Erasmus University Rotterdam

Erasmus School of Economics

Bachelor Thesis BSc2 Econometrics and Economics

Inflation Everywhere: A Machine Learning Approach to Explain Changes in Oil pass-through across Europe

---

July 3, 2022

Author Name:
Rick Plat
Student Number:
492956

Supervisor:
Dr. A. Tetereva
Second assessor:
Dr. E.P. O'Neill

Abstract

This paper proposes a novel extension to the Macroeconomic Random Forest (MRF) algorithm of Coulombe (2021) to model inflation. More specifically, I augment the MRF splitting procedure to accommodate panel data across European countries to exploit commonality in the inflation patterns. We then use this model to assess which macroeconomic state variables are the most important determinants of short-term oil pass-through. Through a forecasting study, we document that pooled MRF outperforms various benchmark models, including plain RF, for periods with inflation of moderate magnitude. I proceed to investigate the pooled MRF output by means of variable importance analysis and surrogate trees on the path of the short-term oil pass-through coefficient. The results indicate that the fuel intensity, exchange rate, debt to GDP ratio, and output gap play a prominent role in the level of oil pass-through in a country.

Contents

1 Introduction

Understanding the empirical relationship between oil prices and inflation rates is crucial for monetary authorities aiming to keep inflation under control. In this regard, oil shocks pose a challenging trade-off for policy makers between either a higher inflation or a higher unemployment rate (Herrera & Pesavento, 2009; Bernanke et al., 1997). Understanding the effect of oil shocks on inflation has recently gained prominence again with the rapidly increasing energy prices and inflation rates across the world. While many past studies have focused on estimating the level of oil price pass-through into consumer prices (Hooker, 2002; Chen 2009), there has been relatively little research into what actually influences this level of oil pass-through. A better understanding of such economic state variables could help economies better manage oil shocks and the associated inflationary influences.

The vast majority of papers to date has focused on traditional time series models in estimating oil pass-through and analyzing what economic state variables are responsible for changes within countries, and differences across countries. However, recent advances in the interpretability of Machine Learning (ML) (Molnar, 2019) paired with increased data availability in macroeconomics make ML a viable new methodology to investigate this topic with. Recent studies show, for example, that Machine Learning has strong potential in modelling inflation dynamics. One such study is that of Medeiros et al. (2019), who document that Random Forest models can consistently outperform the random walk (RW), autoregressive (AR) and unobserved components stochastic volatility (UCSV) model benchmarks in inflation forecasting. This is particularly noteworthy as an extensive literature exists that documents these benchmarks are "exceedingly difficult to improve upon" (Stock & Watson, 2010).

Coulombe (2021) builds further upon these findings by proposing a novel extension to the random forest algorithm, called *Macroeconomic Random Forest* (MRF). The key difference of this algorithm in comparison to "plain" RF is that it performs a linear regression in each leaf of the regression trees, similar to the Local Linear Forests of Friedberg et al. (2019). Firstly, this facilitates interpretation of the model, as MRF now estimates economically meaningful coefficients. Secondly, Friedberg et al. (2019) show that such a linear regression improves the model fit of the algorithm in the case of (strong) smooth signals, which often exist in macroeconomic relationships. MRF is also able to nest many popular non-linearities and has the advantage of letting the data decide which one, or combination, is most appropriate. This stands in contrast with the previously considered time series models, in which the researcher has to employ their own judgement and decide beforehand which type of model to use. For instance, Ferry et al. (2001) model inflation with a regime switching Phillips curve model, while Hooker (2002) makes use of a structural break model, and Eliasson (2001) relies on a smooth transition model. MRF can nest each of these non-linear elements, which has been shown to improve model fit and forecasting performance for quarterly US inflation between 1960Q1 and 2015Q4 (Coulombe, 2021).

This paper presents an extension to the MRF framework, with the goal to assess which economic state variables are the main determinants of change in oil pass-through into consumer prices in Europe. The central idea of this extension is to exploit the commonalities in inflation rates across the similar countries by pooling the data after taking into account long-run differences. The MRF methodology is remarkably suited to such a pooled estimation approach, as rather than requiring identical coefficients over the whole sample period for all countries, it allows to assign those sub-periods of countries that are most similar in their coefficients to a common leaf. Only then does the algorithm proceed to estimate these coefficients. Hence, while the data in each leaf will likely exhibit a higher heteroskedasticity, there is potential to exploit a stronger signal. This way, a more data-rich regression can be achieved for each sub-period in time of every country than would otherwise be possible.

I start by constructing a dataset on CPI inflation, the oil price in euros per barrel and 108 features motivated by the economic literature as being relevant in estimating inflation and oil pass-through to grow the MRF. These include amongst others trade openness, the exchange rate, fuel intensity and the output gap of a country, motivated by Chen (2009) and Gelos & Ustyugova (2017). The dataset comprises quarterly data of 18 European countries from 2004Q1 up to 2021Q4. Importantly, for most of these countries the data availability is insufficient to consider applying MRF and many other ML methods directly. An initial inspection of the inflation rates of these countries shows highly similar patterns over time and high correlations. To make the pooled estimation more meaningful, I adjust each of these inflation series by their long-run mean to fully exploit the commonality in the inflation patterns.

Next, I augment the splitting procedure of MRF to accommodate for pooled data of several countries. This is achieved by splitting observations not only over time, but also over countries, such that country-time pairs are assigned to leafs. Additionally, I build a Phillips curve model augmented with lags for changes in the oil price. The formulation of this Phillips curve model is motivated by the research of Chen (2009), Musso et al. (2009) and Faust & Wright (2013). I use this model in the leafs of the regression trees that constitute the MRF. Finally, through variable importance analysis and the usage of surrogate trees I analyze which variables are the main determinants of changes in the oil pass-through over time for the countries in the sample.

Our main results are that pooled MRF outperforms all benchmark models, including plain RF, for periods with inflation of moderate magnitude. However, for periods with large and rapid changes in inflation, MRF cannot outperform the existing autoregressive and Phillips Curve models. The variable importance analysis and an inspection of the surrogate trees give strong indication that the fuel intensity, exchange rate, debt to GDP ratio, and output gap play a prominent role in the level of oil pass-through in a country. This is largely consistent with the findings of De Gregorio et al. (2008), Chen (2009) and Gelos & Ustyugova (2017). A difference with regards to Gelos & Ustyugova (2017), however, is that I find only a moderate effect for trade openness and the weights of food in the CPI inflation basket of a country.

The remainder of this paper is structured as follows. Section 2 provides a summary of the literature related to oil pass-through and random forests. Then, Section 3 describes the data and which transformations I apply to exploit the commonality in the inflation patterns. Section 4 outlines the augmented MRF methodology and how the Phillips curve will be modelled. Section 5 presents the estimation results, variable importance analysis and surrogate trees for the short-term oil pass-through. Finally, Section 6 provides a conclusion and formulates an answer to the research question.

2 Theoretical Framework

This section first outlines the existing literature on the effect of oil shocks on the inflation rate. Then, I provide a brief overview of the relevant literature on regression trees, random forests and local linear forests.

2.1 Determinants of oil pass-through

One of the first papers to analyze the causes for changes within countries and differences across countries in oil pass-through is De Gregorio et al. (2008). Using Phillips curve models for 34 countries, they find evidence that the oil pass-through has declined in recent decades, in particular for industrial countries. Their analysis shows that the reduction in oil pass-through is likely primarily attributable to a decrease in fuel intensity of many countries as well as a reduction in exchange rate pass-through.

Chen (2009) builds upon these findings and tries to explain the decline in oil pass-through further. In an analysis of 19 developed countries, he regresses the time-varying short-run pass-through coefficients of the countries on several economically characterizing variables and finds that the appreciation of the domestic currency, a more active monetary policy in response to inflation, and a higher degree of trade openness explain the observed decline in oil price pass-through. Contrary to De Gregorio et al. (2008), Chen (2009) finds no evidence that fuel intensity plays a significant role in oil pass-through.

The paper most similar in spirit to ours is that of Gelos & Ustyugova (2017). This paper takes a more systematic approach to determining what economic variables are associated with the level of oil pass-through. For a large set of countries for the period 2001-2010 they consider a range of 20 structural country characteristics and policies within these countries that may contribute to differences in oil pass-through. Then using Phillips curve and VAR models, they test which variables have a significant impact on the oil pass-through in a country. The set of variables considered by Gelos & Ustyugova (2017) contained amongst others the variables used by Chen (2009). They find that countries with higher shares of food consumption, fuel intensities, and pre-existing inflation levels experienced higher oil pass-through. Additionally, they find a small attenuating effect for inflation targeting on oil pass-through. These findings are not in line with those of Chen (2009), as trade openness and appreciation of the domestic currency are not found to be significant determinants of oil-pass through.

2.2 Regression Tree

Before discussing the Random Forest algorithm, it is first important to briefly describe what a *regression tree* is and how it can be obtained. In essence, there are two steps to creating a regression tree (Breiman et al., 1984):

1. Divide the set of all possible values for predictors $X_1, \dots, X_p$ (predictor space) into K regions that don't overlap. Let $R_1, \dots, R_K$ denote these regions.
2. Then the prediction for each observation in a particular region R_k is the average of the values that belong to that region.

In this procedure the regions R_k are usually obtained by applying a top-down, greedy algorithm. More specifically, a recursive splitting procedure is applied such that each parent node is split into two subsamples using the predictor X_j and a cut point c so that the greatest possible reduction in total sum of squared error is achieved (James et al., 2021). In mathematical notation, at each parent node we aim to solve the following minimization problem:

$$\min_{j \in \mathcal{J},\, c \in \mathbb{R}} \left[\min_{\beta_1} \sum_{\{i:\, x_i \in R_1(j,c)\}} (y_i - \hat{y}_{R_1})^2 + \min_{\beta_2} \sum_{\{i:\, x_i \in R_2(j,c)\}} (y_i - \hat{y}_{R_2})^2 \right], \tag{1}$$

where J denotes the set of all predictors, and $R_1(j,c)$ and $R_2(j,c)$ correspond to the region of the predictor space for which $X_j \leq c$ and $X_j \geq c$, respectively. $\hat{y}_{R_k}$ represents the prediction of subregion k, which is simply the average of the observations belonging to subregion k. Note that this procedure does not consider all possible partitionings of the predictor space, as this would be computationally too intensive. Instead, at every step it tries to split a given region into two subsets in a locally optimal way until a certain stopping condition is met.

