2^2 \quad (12)$$

While Zeileis et al. (2008) raise the possibility of not using all n observations, Zeileis et al. (2010) suggest using all $i = 1, ..., N$. The limiting distribution of $\lambda_{supLM}(W_j)$ is $sup_t(t(1-t)^{-1} \left\| W^0(t) \right\|_2^2$. The corresponding rejection regions and p-values can be derived from the latter. Should $\lambda_{supLM}(W_j)$ not fall into the rejection regions, it is being assumed that no parameter instability can be explained by the variable Z_j.

Should this result be the same when considering all potential variables to split by, the splitting process is stopped (Zeileis et al., 2008). The surrogate trees were implemented using R's *partykit* package (Zeileis et al., 2010). Zeileis & Hornik (2007) provide more technical details on the testing procedure.

D The Dataset used

D.1 ADF-tests for new variables

Data	Best Lag	Before transformation		After transformation	
	Order	ADF	P-value	ADF	P-value
1 Year GVT Bond Yield (2,1,1,1)	1	-0.55	0.86	-9.32	0.01
Real GDP (3,1,1,1)	1	-1.47	0.79	-5.32	0.01
Real Gross Capital Formation (3,1,1,1)	1	-1.47	0.79	-5.37	0.01
Real Government Consumption (2,1,1,1)	1	-1.37	0.56	-3.89	0.01
Real Private Consumption (3,1,1,1)	1	-1.42	0.81	-2.63	0.01
Growth Rate of Building Permits (1,-,-,-)	1	-5.98	0.01	-	-
Num. Buildings under Construction (2,1,1,1)	2	-2.19	0.26	-4.81	0.01
Changes in Inventories (1,1,1,1)	1	-4	0.01	-8.52	0.01
Housholds - Financial Assets (3,1,1,1)	1	-2.62	0.32	-5.07	0.01
Housholds - Liabilities (2,1,1,1)	1	-1.61	0.48	-9.05	0.01
Housholds - Debt to Income (2,1,1,1)	3	-2.66	0.09	-2.32	0.02
Real Governmment Debt (3,1,1,1)	2	-2.87	0.22	-6.56	0.01
Real Change in GVT (2,1,1,1)	1	-1.68	0.45	-7.46	0.01
Real Government Revenue (2,1,1,1)	2	-2.94	0.05	-15.55	0.01

Table 10: ADF-Tests Before and After Transformations for Data not included in UK MD. The values in brackets indicate: Type of ADF-Test on original Data [1: no drift, no trend/ 2: drift, no trend/ 3: drift, trend], Type of ADF Test on transformed data, Number of Lags, Number of Differences (both referring to the differencing to achieve stationarity). The maximal lag order of the tests was 10. The choice of the ADF-test was guided by examining the respective series, as suggested in Neusser (2016, p. 147f.) .

D.2 Overview over all used series

The following tables 11 to 14 show all data used in the analysis. This includes all data of the UK-MD dataset and all further features. The tables can provide more precise details on the series like those constituting the respective *important variables* to spit by. Also, the columns give meaning the the pre-processing steps explained in section 3.2. Consult the caption of table 14 for details on the respective columns.