After obtaining a regression tree, it is often recommended to prune some of the splits to prevent overfitting. This is important because while overfitting may result in a strong fit on the training data, the model will often perform poorly on a testing set. Additionally, pruning enhances interpretability of the tree and leads to lower variance of the predictions (James et al., 2021). This does often come at a (small) cost in increased bias. The pruning can be achieved by using *cost-complexity pruning* (Breiman et al., 1984; James et al., 2021). This method aims to find those subtrees $T \subset T_0$ that minimize the following cost-complexity function for a given cost-complexity parameter α:

$$\sum_{k=1}^{|T|} \sum_{x_i \in R_k} (y_i - \hat{y}_{R_k}) + \alpha |T|, \tag{2}$$

where $|T|$ denotes the number of terminal nodes of the subtree, and R_k denotes the subset of the predictor space corresponding to terminal node k. In this function $\alpha = 0$ yields the original tree T_0. The optimal parameter α can be obtained through K-fold cross-validation (James et al., 2021).

2.3 Random Forests

A widely used method in Machine Learning that builds on regression trees is Random Forest (RF) (Breiman, 2001). RF is an ensemble method, which means that it builds on multiple models to create its final prediction, rather than relying on a single model. This often produces more accurate, and in particular, less variant predictions. Random Forest, as the name suggests, relies on many regression tree models to obtain its final prediction.

2.3.1 Bagging & De-correlating

The two key concepts of RF are Bootstrap Aggregation, or simply *bagging*, and *de-correlating*. Bagging is used to reduce the variance of the estimates of a collection of trees. In practice, this can be achieved by training every tree on a different bootstrapped sample. More specifically, from the sample data one draws B new samples with replacement to create B regression trees. Then averaging the predictions of all these trees will often produce a highly accurate estimate with a variance lower than that of a single regression tree (Breiman, 2001). The number of bootstrapped samples B should be chosen sufficiently large to reduce the variance of the ultimate estimate to an acceptable level. James et al. (2021) suggest $B = 100$, as a rule of thumb in practice. In general, the gains in accuracy of the model become negligible as the number of trees considered becomes large.

However, the efficacy of the outlined approach may be jeopardized when the trees produce highly correlated estimates. This can happen, for example, when one predictor is considerably stronger than all the other predictors. In this case, it is likely that all trees will select this strong predictor for the first split of the tree, and consequently all trees will produce correlated estimates. Here, taking the average of many highly correlated estimates does not produce the desired reduction in the variance of the estimate. Random Forests tackle this problem by using de-correlated trees (Breiman, 2001). This means that each tree, beside using a different bootstrapped sample, considers a different set of predictors for each split. Specifically, every tree uses a random (small) subset of m predictors out of the total set of p available predictors at each splitting point. This way different trees give different predictors a chance to be considered for splitting the node. Ultimately, this reduces the variance of the average estimate across all trees. The fraction of m predictors out of the total set of predictors p that is chosen is a hyperparameter that is often referred to as `mtry`. Usually, this parameter is set at one-third for regression settings (Coulombe, 2021). Alternatively, it can be tuned by using k-fold cross-validation on a grid of possible values. An additional benefit of a Random Forest over a regular regression tree, is that in general it is not prone to overfitting once a sufficiently large number of trees B is used, contrary to a single tree that may need pruning (Coulombe, 2021). Thus, a Random Forest combines bagging and de-correlation to obtain a highly accurate estimate with a low variance. Its final prediction is the average of the predictions of each tree.

2.4 Local Linear Forests

While RF has been shown to be an effective method in a variety of fields and applications (Medeiros et al., 2019), one key drawback that remains is that RF cannot fit well a strong, smooth signal (Friedberg et al., 2019). More specifically, when RF is attempting to fit a smooth regression surface, it fails to exploit strong local trends and often fits an incorrect shape to the target. Friedberg et al. (2019) overcome this issue firstly by considering RF as a ensemble method that generates kernel weights, similar to Meinshausen (2006) and Athey et al. (2019). Then, once these weights are obtained, the final prediction follows from a weighted linear regression (WLS). In this regression, Friedberg et al. (2019) also fit a ridge penalty to prevent overfitting to the local trend and reduce variance of the final estimates. Hence, Local Linear Forest first recursively applies an augmented splitting procedure to obtain the terminal nodes L_b for each tree b, and then constructs from these leafs the following weights:

$$\alpha_t(x_0) = \frac{1}{B}\sum_{b=1}^{B}\frac{I(X_t \in L_b(x_0))}{|L_b(x_0)|}, \tag{3}$$

where $I(X_i \in L_b(x_0))$ is an indicator function. Then these weights are used in:

$$\forall t : \arg\min_{\beta_t}\left\{\sum_{t=1}^{T}\alpha(x_0)(Y_t - X_t\beta_t)^2 + \lambda||\beta_t||_2^2\right\} \tag{4}$$

to obtain the final estimates of the conditional mean. Friedberg et al. (2019) present evidence that Local Linear Forests outperform plain RF when strong smooth signals exist in both an application to wage estimation as well as various simulations.

3 Model Data & Features of MRF

The data for our empirical analysis is of a qaurterly frequency and comprises 18 European countries. The sample period extends from 2004-Q1 up to 2021-Q4 (72 observations per country, 1296 observations in total). Additionally, I use 108 unique features to grow the MRF. The data are collected from the OECD database, FRED-QD and the IMF Financial Statistics database.[1] The countries are selected based upon data availability and similarities in the time pattern of inflation. Initially, out of all European countries, 20

[1]The OECD database can be accessed online at https://data.oecd.org/. The FRED-QD is publicly available at the Federal Reserve of St-Louis's website. Lastly, the IMF Financial can be accessed online at https://data-imf-org.eur.idm.oclc.org/.

satisfied the data availability condition. Out of these 20 countries, we keep those countries that exhibit a correlation of 0.70 or higher with at least one other country in the sample. This is to ensure that the observations of the countries can be meaningfully pooled and then estimated. Because of this, Norway and Switzerland were excluded from the sample, which yields our ultimate sample of 18 European countries. The sample period is chosen to include several economic recessions as well as the recent peak in inflation starting in 2021-Q1, with the associated rise in the oil price in euros. This period is particularly relevant, as the oil pass-through may have changed here relative to the stable, low-inflation years before.

In our analysis we focus on inflation as measured by CPI, similar to Chen (2009) and Gelos & Ustyugova (2017). This is done firstly because revisions to CPI inflation are infrequent and often negligible, while for other inflation measures revisions can be large and include benchmark changes or changes in conceptual definitions (Faust & Wright, 2013). Secondly, the usage of CPI inflation accommodates the usage of CPI basket weights as features when growing the Macroeconomic Random Forest. These variables have been found to play an important role in oil pass-through by Gelos & Ustyugova (2017). In the Phillips curve models, we use the first lag of quarterly unemployment as a measure of economic slack, as is common in the literature (Stock & Watson, 2008; Faust & Wright, 2013). For the oil price terms in the Phillips curve, we use the price per oil barrel in euros, similar to Musso et al. (2009).

3.1 Data Characteristics

Figure 1 plots the quarterly inflation and the unemployment rate for the countries in our sample. We observe that all countries share important similarities in their inflation patterns over time. Firstly, all countries exhibit a large and rapid drop in inflation starting in 2008-Q2, during the financial crisis in this year. Secondly, as countries recovered, inflation rose back up gradually to levels slightly lower than in 2007. Thirdly, all countries experience a gradual decrease in inflation after approximately 2012-Q1, which has been associated with decreases in energy prices as well as increases in economic slack (Koester et al., 2021). Lastly, the inflation of each country increases noticeably after the reopening of economies and various fiscal stimuli in many European countries in 2021-Q1 following the Covid-19 pandemic.

In the unemployment rate curves of many countries, we observe a clear negative correlation with the inflation rates. This can be observed most noticeably overall in the period 2012Q1- 2019-Q1, where for many countries inflation is often a mirror image compared to the unemployment rate. See for example Belgium, Luxembourg and France. Importantly, there are also clear differences in the average levels of inflation and unemployment across countries. Countries such as Spain, Portugal and Greece experience historically high unemployment rates, often in excess of 15%, while countries like the Netherlands, Denmark and the United Kingdom, have much lower unemployment rates of around 5% on average.

To investigate the commonality in the inflation patterns further, Table 1 shows the correlations between the CPI inflation of the countries. We observe that almost all countries have at least one other country with which they exhibit a correlation above 0.75, while most countries also have several countries with which a correlation above 0.80 exists. The exceptions are Hungary, Sweden and The Netherlands. These countries, however, still exhibit a correlation of more than 0.70 with at least one other country in the sample. This gives indication that the pooled MRF approach may yield benefits in estimation, as overall most countries show high correlations in their CPI inflation rates.

Table 1: Correlations between CPI inflation of 18 sample European countries, 2004Q1-2021-Q4 (Correlations in excess of 0.75 or higher in bold)

	AUT	BEL	CZE	DEU	DNK	ESP	FIN	FRA	UK	GRC	HUN	IRL	ITA	LUX	NLD	PRT	SVN	SWE
AUT	1.00																	
BEL	**0.81**	1.00																
CZE	0.67	0.67	1.00															
DEU	**0.85**	0.70	**0.75**	1.00														
DNK	0.67	**0.75**	0.61	0.60	1.00													
ESP	**0.78**	**0.80**	0.66	**0.77**	**0.78**	1.00												
FIN	0.74	0.70	0.63	0.69	0.71	0.60	1.00											
FRA	**0.78**	**0.80**	0.71	0.74	0.74	**0.83**	0.64	1.00										
GBR	0.74	0.71	0.51	0.65	**0.77**	0.66	**0.77**	0.63	1.00									
GRC	0.53	0.69	0.44	0.47	**0.77**	0.74	0.47	0.70	0.62	1.00								
HUN	0.48	0.40	0.59	0.58	0.61	0.60	0.49	0.62	0.49	0.64	1.00							
IRL	0.62	0.63	0.58	0.70	0.46	0.72	0.63	0.63	0.42	0.43	0.42	1.00						
ITA	**0.77**	**0.82**	0.67	0.71	0.86	**0.86**	**0.76**	**0.84**	0.74	0.71	0.64	0.63	1.00					
LUX	**0.83**	**0.81**	0.66	**0.80**	**0.82**	**0.87**	0.70	**0.85**	**0.77**	0.71	0.60	0.66	**0.86**	1.00				
NLD	0.70	0.55	0.60	0.69	0.46	0.51	0.62	0.55	0.66	0.24	0.37	0.39	0.55	0.64	1.00			
PRT	0.65	**0.75**	0.48	0.60	0.74	**0.81**	0.62	**0.75**	0.59	0.69	0.50	**0.75**	0.83	**0.81**	0.28	1.00		
SVN	0.65	0.68	**0.80**	0.73	0.68	**0.76**	0.62	**0.81**	0.52	0.65	0.72	0.66	**0.79**	0.74	0.46	0.64	1.00	
SWE	0.67	0.73	0.65	0.68	0.55	0.57	0.73	0.62	0.71	0.46	0.38	0.56	0.56	0.67	0.56	0.54	0.54	1.00

Figure 2 plots the oil price per barrel between 2004Q1 and 2021Q4. When we compare this time series to those of the inflation rates of the countries, we observe a highly similar pattern. The oil price also peaks just before the financial crisis in 2008Q1, to the rapidly decrease and rebound to its previous level in the years following the financial crisis. Additionally, the oil price rapidly increases starting in 2020Q1. A difference between the oil price time pattern and that of the inflation rates of many countries is that the oil price decreases much more slowly following 2012Q1.