Original Freq.	Name	Description	Group	Source	Pre-transf.	Stat.-Trans.	Agg. Type
monthly	EMP_TEMP	LFS: Temporary employees: UK: All: Thousands: SA	3	ONS	0	5	mean
monthly	UNEMP_RATE	Unemployment rate (aged 16 and over, SA)	3	ONS	0	2	mean
monthly	UNEMP_DURA_6mth	LFS: Unemployed up to 6 months: UK: All: Aged 16 and over: Thousands: SA	3	ONS	0	5	mean
monthly	UNEMP_DURA_6.12mth	LFS: Unemployed over 6 and up to 12 months: UK: All: Aged 16+: Thousands: SA	3	ONS	0	5	mean
monthly	UNEMP_DURA_12mth.	LFS: Unemployed over 12 months: UK: All: Aged 16 and over: Thousands: SA	3	ONS	0	5	mean
monthly	UNEMP_DURA_24mth.	LFS: Unemployed over 24 months: UK: All: Aged 16 and over: Thousands: SA	3	ONS	0	5	mean
monthly	EMP_RATE	Employment rate (aged 16 to 64, SA)	3	ONS	0	2	mean
monthly	EMP_ACT	LFS: Economically Active: UK: All: Aged 16-64: Thousands: SA	3	ONS	0	5	mean
monthly	EMP_ACT_RATE	LFS: Economic activity rate: UK: All: Aged 16-64 (%): SA	3	ONS	0	2	mean
monthly	CLAIMS	Claimant Count : K02000001 UK : People : SA : Thousands	3	ONS	0	5	mean
monthly	CLAIMS_RATE	Claimant Count : K02000001 UK : People : SA : Percentage (%)	3	ONS	0	2	mean
monthly	TOT_WEEK_HRS	LFS: Total actual weekly hours worked (millions): UK: All: SA	7	ONS	0	5	mean
monthly	AVG_WEEK_HRS	LFS: Avg actual weekly hours of work: UK: All workers in main & 2nd job: SA	7	ONS	0	5	mean
monthly	AVG_WEEK_HRS_FULL	Average actual weekly hours of work for full-time workers (SA)	7	ONS	0	5	mean
monthly	AWE_ALL	(Average Weekly Earning) AWE: Whole Economy Level (£ per Person) SA	7	ONS	0	5	mean
monthly	AWE_CONS	AWE: Construction Level (£ per Person): SA Total Pay Excluding Arrears	7	ONS	0	5	sum
monthly	AWE_MANU	AWE: Manufacturing Level (£ per Person): SA Regular Pay Excluding Arrears	7	ONS	0	5	sum
monthly	AWE_PRIV	AWE: Private Sector Level (£ per Person): SA Regular Pay Excluding Arrears	7	ONS	0	5	sum
monthly	AWE_PUB	AWE: Public Sector Level (£ per Person): SA Total Pay Excluding Arrears	7	ONS	0	5	sum
monthly	AWE_SERV	AWE: Services Level (£ per Person): SA Total Pay Excluding Arrears	7	ONS	0	5	sum
monthly	VAC_TOT	UK Vacancies (thousands) - Total	3	FRED	0	5	mean
monthly	VAC_CONS	UK Job Vacancies (thousands) - Construction	3	ONS	0	5	mean
monthly	VAC_MANU	UK Job Vacancies (thousands) - Manufacturing	3	ONS	0	5	mean
monthly	IOP_PROD	(Index of Production) IOP: B-E: PRODUCTION: CVMSA	2	ONS	0	5	mean
monthly	IOP_CAP_GOOD	IOP: MIG-CAG:Main Industrial Groupings - Capital Goods: CVMSA	2	ONS	0	5	mean
monthly	IOP_ENER	IOP: MIG-NRG:Main Industrial Groupings - Energy: CVMSA	2	ONS	0	5	mean
monthly	IOP_GOOD	IOP: MIG-COG:Main Industrial Groupings - Consumer Goods: CVMSA	2	ONS	0	5	mean
monthly	IOP_INT_GOOD	IOP: MIG-IG:Main Industrial Groupings - Intermediate Goods: CVMSA	2	ONS	0	5	mean
monthly	EMP	Number of People in Employment (aged 16 and over, SA)	3	ONS	0	5	mean
monthly	EMP_PART	LFS: In employment: Part-time: UK: All: Thousands: SA	3	ONS	0	5	mean
monthly	IOP_DUR	IOP: MIG-CD:Main Industrial Groupings - Consumer Durables: CVMSA	2	ONS	0	5	mean
monthly	IOP_MACH	IOP: CK:Manufacture of machinery and equipment n.e.c.: CVMSA	2	ONS	0	5	mean
monthly	IOP_MANU	IOP: C:MANUFACTURING: CVMSA	2	ONS	0	5	mean

Table 11: See Table 14 for an explanation.