Next, I describe the most important features that will be employed by the Macroeco-

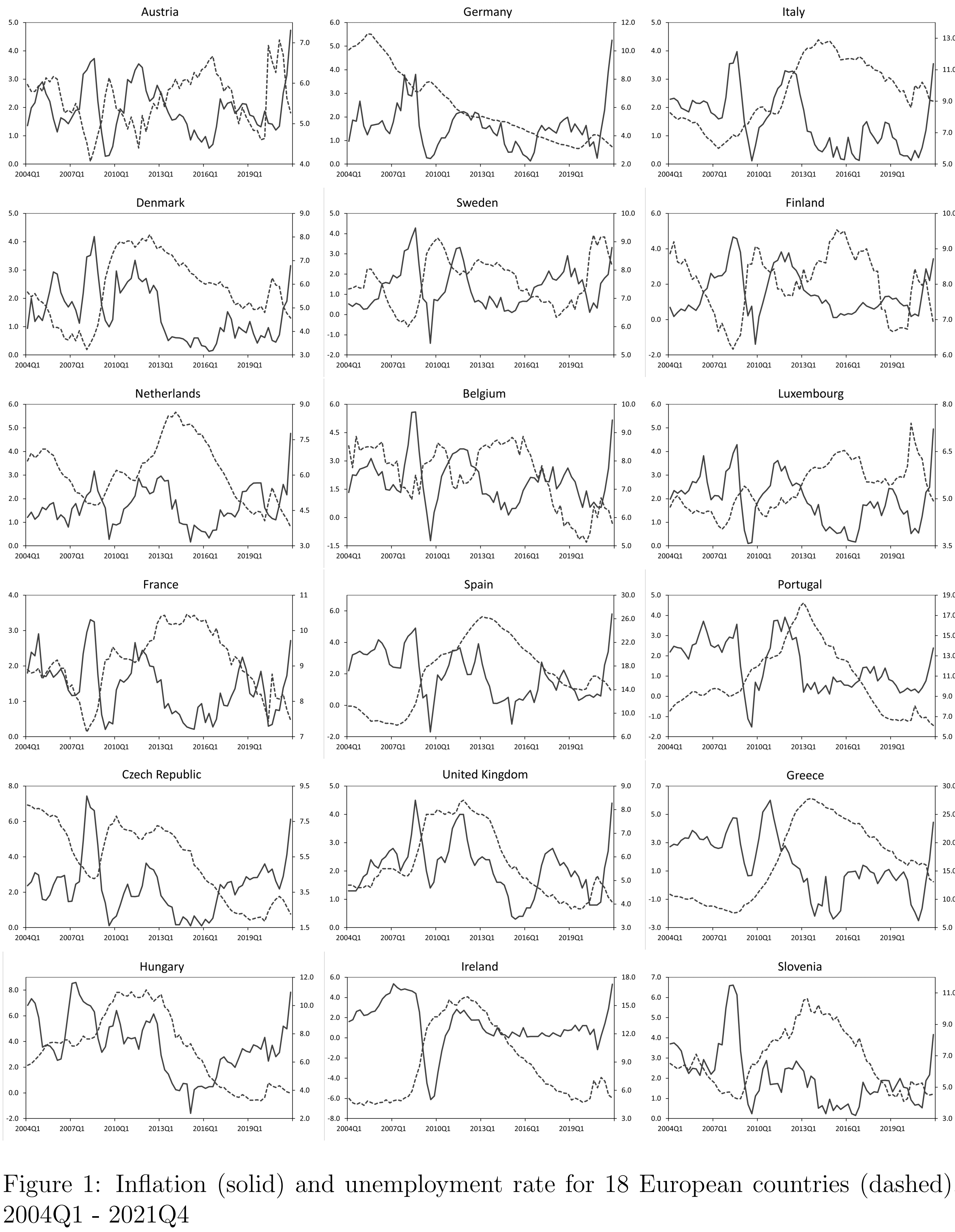

Figure 1: Inflation (solid) and unemployment rate for 18 European countries (dashed), 2004Q1 - 2021Q4

nomic Random Forest algorithm. For each of these I briefly explain the economic rationale for their relevance relating to changes in the oil pass-through. Afterwards, I explain the

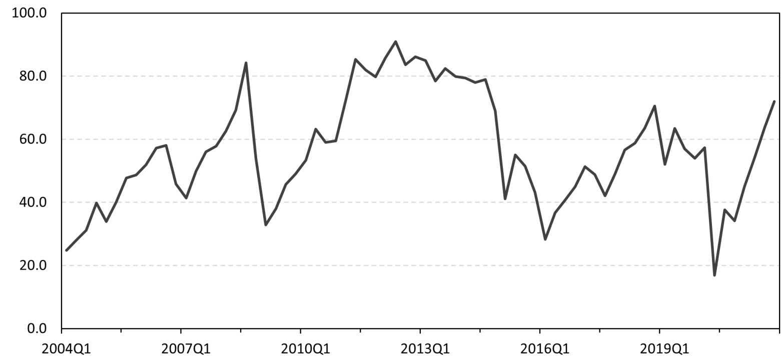

Figure 2: Oil price in euros per barrel, 2004Q1 - 2021Q4

importance of using Moving Average Factors in growing the MRF. Beside the features relevant to oil pass-through, various features are also of importance to modelling the relationship between inflation, its lags and the measure of economic slack (unemployment rate). These include amongst others the M1 and M3 money supply, the normalized capacity utilization of each country, the PPI, and short and long term interest rates for each country. See appendix A for a complete list of the features and associated transformations.

3.2 Most Important Features for Oil pass-through

- *Financial Development*: As Gelos & Ustyugova (2017) point out, a better developed financial system often allows for more effective monetary policy. Hence, a more developed financial system can allow a country to control the effects of an oil-shock better. We rely on the Financial Development index data of the IMF Financial Statistics database for this variable.

- *Food and Transport weights in CPI basket*: Food prices have been found to be positively correlated with oil prices and oil price uncertainty (Alghalith, 2010). Similarly, as many modes of transport run on oil-based fuels, transport costs are positively correlated with the oil price. Hence, oil pass-through into CPI inflation often positively correlates with the weights of Food and Transport in the CPI basket (Gelos & Ustyugova, 2017).

- *Fuel Intensity*: Countries that consume more fuel per capita are likely to be more sensitive to oil shocks (Hooker, 2002; Chen, 2009).

- *Trade openness*: A significant negative relationship has been found between a country's trade openness, as measured by the sum of exports and imports as a percentage of GDP, and a country's oil pass-through (Romer, 1993, Chen, 2009). More specifically, importing goods cheaply abroad can offset the inflationary effects of oil shocks. We use the sum of exports and imports as a percentage of GDP to measure trade openness.

- *Public debt/GDP*: Celasun et al. (2004) find that improvements in fiscal balances and fiscal credibility significantly decrease inflation expectations. Consequently, oil shocks can have a smaller effect on countries with lower public debt as a share of GDP.

- *Output gap*: This variable is often used as a measure of economic slack in inflation models. When the output gap is small, an oil shock will likely have a larger pass-through into consumer prices (Gelos & Ustyugova, 2017). To estimate this variable I rely on the Hodrick-Prescott filtered output gap, with a penalty parameter equal to 100, as is common in the literature.

- *Low inflation environment*: Taylor (2000) documents that in countries with a lower inflation level, producers are less likely to pass through changes in costs to consumers. The rationale here is that firms change their prices depending on how persistent they expect the change in costs to be. In low inflation environments changes in costs are generally believed to be less persistent (Chen, 2009; Gelos & Ustyugova, 2017).

- *Exchange rate*: Oil transactions between countries are usually conducted in terms of dollars. Hence, a shock in the oil price will often have larger consequences for inflation if the exchange rate with respect to the dollar depreciates (Chen, 2009; Gelos & Ustyugova, 2017).

3.3 Moving Average Factors

As Coulombe (2021) points out, one drawback of random forests in time series applications is the difficulty it faces when extracting the relevant information from many lags of a feature. In the current application of inflation modelling, for example, it could be that the first 8 lags of capacity utilization are relevant due to sticky prices or rigidities in the labour market. A random forest may then spend a split first on the 2nd lag, then the 5th lag, followed by a split on the 3rd, etc.

Coulombe et al. (2021) propose instead to consider a weighted average of the first P lags of a variable that best capture the temporal pattern of the specific feature. This can be achieved by applying principal component analysis to these P lags of a feature (Coulombe et al., 2021). This has two advantages: first it facilitates a more straightforward interpretation. Secondly it has been shown that the inclusion of such Moving Average Factors (MAFs) improves the root mean squared error (RMSE) of the model in predicting on both long and short horizons (Coulombe et al., 2021). Hence, we apply PCA to various lag polynomials for which we believe the temporal pattern could be of relevance. Here, the availability of a sufficient number of lags for each feature is sometimes binding (See Appendix A). Similar to Coulombe (2021), we use $P = 8$ lags for quarterly data, which corresponds to two years.

4 Methodology

This section introduces pooled Macroeconomic Random Forests (MRF). First, I describe the Phillips curve model I use in the leafs of MRF, and how it can be augmented to exploit commonality across the European countries. Next, I explain how the MRF algorithm works and how I adapt it to pooled data. Finally, I provide an explanation of how variable importance analysis and surrogate trees will be used to analyze the MRF output and to determine which economic state variables determine changes in the oil pass-through of European countries.