Original Freq.	Name	Description	Group	Source	Pre-transf.	Stat.-Trans.	Agg. Type
monthly	IOP_MINE	IOP: B:MINING AND QUARRYING: CVMSA	2	ONS	0	5	mean
monthly	IOP_NON_DUR	IOP: MIG-CND:Main Industrial Groupings - Consumer Non-Durables: CVMSA	2	ONS	0	5	mean
monthly	IOP_PETRO	IOP: CD:Manufacture of coke and refined petroleum product: CVMSA	2	ONS	0	5	mean
monthly	IOP_OIL_EXTRACT	IOP: 06:Extraction Of Crude Petroleum And Natural Gas: CVMSA	2	ONS	0	5	mean
monthly	IOS	(Index of Services) IoS: Services: Index-1dp	2	ONS	0	5	mean
monthly	IOS_45	IoS: 45: Wholesale And Retail Trade And Repair Of Motor Vehicles And Motorcycles	2	ONS	0	5	mean
monthly	IOS_46	IoS: 46: Wholesale trade except of motor vehicles and motorcycles: Index-1dp	2	ONS	0	5	mean
monthly	IOS_47	IoS: 47: Retail trade except of motor vehicles and motorcycles: Index-1dp	2	ONS	0	5	mean
monthly	IOS_G	IoS: G: Wholesales, Retail and Motor Trade: Index-1dp	2	ONS	0	5	mean
monthly	IOS_EDUC	IoS: O-Q: PAD, Education and Health Index-1dp	2	ONS	0	5	mean
monthly	IOS_PNDS	IoS: H-N and R-U: PNDS: Private Non-Distribution Services: Index-1dp	2	ONS	0	5	mean
monthly	RSI	(Retail sales index) RSI:Volume SA:All Retailers inc fuel	5	ONS	0	5	mean
monthly	CAR_REGIS	Car registration: Passenger cars for the United Kingdom, Number, SA	5	FRED	0	5	mean
monthly	RETAIL_TRADE_INDEX	Total Retail Trade in the United Kingdom, Index 2015=100,SA	5	FRED	0	5	sum
monthly	AVG_WEEK_RETAIL_SALE	All retailing including automotive fuel, VALUE SA - Average Weekly	5	ONS	0	5	sum
monthly	AVG_WE_RET_SALE_NF	Retailing Predominantly non-food stores, VALUE SA - Average Weekly	5	ONS	0	5	mean
monthly	CPIH_ALL	ALL ITEMS 2015=100, consumer price inflation with occupiers' housing costs	6	ONS	0	5	mean
monthly	CPI_ALL	CPI INDEX 00: ALL ITEMS 2015=100	6	ONS	0	5	mean
monthly	CPI_EX_ENER	CPI INDEX: Excluding energy (SP) 2015=100	6	ONS	0	5	mean
monthly	CPI_GOOD	CPI INDEX: Goods 2015=100	6	ONS	0	5	mean
monthly	CPI_DUR	CPI INDEX: Durables (G) 2015=100	6	ONS	0	5	mean
monthly	CPI_NON_DUR	CPI INDEX: Non-durables (G) 2015=100	6	ONS	0	5	mean
monthly	CPI_SERV	CPI INDEX: Services 2015=100	6	ONS	0	5	mean
monthly	CPI_CLOTH	CPI INDEX: Clothing & footwear goods (G) 2015=100	6	ONS	0	5	mean
monthly	CPI_TRANS	CPI INDEX 07 : TRANSPORT 2015=100	6	ONS	0	2	mean
monthly	RPI_ALL	RPI All Items Index: Jan 1987=100	6	ONS	0	5	mean
monthly	RPI_GOOD	RPI: All Goods (Jan 1987=100)	6	ONS	0	5	mean
monthly	RPI_SERV	RPI: All Services (Jan 1987=100)	6	ONS	0	5	mean
monthly	RPI_HOUSE	RPI: Housing (Jan 1987=100)	6	ONS	0	5	mean
monthly	EXP_TOT	Total Trade (TT): WW: Exports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean
monthly	EXP_GOOD	Trade in Goods (T): WW: Exports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean
monthly	IMP_ALL	Total Trade (TT): WW: Imports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean
monthly	IMP_GOOD	Trade in Goods (T): WW: Imports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean
monthly	EXP_FUEL	Trade in Goods: Fuels (3): WW: Exports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean

Table 12: See Table 14 for an explanation.