4.1 Pooled Phillips Curve

A model that is commonly used in the literature to investigate oil pass-through is that of a Phillips curve with added lags of the change in the oil price (Chen, 2009; Musso et al., 2009; Faust & Wright, 2013). This motivates the following general model specification:

$$\pi_t = \alpha + \sum_{i=1}^{p} \delta_i \pi_{t-i} + \gamma u_{t-1} + \sum_{i=1}^{p} \theta_i \Delta o_{t-i} + \varepsilon_t, \tag{5}$$

where π_t denotes the CPI inflation of a country at time t, α is an intercept and u_{t-1} represents the first lag of the unemployment rate of a country. Lastly, Δ denotes the first difference and o_t denotes the log of the price of an oil barrel in euros. From this model the short-term oil pass-through can be obtained simply as θ_1. The long-run oil pass-through can be computed as $\phi = \sum_{i=1}^{p} \theta_i / (1 - \sum_{i=1}^{p} \delta_i)$. We use $p = 4$ lags, as this minimizes the Schwarz Information Criterion (SIC) for the vast majority of countries we consider.

Next, it is important to augment this model in such a way that it allows for meaningful pooled estimation. Clearly, a common intercept across countries is too restrictive. This is because each country has a different long-run level of inflation (See Section 3.1). Hence, we first adjust the Phillips curve model of Equation (3) to take into account varying long-run levels of inflation. We can achieve this by subtracting a long-run inflation factor π_{t-1}^{LR} from each country's respective inflation series at every point in time. This method has been shown to work well in pooled estimation of volatility across asset classes (Bollerslev, 2017). We use the expanding window sample mean from the start of the sample up until quarter $t-1$ as the long-run inflation π_{t-1}^{LR}. This gives the following augmented Phillips Curve model:

$$\pi_t - \pi_{t-1}^{LR} = \sum_{i=1}^{4} \delta_i(\pi_{t-i} - \pi_{t-1}^{LR}) + \gamma u_{t-1} + \sum_{i=1}^{4} \theta_i \Delta o_{t-i} + \varepsilon_t. \tag{6}$$

In the remainder of the paper we denote the model of Equation (6) as

$$y_t = X_t \beta_t + \varepsilon_t, \tag{7}$$

where $y_t = [\pi_t - \pi_{t-1}^{LR}]$ and X_t denotes the 1×9 row vector of the regressors in Equation (6). β_t represents the associated 9×1 column vector of coefficients in Equation (6). This is for the purpose of brevity and ease of notation when describing the splitting procedure.

4.2 General Model MRF

Next I describe the general MRF framework. The key difference between MRF and plain RF, is that the regression trees upon which the model builds, do not simply predict the average of the observations in the particular leaf. Instead, MRF performs a regression in each leaf, similar to the Local Linear Forests of Friedberg et al. (2019), and so shifts the focus of the random forest towards predicting the beta coefficients of this regression, rather than the dependent variable. Additionally, as previously discussed (Section 2.4), this linear regression allows the MRF model to better capture strong smooth signals, which often exist in macroeconomic relationships (Coulombe, 2021). This gives the following general model (Coulombe, 2021):

$$\begin{aligned} y_t &= X_t \beta_t + \varepsilon_t \\ \beta_t &= \mathcal{F}(S_t) \end{aligned}$$

where $\mathcal{F}$ denotes a random forest and S_t is the set of features we consider for splitting. In our case, these are economic characteristics of countries that determine the time variation in β_t (See Section 3.2 and Appendix A for more information). In general, it holds that $X_t \subset S_t$. Hence, the regressors used within the leafs are also features themselves. By performing in each leaf the regression above, MRF obtains General Time Varying Parameters (GTVP) (β_t). Following Coulombe (2021) and Friedberg et al. (2020), the general tree splitting procedure becomes:

$$\min_{j \in \mathcal{J}^-,\, c \in \mathbb{R}} \left[\min_{\beta_1} \sum_{\{t \in l | S_{j,t} \leq c\}} (y_t - X_t\beta_1)^2 + \lambda \, ||\beta_1||_2^2 + \min_{\beta_2} \sum_{\{t \in l | S_{j,t} > c\}} (y_t - X_t\beta_2)^2 + \lambda \, ||\beta_2||_2^2 \right]. \quad (8)$$

In this equation $\mathcal{J}^-$ denotes the random subset of regressors of S_t that are used in the particular tree split under consideration. Furthermore, l represents the parent node we aim to split. The goal of this minimization problem is to find the optimal regressor S_j for $j \in \mathcal{J}^-$ by which to split the parent node l and at which cut point value c of the regressor S_j this split should be made. We apply this minimization problem recursively to the two split samples produced by the splitting procedure. This continues until a certain stopping criterion is met. This ultimately produces a tree. Note that Equation (8) also contains a ridge penalty represented by the λ term. This term is included to prevent overfitting to a possible strong trend and reduce the variance of the parameter estimates (Friedberg et al., 2019).

In general, MRF aims to use trees with a very high depth (Coulombe, 2021). While for normal regression trees it is then recommend to prune some of the terminal nodes to prevent overfitting and improve interpretability (Breiman, 1984), this is not necessary for a sufficiently diversified ensemble of trees. More specifically, Goulet Coulombe (2020) presents evidence that the out-of-sample prediction of a "non-pruned" RF is identical to that of an optimally pruned one, when the trees are sufficiently diversified. In our case this means that if a reasonably large number of trees B is chosen paired with a reasonable fraction of predictors at each split (also called `mtry`, see Section 2.3.1 on decorrelating trees), the trees will not be over-split. Accordingly, the model will not predict time-variation when it is not there (Coulombe, 2021). The final prediction of MRF is the average of all β_t predictions over the B trees.

In comparison to plain RF, MRF inherits many desirable properties. Firstly, MRF like RF has relatively few tuning parameters, which Coulombe (2021) argues are of little importance to the performance and robustness of the algorithm. Secondly, MRF performance is relatively unaffected by the inclusion of many irrelevant features in S_t (Friedman et al., 2001). Thirdly, with a sufficiently large `mtry`, the model can handle sparsity in the features and so discard less meaningful predictors (Coulombe, 2021).

4.3 Augmenting the splitting rule

Next, the single country splitting procedure needs to be adapted to accommodate for the pooled inflation data. The MRF framework can be remarkably suited to a pooled estimation approach, as rather than requiring coefficients to be the same over the whole time series of every country, we let the algorithm decide based on the data and features which country's coefficients are most similar during specific sub-periods in time. This stands in constrast with a normal pooled OLS approach, where the coefficients of the countries need to be highly similar across the whole time period under consideration.

The goal of the original splitting procedure of Equation (8) was to find at each node the best variable S_j out of the random subset of predictors J^- to split the quarterly observations across two leafs, and at which value c of that variable the split should occur. Hence, fundamentally, observations corresponding to time indices t were assigned to leafs. However, after pooling the data, each quarter t exists for every country in the sample. Therefore, the splitting procedure should now not only split according to time index t but also according to country q. This is simply because quarter t of one country does not necessarily have to be in the same leaf as the quarter t of another per se. For this purpose we consider country-quarter pairs (q, t) that will be assigned to leafs. Additionally, we need to consider that each feature $S_{j,t}$ exists for every country q. Hence, this becomes $S_{j,q,t}$. Jointly, this gives the following augmented splitting rule:

$$\begin{aligned} \min_{j \in \mathcal{J}^-,\, c \in \mathbb{R}} \Bigg[& \min_{\beta_1} \sum_{\{(q,t) \in l \mid S_{j,q,t} \leq c\}} (y_{q,t} - X_{q,t}\beta_1)^2 + \lambda \, ||\beta_1||_2^2 \\ & + \min_{\beta_2} \sum_{\{(q,t) \in l \mid S_{j,q,t} > c\}} (y_{q,t} - X_{q,t}\beta_2)^2 + \lambda \, ||\beta_2||_2^2 \Bigg]. \end{aligned} \tag{9}$$

The goal of this problem is to find the best variable S_j out of the random subset of predictors J^- to split the sample with, and at which value c of that variable the split should occur. Now, however, the variable S_j can take values c across countries. In practice, we implement this minimization problem by first sorting the unique values of $S_{j,q,t}$ from high to low for each j, and then considering values c at intervals of a length of 5% of the number of unique values $S_{j,q,t}$. Then for each value c, we split the observations over two subsets as detailed in Equation (9) and perform two ridge regressions on the subsets. Again, the ridge penalty is included to prevent overfitting and to decrease the variance of the parameter estimates (Friedberg et al., 2019). We use 5-fold cross-validation to tune the ridge parameter λ. For this purpose I use the `cv.glmnet` function from the `glmnet` package, which performs the optimization problem of finding the appropriate λ parameter (Friedman et al., 2022).

To stop the splitting procedure I use two stopping criteria. Firstly, the `minimum node size` for a leaf to be considered for splitting is 140 observations. Secondly, for a split to be considered valid, it should have at least 6 observations per regressor (`minimum leaf fraction (MLF)` = 6). Both of these stopping criteria are higher than the recommendation of Coulombe (2021) to use `minimum node size` = 10 for quarterly data and a `MLF` between 1 and 2. This is for three reasons. First, pooling the data leads to a considerably larger number of available observations than would otherwise occur using quarterly data of a single country. More specifically, after pooling we have 1296 observations for all variables, while for a country with 50 years of quarterly data (which is often unavailable), one would only have 200 observations. Therefore, the tree already attains sufficient depth by stopping splitting earlier. Secondly, Coulombe (2021) focuses on small models with few explanatory variables, such that estimation is possible with fewer observations. This paper investigates a relatively larger model with nine parameters (compared to four parameters of Coulombe (2021)). Hence, a larger minimum number of observations is justified to facilitate estimation. Thirdly, Coulombe (2021) employs a random walk regularization in estimation, that includes the first two lags and first two forwards of each observation t in a leaf. This means that often in regressing, effectively many more observations are used than is implied by the minimum node size and `MLF`. In a pooled setting, however, including lags and forwards of each t of a country is less meaningful. Hence, this regularization is not included. This necessitates a larger minimum node size and minimum leaf fraction for meaningful estimation. Based on trial runs with `MLF` $\in \{20, 30, 40, 50, 60, 70, 80, 90\}$, I find that MRF performance does not improve appreciably for `MLF` ≥ 60. For lower `MLF` we find a sharp increase in model performance based on both in-sample fit and out-of-sample forecasting performance.

Lastly, in growing the MRF I use a fraction `mtry` $= 0.25$ of the total number of available features in determining each split. This is slightly lower than the recommended value of one third in the literature (James et al., 2021). However, based on experimentation, and in line with Coulombe (2021), we find that MRF performance does not change appreciably when considering `mtry` $\in \{0.1,\ 0.25,\ 0.33,\ 0.5\}$, while computational burden is strongly affected. A possible explanation for this is that macroeconomic data often has a factor structure. For example, if for a certain split the variable *unemployment* is not selected, there are many other correlated variables such as *output gap* or other labour indicators that will produce similar results. This is because these variables often represent a similar latent factor (Coulombe, 2021).

Note that, ultimately, the algorithm is not optimally tuned. However, as Coulombe (2021) points out, this is generally not problematic for Random Forest algorithms as often the performance gains from optimizing the tuning parameters are "miniscule". Hence, most parameters are set based on theoretical arguments and a limited number of trial runs that assess both in-sample fit, and prediction accuracy for those observations not used in growing the trees (due to sampling with replacement). For this purpose I evaluated the root mean squared error in both cases.

4.4 Assessing relative performance of Pooled MRF

To assess the relative performance and viability of Pooled MRF, we consider both the in-sample fit and a forecasting exercise in which we forecast the period 2018Q1-2021Q4. This period includes part of the recent peak in inflation and energy prices and is chosen to assess the performance of the model both in a period of stability as well as rapid change in inflation. In the forecasting exercise we compare pooled MRF to the following models:

- The Phillips curve model of Equation (6), where the coefficients are estimated once based on the training set of 2004Q1-2018Q4 for each country separately.
- A Random Walk (RW) model considered by Faust and Wright (2013) in which $\pi_{t+1} = \pi_t$. This model is found to outperform the Atkeson & Ohanian (2001) model, which is often used as a benchmark for out-of-sample inflation forecasting.
- A plain Random Forest model that is grown using the same features, `mtry` and number of trees as pooled MRF. This model also uses the pooled data.
- An AR(4) model that is estimated on the training set 2004Q1-2018Q4 for each country and then makes one-quarter ahead forecasts. Faust & Wright (2013) use a highly similar model as their benchmark for inflation forecasting.

Here the pooled MRF and plain RF models are estimated only once based on the 2004Q1-2018Q4 pooled data. We compare the models by assessing differences in the Root Mean Squared error, which can be computed as:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}. \tag{10}$$

As for each country only 16 out-of-sample observations exist, we compare the predictive accuracy of the models using a small-sample adjusted Diebold-Mariano test (Harvey et al., 1997). This adjusts for the asymptotic results used in the usual Diebold-Mariano test (Diebold & Mariano, 1995). In this test we consider the squared estimation error as the loss function. Furthermore, we test the two sided null hypothesis of equal predictive accuracy. The adjusted test statistics S^* can be computed based on the usual Diebold Mariano test statistic S as:

$$S^* = \left(\frac{n - 1 - 2h + n^{-1}h(h-1)}{n} \right) S, \tag{11}$$

where n is the number of observations we forecast, and h is the forecasting horizon. In our case this gives $n = 16$ and $h = 1$. The test statistic follows a Student's t distribution with $(n-1)$ degrees of freedom.

4.5 Interpreting MRF

Generally, Random Forests are regarded as black box models and, hence, often require external devices to facilitate interpretation (Molnar, 2019). While MRF improves interpretability by providing economically meaningful and time-varying coefficients, understanding what drives this time-variation, similar to RF for the dependent variable y_t, still requires external devices. In this section I outline how the MRF output can be analyzed to answer what economic variables are the main determinants of changes in oil pass-through. I focus on Variable Importance (VI) as well as surrogate trees.

4.5.1 Variable Importance

The central idea of Variable Importance in random forests is analyzing how the predictive accuracy of the model changes on a testing set when we replace a feature S_j by a random permutation of this feature.

The measure of VI we consider is that of VI_{OOB}, where the observations not used in growing the trees (due to sampling with replacement) of the RF, also called out-of-bag, are used as the testing set (Breiman, 2001). To analyze the VI, we first compute the Root Mean Squared Error (RMSE) of this testing set before permuting any feature. We can compute the RMSE as

$$RMSE = \sqrt{\frac{1}{\#OOB} \sum_{q \in Q} \sum_{t \in OOB(q)} (y_{q,t} - \hat{y}_{q,t})^2}, \tag{12}$$

where Q is the set of all countries, $OOB(q)$ denotes the set of out-of-bag observations for country q and $\#OOB$ denotes the total number of out-of-bag observations. Then we permute each feature j one at a time and compute the RMSE again. Finally, the VI for a tree can be computed as the percentage reduction in RMSE when the variable is included compared to when it is randomly permuted. The VI measure for the pooled Random Forest is the average of VI across all trees. A key drawback of this method is that when features are highly correlated, VI can experience difficulty in deciding which of the two features is most important. Therefore, the results of this analysis should be interpreted with caution.

4.5.2 Surrogate Trees

Another way to assess which variables are most important for oil pass-through is by growing a so-called *surrogate tree* on the path of the $\beta_{k,t}$ coefficients that relate to short-term oil pass-through (Ribeiro et al., 2016). The goal of this surrogate tree is to gauge which features are most important in explaining the path of the coefficients. Hence, the surrogate tree focuses on explaining the model rather than the data (Coulombe, 2021). We use a cost-complexity parameter of 0.04 in growing this surrogate tree. This parameter is set to take into account the trade-off between the fit of the surrogate tree on the GTVP path and interpretation. In short, a surrogate tree is a regression tree that tries to replicate the MRF model predictions of $\beta_{k,t}$ by only using a handful of leafs.

5 Results

This section describes the findings relating to which variables are most important in the inflation forecasts of pooled MRF, and in the path of short-term oil pass-through coefficient. As discussed, I first assess the in-sample fit and out-of-sample forecasts of pooled MRF relative to four benchmark models. Next, I present the results of the variable importance analyses. Finally, we use the features most relevant for oil pass-through to grow a surrogate tree.

5.1 Relative Performance MRF

Figure 3 plots the in-sample predictions of pooled MRF in comparison to the PC model and true inflation for France, Ireland and Hungary. These countries are chosen to illustrate the strengths and weaknesses of pooled MRF most strikingly, although similar patterns exist for the other countries (See Appendix B for the in-sample predictions of each country).

For France, we observe that both pooled MRF and the augmented PC model capture the time-pattern of the adjusted inflation well. Both models manage to fit the peaks and lows of the true inflation, albeit with a slight delay due to the autoregressive nature of the models. A notable difference between pooled MRF and the Phillips curve model appears to be that the pooled MRF forecasts are considerably more smooth than those of the augmented PC model. This can be most clearly seen for the period 2016Q1-2018Q4, where both the actual inflation and the PC model exhibit small peaks and lows, while pooled MRF fits a smooth averaging line through the pattern. As the augmented PC model is in essence an AR model, this often means that the fit is worsened, as the model still attempts to capture the previous period peak, while this period true inflation has dropped again.

For Ireland, we observe a key weakness of pooled MRF. More specifically, we observe that pooled MRF exhibits difficulty in capturing extreme absolute values of inflation, where augmented PC does not. Following the financial crisis of 2008, Ireland experienced a period of deflation, which produced a deep trough for the adjusted inflation pattern. Pooled MRF does not fit this decrease timely and fully. A potential explanation for this is that too few observations exist with such a rapid and deep decline in inflation across the countries. Hence, the observations corresponding to Ireland in the period 2008Q1-2010Q1 are estimated in a common leaf with observations that have a less extreme decline in inflation. The result is that the coefficients for Ireland in this specific period are not of the magnitude required to fit the rapid decline in inflation, and hence the trough is never fully captured.

For Hungary, it can be observed that both Pooled MRF and the augmented PC model fit the inflation pattern well. This is an interesting result, because in Section 2 we noted that Hungary had the lowest correlation with the other countries in the sample. The strong fit of Hungary shows the flexibility of pooled MRF for inflation forecasting in two ways. Firstly, even when the inflation pattern of a country exhibits low correlation with those of another country, if at specific points in time the coefficients correspond to those of other countries at different points in time, pooled MRF can still produce a good fit. Secondly, if a country has no countries that ever exhibit a similar autoregressive structure, MRF can assign the observations of the specific country to an isolated leaf for estimation. An inspection of the MRF trees, however, shows that this was not the case for Hungary. Therefore, the first explanation is more probable.

Overall, an inspection of all three countries gives indication that pooled MRF performs best when inflation is changing in a gradual manner. This is in line with the findings of Coulombe (2021) for regular MRF. Specifically, he presents evidence that MRF performs best when a time-series exhibits strong persistence. A possible explanation for the relatively strong performance of pooled MRF in such cases is that many observations exist for which inflation is changing only gradually, such that the coefficients can be estimated well. Additionally, in periods of gradual change, inflation exhibits altering small peaks and lows, which pooled MRF can smooth out nicely. In particular AR models struggle with such small altering ups and downs, due to the lagged capturing of either movement (See France 2016Q1-2018Q1).

Figure 3: Adjusted inflation (green), pooled MRF prediction (orange) and augmented PC prediction (blue) for France, Ireland and Hungary, 2004Q1-2021Q4

Next, we compare the pooled MRF model to the augmented PC model, an AR(4) model, an adapted version of the Atkeson and Ohanian (2001) random walk model, as used by Faust & Wright (2013), and plain RF. First, we compute the predictive accuracy of all models using the RMSE. Second, we compare the predictive accuracy of the models for the period 2018Q1-2021Q4, using a small-sample adjusted Diebold Mariano test (Harvey et al., 1997). This period poses a considerable challenge for pooled MRF, as the rapidly increasing inflation period of 2021Q1-2021Q4 is included in the sample. Therefore, the results should be interpreted as being conservative estimates of the performance of pooled MRF in general.

Table 2 shows the RMSE for each model and each country for the one quarter ahead forecasts corresponding to 2018Q1-2021Q4. We observe that for most countries, the pooled MRF model does not differ significantly in its predictive accuracy in comparison to the augmented PC, plain RF and AR (4) models. The PC, RF and AR(4) models outperform pooled MRF for 5, 6 and 6 countries, respectively. Pooled MRF does, however, appear significantly more accurate than the random walk model. The relatively poor performance of this models is likely attributable to short forecast horizon of one-quarter ahead (Faust & Wright, 2013). Hence, overall the AR(4) model and the augmented Phillips curve perform best, followed by RF and the Pooled MRF model and finally the RW model.