Original Freq.	Name	Description	Group	Source	Pre-transf.	Stat.-Trans.	Agg. Type
monthly	IMP_FUEL	Trade in Goods: Fuels (3): WW: Imports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean
monthly	EXP_OIL	Trade in Goods: Crude oil (33O): WW: Exports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean
monthly	IMP_OIL	Trade in Goods: Crude oil (33O): WW: Imports: BOP: CVM: SA in million pounds	1	ONS	0	5	mean
monthly	EXP_MACH	Trade in Goods: Machinery and Transport (7): WW: Exports: in million pounds	1	ONS	0	5	mean
monthly	IMP_MACH	Trade in Goods: Machinery and Transport (7): WW: Imports:in million pounds	1	ONS	0	5	mean
monthly	EXP_METAL	Trade in Goods: Metal ores & scrap (28): WW: Exports: in million pounds	1	ONS	0	5	mean
monthly	IMP_METAL	Trade in Goods: Metal ores & scrap (28): WW: Imports: in million pounds	1	ONS	0	5	mean
monthly	EXP_CRUDE_MAT	Trade in Goods: Crude Materials (2): WW: Exports: million pounds	1	ONS	0	5	mean
monthly	IMP_CRUDE_MAT	Trade in Goods: Crude Materials (2): WW: Imports: million pounds	1	ONS	0	5	mean
monthly	GBP_BROAD	Monthly average Broad Effective exchange rate index, Sterling (Jan 2005 = 100)	11	BOE	0	2	mean
monthly	GBP_CAN	Monthly average Spot exchange rate, Canadian Dollar into Sterling	11	BOE	0	2	mean
monthly	GBP_EUR	Monthly average Spot exchange rate, Euro into Sterling (XUMAERS)	11	BOE	0	2	mean
monthly	GBP_JAP	Monthly average Spot exchange rate, Japanese Yen into Sterling	11	BOE	0	2	mean
monthly	GBP_US	Monthly average Spot exchange rate, US$ into Sterling (XUMAUSS)	11	BOE	0	2	mean
monthly	OIL_PRICE	Crude Oil Prices: Brent - Europe, Dollars per Barrel	6	BOE	0	5	mean
monthly	BANK_RATE	Monthly average of official Bank Rate (IUMABEDR)	8	BOE	0	2	mean
monthly	CONS_CREDIT_ex_student_loan	Outstanding of total (excluding the Student Loans Company) sconsumer credit lending, SA	9	BOE	0	5	sum
monthly	TOT_LENDING_APP	Monthly number of total sterling approvals for secured lending to individuals SA	9	BOE	0	5	sum
monthly	TOT_HOUSE_APP	Monthly number of total sterling approvals for house purchase to individuals SA (LPMVTVX)	9	BOE	0	5	sum
monthly	MORT_FIXED_RATE_5YRS	Interest rate of UK financial institutions (excl. Central Bank) 5 year fixed rate mortgage to households	9	BOE	0	2	mean
monthly	MORT_FIXED_RATE_2YRS	Interest rate of UK financial institutions (excl. Central Bank) 2 year fixed rate mortgage to households	9	BOE	0	2	mean
monthly	M1	Amounts outstanding of monetary financial institutions' sterling and all foreign currency M1	9	BOE	0	5	sum
monthly	M2	Monthly amounts outstanding of monetary financial institutions' sterling and all foreign currency M2	9	BOE	0	5	sum
monthly	M3	Monthly amounts outstanding of monetary financial institutions' sterling and all foreign currency M3	9	BOE	0	5	sum
monthly	M4	amounts outstanding of M4 (financial institutions' M4 liabilities to private sector)	9	BOE	0	5	sum
monthly	LIBOR_3mth	3-Month London Interbank Offered Rate (LIBOR), based on British Pound, Percent, Monthly	8	ECB	0	2	mean
monthly	BGS_5yrs_yld	Monthly average yield from British Government Securities, 5 year Nominal Par Yield	8	BOE	0	2	mean
monthly	BGS_10yrs_yld	Monthly average yield from British Government Securities, 10 year Nominal Par Yield	8	BOE	0	2	mean
monthly	BGS_20yrs_yld	Monthly average yield from British Government Securities, 20 year Nominal Par Yield	8	BOE	0	2	mean
monthly	FTSE_ALL	UK FTSE All Share (F̂TAS)	12	YAHOO!	0	5	mean
monthly	FTSE250	FTSE 250 (F̂TMC)	12	YAHOO!	0	5	mean
monthly	VIX	CBOE Volatility Index (V̂IX)	12	YAHOO!	0	1	mean
monthly	SP500	S&P 500 (ĜSPC)	12	YAHOO!	0	5	mean
monthly	UK_focused_equity	iShares MSCI United Kingdom ETF (EWU)	12	YAHOO!	0	5	mean
monthly	EUR_UNC_INDEX	Economic Policy Uncertainty Index for Europe, Index, Monthly, Not SA	14	FRED	0	2	mean
monthly	BCI	Business confidence index (BCI)Amplitude adjusted, Long-term average = 100	14	OECD	0	2	mean
monthly	CCI	Consumer confidence index (CCI)Amplitude adjusted, Long-term average = 100	14	OECD	0	2	mean
monthly	CLI	Composite leading indicator (CLI)Amplitude adjusted, Long-term average = 100	14	OECD	0	2	mean