Table 2: Root Mean Squared Errors for predicting adjusted inflation one quarter ahead over period 2004Q1-2021Q4

Country/Model	**P. MRF**	**Aug. PC**	**RW**	**RF**	**AR(4)**
AUT	0.57	0.27	0.75	0.43	0.29
BEL	1.11	0.96	1.89*	0.81	0.66
GER	1.05	0.57	2.12*	0.93	0.61
FIN	0.54	0.46	4.43***	0.44	0.33
FRA	0.30	0.23	1.41***	0.29	0.23
DNK	0.62	0.40	0.51	0.25***	0.25***
GBR	0.56	0.14*	1.68	0.42	0.39
ITA	0.81	0.18***	3.31***	0.32***	0.20***
LUX	0.77	0.51	1.43*	0.60	0.62
NLD	0.68	0.46	2.55***	0.60	0.55
SWE	0.48	0.22	3.21***	0.51	0.37
ESP	1.69	2.46	2.48	1.04**	0.81**
CZE	1.28	0.67	2.70*	1.13	0.54
PRT	1.00	0.28***	2.02**	0.21***	0.16***
GRC	3.45	1.04***	3.86	1.44***	0.86***
HUN	1.39	1.85	4.66*	1.45	1.68
IRL	1.69	0.54*	2.49	1.26	1.31
SVN	1.22	0.82	1.83**	0.39***	0.46**

Note. * indicates $p < 0.1$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$.

Interestingly, the pooled MRF model appears to perform best for Hungary. This is particularly remarkable, because Hungary is the country that has the least commonality in its inflation pattern in comparison to the other countries. A potential explanation for this is that for this country the rapid rise in inflation of 2021, set in slightly later.

While for most countries inflation started to increase in 2020Q4, Hungary's inflation remained approximately equal between 2020Q4 and 2021Q1. This in line with the findings previously discussed, that pooled MRF performs best for inflation of moderate magnitude and with moderate change.

Lastly, as we observed the performance of pooled MRF to be particularly strong in capturing the inflation levels that are not extreme in magnitude, we also consider the out-of-sample predictive performance over all countries based on the magnitude of the adjusted inflation. To achieve this we split the out-of-sample data into quartiles and compute RMSE and small-sample adjusted DM test for each quartile. In each of these tests we compare the predictive accuracy of pooled MRF to those of the other 4 models.

Table 3 presents the RMSEs and associated small-sample DM test results. We observe that the augmented PC and AR(4) models are significantly more accurate when predicting values of inflation in the highest and lowest quartiles. Interestingly, however, Pooled MRF performs equally well as the augmented PC model for observations of inflation in the second quantile, and outperforms all models significantly at the 1% level for inflation observations in the third quartile. More specifically, pooled MRF attains a RMSE 80% smaller than the next best model for this quartile. These findings are in line with our previous results and the findings of Coulombe (2021).

Table 3: Root Mean Squared Error for each quantile based on magnitude of the adjusted inflation value to be forecasted

Quantile/Model	**P. MRF**	**Aug.PC**	**RW**	**RF**	**AR(4)**
1	1.76	0.99**	2.85***	0.29***	0.55***
2	0.43	0.42	1.34***	0.18***	0.21***
3	0.06	0.34***	1.26***	0.36***	0.31***
4	2.02	0.93***	4.17***	1.96	1.23***

Note. * indicates $p < 0.1$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$.

5.2 Variable Importance Analysis

Next, we analyze the pooled MRF output through variable importance analysis. Figure 4 plots the relative reduction in RMSE based on a random permutation of each feature. The results indicate that the first lag of exports as a percentage of GDP has the largest effect on the fit of the pooled MRF. This is consistent with the findings of Romer (1993) and Chen (2009), who find that trade-openness significantly reduces oil pass-through and so affects inflation. Accordingly, the first lag of exports and the first lag of the percentage change in exports are both in the top 10 of most important features.

Another variable that is strongly related to oil pass-through is fuel intensity. In line with the results of Chen (2009), we find that fuel intensity is an important feature in modeling inflation. Almost certainly this is attributable to its effect on oil pass-through. Economic theory suggests that if a country consumes more oil per capita, an increase in the oil price will lead to a higher experienced inflation in this country. Similarly, Hooker (2002) related the attenuation of oil pass-through after the 1980s to a decreasing fuel intensity. As fuel intensity is the fifth most important variable, this confirms the significant role for fuel intensity.

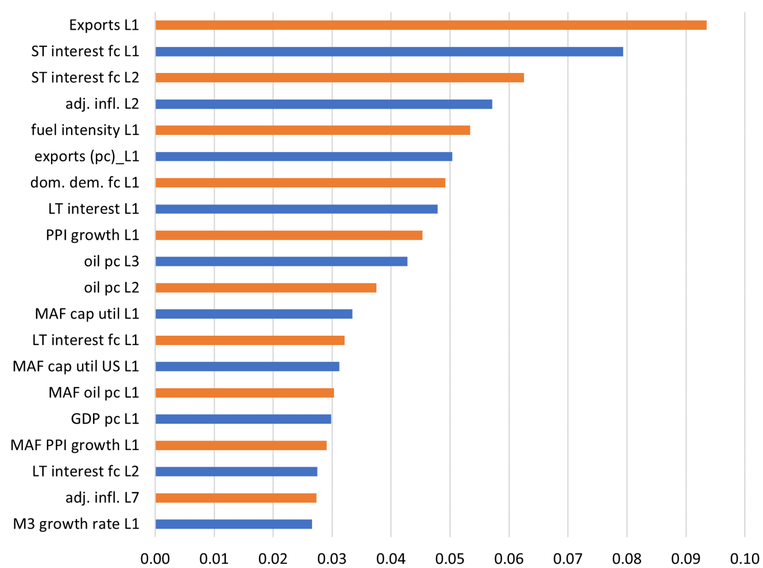

Figure 4: Relative reduction in RMSE based on random permutation of each feature for 2004Q1-2021Q4 adjusted inflation

Lastly, we observe that the second and seventh lag of adjusted inflation are important features in the pooled MRF model. This can be related to the findings of Chen (2009) and Gelos & Ustyugova (2017), who find that a low inflation environment significantly reduces oil pass-through. From an economic perspective, the historic level of inflation is related to how persistent producers will think a price shock to be, which affects oil pass-through (Taylor, 2000).

An additional observation is the the short-term interest rate forecast provided by the OECD appears of significant importance. This is particularly striking as both the first and second lag of this variable score very high in the variable importance analysis. This provides strong indication that the short term interest rate is an important determinant of inflation. This is also in line with economic intuition, think for example of the famous Taylor Rule model.

5.3 Surrogate Trees

We proceed to analyze the short-term oil pass-through and which features are most relevant in this regard by growing surrogate trees on the short term oil pass-through for each country between 2004Q1 up to 2021Q4. For this purpose, we consider the features outlined in Section 2, which past studies have found to be most important in determining oil pass-through.

Figure 5 plots the coefficient θ_1 over time and the surrogate tree fit of this coefficient. In general for most countries we observe that the oil pass-through gradually decreased from 2007Q1 onward up to 2016Q1 approximately, although the path is volatile. Additionally, we observe that for a majority of the countries, the short-term oil pass-through did not change considerably following the Covid-19 crisis, although a small increase can be observed. This is an interesting finding, as it suggests that the currently increasing energy prices are primarily a result of an increasing level in oil prices, rather than an increase in oil pass-through.

Table 4 shows the correlation between the surrogate tree fitted path of θ_1 and its actual path. We observe that overall for most countries, the surrogate trees provide a strong fit to the path of θ_1. This is confirmed by the high correlations between the surrogate tree fit and the actual path of θ_1. Indeed, most correlations are larger than or equal to 0.70. Still, the features are not a perfect fit to the path of θ_1. Most notably for Italy, it can be seen that the surrogate tree misses many of the peaks and troughs after 2016Q1. On the other hand, the surrogate tree for countries such as Austria, Czech Republic, Ireland, Slovenia and Belgium, the peaks and troughs are very well captured.

Table 4: Correlation between path of θ_1 and the fitted path of the surrogate tree

Country	AUT	BEL	GER	FIN	FRA	DNK	GBR	ITA	LUX	NLD	SWE	ESP	CZE	PRT	GRC	HUN	IRL	SVN
Correlation	0.71	0.72	0.64	0.57	0.71	0.64	0.71	0.58	0.65	0.59	0.59	0.76	0.80	0.71	0.64	0.75	0.66	0.71

Next, we analyze the features of main importance in growing the surrogate trees more in depth. Figure 6 shows how many times a tree has selected a certain feature for splitting at least once. We observe that similar to the feature importance analysis from before, the first lag of fuel intensity appears important. This feature is selected by more than half of the surrogate trees. Second, we see that the exchange rate is of significant importance,

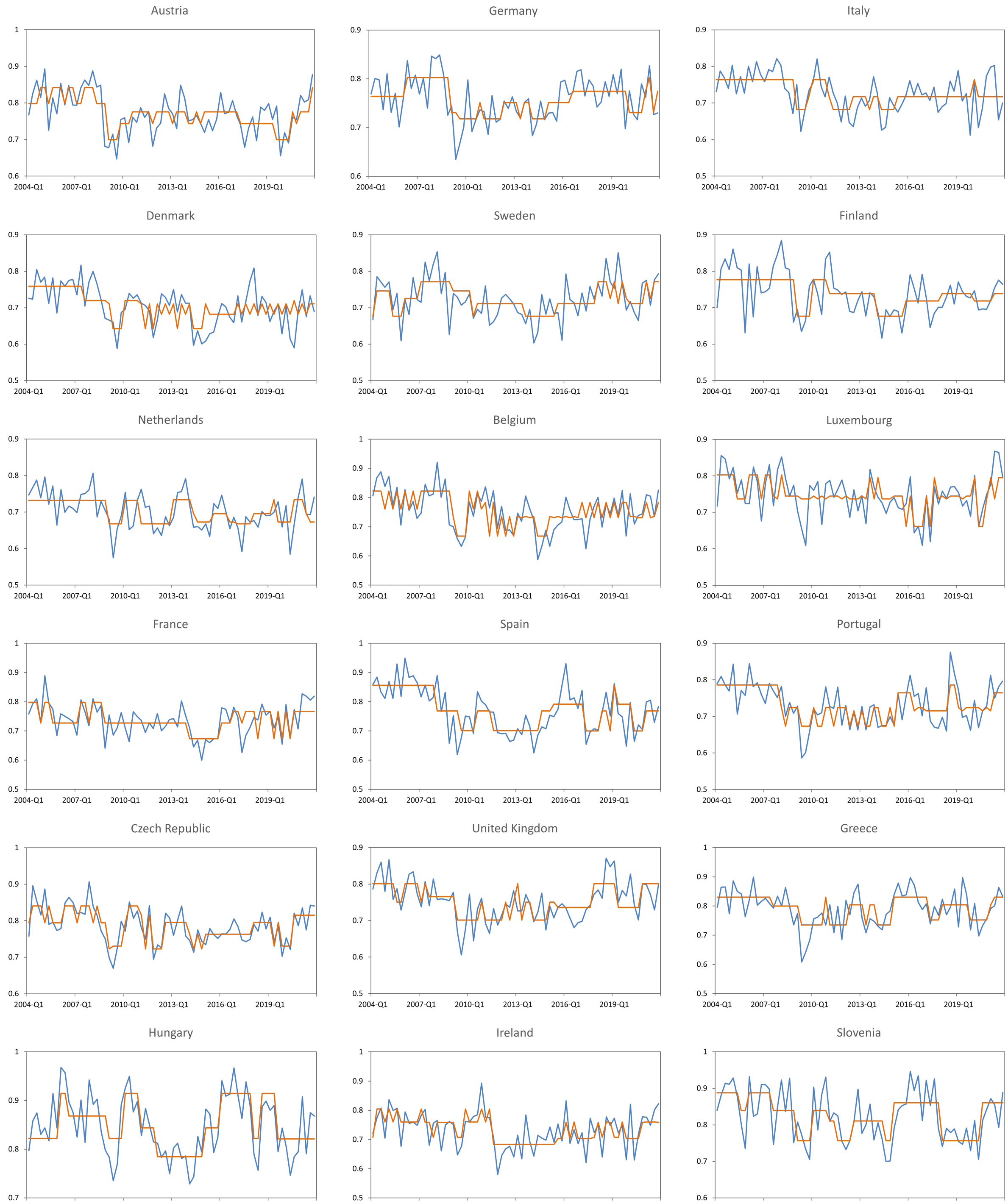

Figure 5: θ_1 coefficient for short-term oil pass-through (blue) and surrogate tree fit (orange) for each country, 2004Q1-2021Q4

as both the first lag and second lag of this feature are in the top 4 of most used features, each being used in almost half of the surrogate trees. This is particularly interesting, because these features are highly correlated. Hence, that both appear so high, confirm the importance of this feature. Third, we observe that the output gap is second highest in the list of features used. This relates to the argument of Gelos & Ustyugova (2017) that oil pass-through is significantly higher when the economy has a positive output gap. More specifically, producers can more easily pass on the increased energy price to consumers in such economic conditions. Lastly, the ratio of debt to GDP also appears to play an important role across the countries in replicating the path of short-term oil pass-through. This can be seen by the first and second lag both appearing relatively high in the ranking. The other features appear to be of importance only to a handful of specific countries.

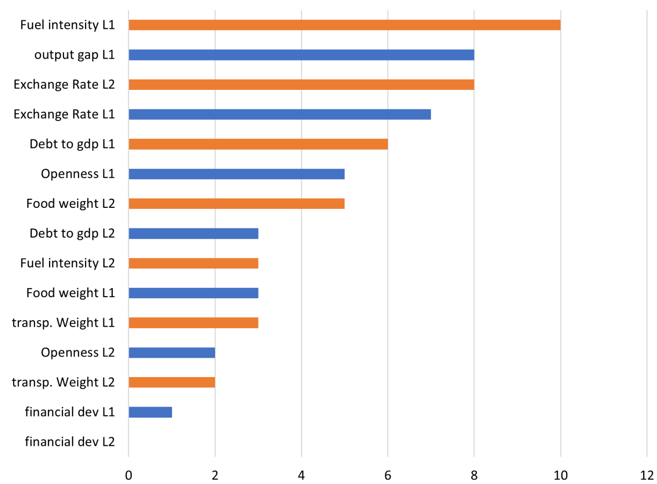

Figure 6: Frequency of selection of each feature across all countries in growing the surrogate trees

For illustrative purposes, we now consider the surrogate trees of Finland (top left), Denmark (top right), Luxembourg (bottom left) and France (bottom right). Each leaf shows the estimated θ_1 coefficient as well as the number of observations in the leaf.

Inspection of the surrogate trees for Finland, Denmark, and Finland confirm the previous arguments that a higher fuel intensity is correlated with a higher short term oil pass-through. More specifically, we see that for a lower fuel intensity, each country has a lower oil pass-through. Moreover, Finland and Denmark use fuel intensity for the first split, which provides further indication for the strength of this feature.

If we look at the surrogate trees of Luxembourg and France, we observe secondly that a higher debt/GDP ratio is associated with a lower oil pass-through. This stands in contrast with the findings of Celasun et al. (2004), who argue that if a country has a stronger fiscal balance, it can intervene better and more credibly to contain an oil-

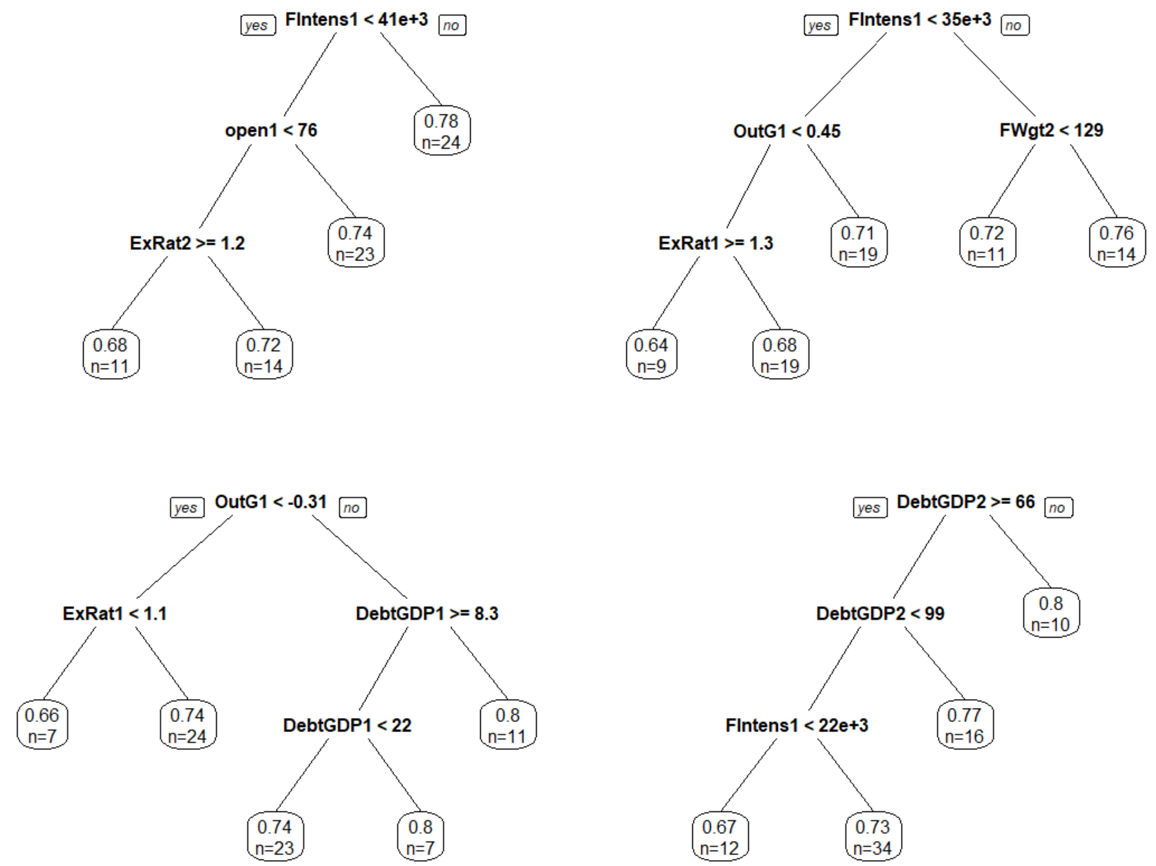

Figure 7: Surrogate Tree of Finland (top left), Denmark (top right), Luxembourg (bottom left) and France (bottom right) for θ_1

shock. The observed pattern in the surrogate trees, however, can likely be attributed to the contemporaneous rapid increase in (1) inflation, (2) sovereign debts relative to GDP and (3) the oil price following the Covid-19 pandemic. The model, therefore, tries to capture the higher inflation by associating a higher debt to GDP ratio with a higher oil pass-through.

For Finland, Denmark and Luxembourg we see that in the left subtrees, the exchange rate is negatively associated with the exchange rate. This is in line with expectations as a higher exchange rate in terms of dollars, means oil can be more cheaply acquired. Therefore, a higher exchange rate is associated with a lower oil pass-through. This is also consistent with the findings of Chen (2009) and Gelos & Ustyugova (2017). Additionally, we observe in the left subtree of Finland that trade openness positively correlates with lower oil pass-through. This is in line with the arguments of Romer (1993) and Chen (2009) that goods can often be imported more cheaply to offset oil shocks. Lastly, we observe in the surrogate trees of Denmark and Luxembourg that a larger (standardized) output gap is associated with a higher oil pass-through. This makes sense, because a positive output gap is associated with very little economic slack, such that an oil shock leads to higher pass-through in consumer prices (Gelos & Ustyugova, 2017). Overall, the surrogate trees illustrate the prominent role that fuel intensity, the exchange rate and the output gap play in the paths of the short-term oil pass-through coefficients.

6 Conclusion

This paper proposes a novel extension to the Macroeconomic Random Forest (MRF) algorithm to model inflation. More specifically, I augment the MRF splitting procedure to accommodate panel data across European countries to exploit commonality in the inflation patterns. We then use this model to assess which macroeconomic state variables are the most important determinants of short-term oil pass-through. Through a forecasting study we document that pooled MRF outperforms all benchmark models, including plain RF, for periods with inflation of moderate magnitude. However, for periods with extreme changes in the inflation rate, pooled MRF cannot outperform the existing autoregressive and Phillips curve models. Through an in depth assessment of the pooled MRF output by means of variable importance analysis and surrogate trees on the path of the short-term oil pass-through coefficient, we find evidence that the fuel intensity, exchange rate, debt to GDP ratio, and output gap play a prominent role in the level of oil pass-through in a country. These results are largely consistent with the findings in the existing literature of Gelos & Ustyugova (2017) as well as Chen (2009). A difference with Gelos & Ustyugova (2017) is that the weights of food in the CPI inflation basket are found to be only of minor importance.

For further research we suggest to consider a larger set of features. A major challenge in constructing the model rested with data availability for all features across all countries for the time period under consideration. This meant that certain variables that are often found to be of significant importance in inflation modelling, such as the number of construction permits and various indicators on the housing market, could not be used. Additionally, further research could investigate applying the pooled MRF framework to a larger set of countries across the world. Our results indicate that pooled MRF works remarkably well even for countries with relatively low correlations in their inflation pattern with other countries. This provides a basis for the inclusion of more countries, which could in turn improve fit for the countries already considered in the model.