Table 13: See Table 14 for an explanation.

Original Freq.	Name	Description	Group	Source	Pre-transf.	Stat.-Trans.	Agg. Type
monthly	PPI_MANU	Producer price indices (PPI)Manufacturing, domestic market, 2015=100	6	OECD	0	5	mean
monthly	PPI_MACH	PPI Machinery and Equipment N.E.C. for Domestic Market (G6VG)	6	ONS	0	5	mean
monthly	PPI_OIL	PPI Coke and Refined Petroleum Products for Domestic Market (G6ST)	6	ONS	0	5	mean
monthly	PPI_METAL	PPI Basic Metals for Domestic Market (G6SZ)	6	ONS	0	5	mean
monthly	PPI_MOTOR	PPI Motor Vehicles, Trailers and Semi-Trailers for Domestic Market (G6WH)	6	ONS	0	5	mean
monthly	SPREAD	10 Year Gov't Bond Yield - Bank Rate	8	calculation	0	1	mean
monthly	BGS_1yrs_yld	Monthly average yield from British Government Securities, 1 year Nominal Par Yield	8	Investing.com	0	2	mean
quarterly	gdp	Real Gross Domestic Product for United Kingdom (Million Pounds - 2010 Chained, SA)	1	FRED	0	5	-
quarterly	inv	Gross Fixed Capital Formation in United Kingdom (Pounds, SA)	1	FRED	3	5	-
quarterly	gvt_con	Government Final Consumption Expenditure in United Kingdom (Pounds, SA)	1	FRED	3	5	-
quarterly	pr_con	Private Final Consumption Expenditure in the United Kingdom (british pounds, chained 2015, SA)	1	FRED	1	5	-
quarterly	dw_con	Dwellings and Residential Buildings Permits Issued for Construction for the United Kingdom (Growth Rate)	4	FRED	0	1	-
quarterly	dw_start	Total Buildings under Construction, Started for the United Kingdom (Num. Of Buildings)	4	FRED	0	5	-
quarterly	ch_inv	Real Changes in Inventories for Great Britain (Pounds)	5	FRED	2	4	-
quarterly	hh_assets	Households: Total financial assets (current prices - Million Pounds)	10	ONS	0	1	-
quarterly	hh_liab	Households: Total net acquisition of financial liabilities (current prices - Million Pounds)	10	ONS	0	1	-
quarterly	hh_debt_ti	Households: Debt to Income Ratio	10	ONS	0	2	-
quarterly	gvt_debt	General government condolidated gross debt (Maastricht) - Million Pounds	13	ONS	2	5	-
quarterly	chg_gvt_debt	General Government Gross Debt:Change in Debt (MillionPounds)	13	ONS	2	1	-
quarterly	gvt_rev	Government: Total Revenue (Million Pounds)	13	ONS	2	5	-

Table 14: Together with Tables 11, 12, 13: Overview of all data used. *Group* indicates overarching group (1: NIPA, 2: Industrial Production, 3: Employment, 4: Housing, 5: Inventories, Orders and Sales, 6: Prices, 7: Earning Productivity, 8: Interest Rates, 9: Money Credit, 10: Household Balance, 11: Exchange Rates, 12: Stock Markets, 13: Non-household balance, 14: other (e.g. Sentiment Scores). *Pre-transf.* indicates the transformations performed on the series not included in UK MD (0: no change, 1: to 1M Pounds, 2: Nominal to Real, 3: 2 & 3). *Stat-Transf.* shows the transformation needed to make the series stationary (1: None, 2: First difference, 4: Logarithm, 5: 2 & 4). *Aggr.Type* displays the aggregation function to turn the raw, monthly data to quarterly data.