References

Alghalith, M. (2010). The interaction between food prices and oil prices.

Atkeson, A., Ohanian, L. E. et al. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve bank of Minneapolis quarterly review*, *25*(1), 2–11.

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

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. CRC Press.

Celasun, O., Gelos, R. G., & Prati, A. (2004). Obstacles to disinflation: What is the role of fiscal expectations? *Economic Policy*, *19*(40), 442–481.

Chen, S.-S. (2009). Oil price pass-through into inflation. *Energy economics*, *31*(1), 126–133.

Coulombe, P. G. (2021). The macroeconomy as a random forest. *Working Paper*.

Coulombe, P. G., Leroux, M., Stevanovic, D., & Surprenant, S. (2021). Macroeconomic data transformations matter. *International Journal of Forecasting*, *37*(4), 1338–1354.

De Gregorio, J. (2012). Commodity prices, monetary policy, and inflation. *IMF Economic Review*, *60*(4), 600–633.

Diebold, F., & Mariano, R. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, ***13***(3), 253–263.

Eliasson, A.-C. (2001). *Is the short-run phillips curve nonlinear? empirical evidence for australia, sweden and the united states* (tech. rep.). Sveriges Riksbank Working Paper Series.

Faust, J. H., J. Wright. (2013). Forecasting inflation. *Handbook of economic forecasting* (pp. 2–56).

Ferri, P., Greenberg, E., & Day, R. H. (2001). The phillips curve, regime switching, and the nairu. *Journal of Economic Behavior & Organization*, *46*(1), 23–37.

Friedberg, R., Tibshirani, J., Athey, S., & Wager, S. (2020). Local linear forests. *Journal of Computational and Graphical Statistics*, *30*(2), 503–517.

Gelos, Y., G. Ustyugova. (2017). Inflation responses to commodity price shocks–how and why do countries differ? *Journal of International Money and Finance*, *72*, 28–47.

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

Herrera, A., & Pesavento, E. (2009). Oil price shocks, systematic monetary policy, and the “great moderation”. *Macroeconomic Dynamics*, *13*(1), 107–137.

Hooker, M. (2002). Are oil shocks inflationary? asymmetric and nonlinear specifications versus changes in regime. *Journal of money, credit and banking*, 540–561.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning* (Vol. 112).

Jerome Friedman, T. H., & Tibshirani, R. (2022). *Lasso and elastic-net regularized generalized linear models* [R package version 4.1-4]. https://cran.r-project.org/web/packages/glmnet/glmnet.pdf

Koester, G., Lis, E., Nickel, C., Osbat, C., & Smets, F. (2021). Understanding low inflation in the euro area from 2013 to 2019: Cyclical and structural drivers. *ECB Occasional Paper*, (280).

LeBlanc, M. D., M. Chinn. (2004). Do high oil prices presage inflation? the evidence from G-5 countries. *UC Santa Cruz Economics Working Paper*, (561), 04–04.

Lòpez-Villavicencio, M., A. Pourroy. (2019). Inflation target and (a)symmetries in the oil price pass-through to inflation. *Energy Economics*, *80*, 860–875.

McCracken, M. W., & Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, *34*(4), 574–589.

Mishkin, F. S. (2009). *Will monetary policy become more of a science?* Routledge.

Molnar, C., Casalicchio, G., & Bischl, B. (2020). Interpretable machine learning–a brief history, state-of-the-art and challenges. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 417–431.

Musso, A., Stracca, L., & van Dijk, D. J. (2007). Instability and nonlinearity in the Euro Area phillips curve.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "why should i trust you?" explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1135–1144.

Romer, D. (1993). Openness and inflation: Theory and evidence. *The quarterly journal of economics*, *108*(4), 869–903.

Stock, J. H., & Watson, M. W. (2008). Phillips curve inflation forecasts.

Taylor, J. B. (2000). Low inflation, pass-through, and the pricing power of firms. *European economic review*, *44*(7), 1389–1408.

Appendix A List of Features and Transformations

Table 1: Features and corresponding transformations

Feature	Transformation	Lags	MAF
Inflation	-	2	-
Inflation	Subtraction of long-run mean	8	2
output gap	Hodrick-Prescott filtered	2	-
output gap	Normalized for each country	2	2
oil price	-	2	-
oil price	log of first difference	4	2
Euro/dollar exchange rate	-	2	2
Unemployment rate	-	2	-
Unemployment rate	Normalized for each country	2	2
Unemployment rate	growth rate compared to last quarter	2	-
Capacity Utilization	-	2	2
Capacity Utilization	growth rate compared to last quarter	2	-
Capacity Utilization US	-	2	2
Long term interest	-	2	-
Long term interest forecast	-	2	-
Short term interest	-	2	-
Short term interest forecast	-	2	-
Unit Labour Cost growth rate	-	2	-
M1 growth rate	-	2	-
M3 growth rate	-	2	-
Domestic demand forecast	-	2	2
PPI growth	-	2	2
Imports percentage change	-	2	-
Exports percentage change	-	2	-
Exports	% of GDP	2	-
Imports	% of GDP	2	-
Trade Openness	% of GDP	2	-
Current Account Balance	% of GDP	2	-
GDP	growth rate compared to last quarter	2	2
GDP forecast	-	2	-
Share Prices	-	2	-
Share Prices	growth rate compared to last quarter	2	-
Housing prices	-	2	-
Housing prices	growth rate compared to last quarter	2	-
Debt/GDP	-	2	-
Fuel Intensity	Per capita	2	-
Transport weight in CPI basket	-	2	2
Food weight in CPI basket	-	2	2
Financial development	index	2	-

Appendix B In-sample predictions pooled MRF

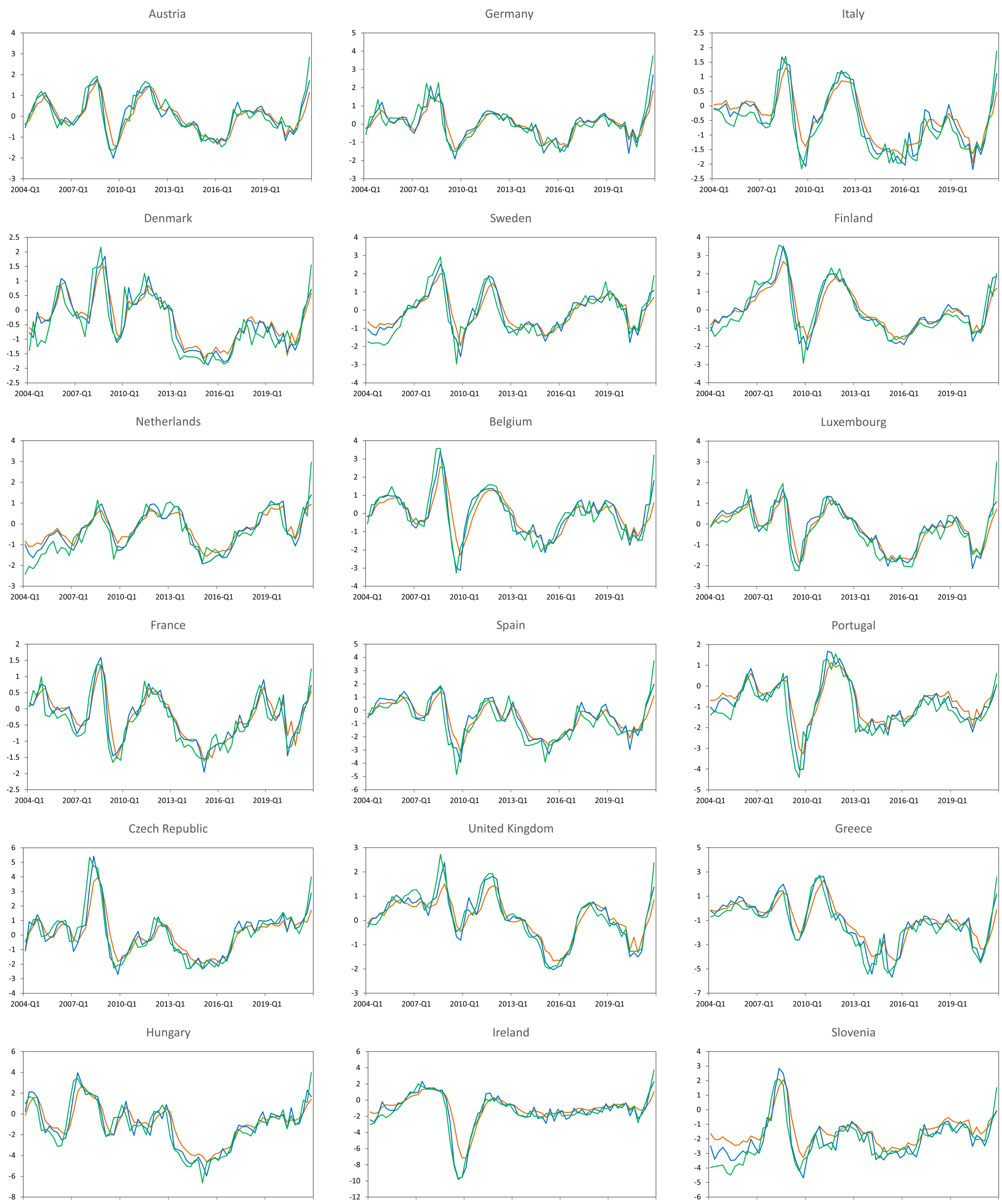

Figure B.1: Inflation (solid) and unemployment rate for 18 European countries (dashed), 2004Q1 - 2021Q4 (Correlations in excess of 0.75 or higher in bold)

Appendix C Code Description

In this Appendix I outline what each of my codes is used for and how it relates to the results obtained in this paper.

- *PMRF and VI*: This code is used firstly to obtain the pooled MRF estimates on the in-sample data and the full sample data. Additionally, this code provides the out-of-bag observations of each tree used and how these can be used to compute the variable importance of each feature.

- *Quarterly Data Preparation*: This code outlines how all the data obtained from the databases are augmented, how MAFs are computed and what variable relates to which economic indicator.

- *DM tests*: This code obtains all the fitted values for each of the 4 models I consider in the forecasting exercise and computes both the RMSEs and the small-sample adjusted DM tests.

- *Surrogate Trees*: This code outlines how all the surrogate trees are grown, and how the plots are obtained.