Forecasting Macroeconomic Variables: A Systematic Comparison of Machine Learning Methods

Teja Konduri[1] and Qian Li[2]

[1]University of Notre Dame

[2]Indiana Legislative Services Agency

July 17, 2024

Abstract

This paper evaluates the performance of an extensive set of machine learning algorithms in forecasting macroeconomic variables relative to baseline econometric models. We conduct a pseudo-out-of-sample forecast for fifteen real, nominal, and financial variables. The findings can be summarized in three points. First, machine learning models perform better than the benchmark model in forecasting real variables but worse than the baseline models in forecasting nominal variables (price indices) and financial variables. Second, machine learning models forecast better than benchmark models during periods of high volatility, like recessions and the COVID-19 pandemic. Third, models that employ dimension reduction frequently appear in the top five most accurate models in forecasting real variables, especially at longer horizons.

Keywords: Macroeconomic Indicators Forecast, Autoregressive, Random Walk, Machine Learning Methods, Data-rich Environment, Ensemble Methods, Dimension Reduction, Diffusion Index

1 Introduction

The evolution of economic prediction methods has shifted from traditional econometric models, like Auto-Regressive (AR) forecasts, to advanced machine learning techniques, marking a significant change in economic forecasting. This shift is driven by the growing complexity and abundance of data, demanding powerful analytical tools capable of capturing nonlinear patterns and efficiently using vast datasets. While traditional econometric models have laid the groundwork, they are often hampered by the "curse of dimensionality," where the increase in the number of predictors leads to an exponential increase in the parameters that need to be estimated. This proliferation of parameters increases the risk of overfitting, where models may fit the training data very closely but fail to generalize to new, unseen data due to their sensitivity to noisy or unrepresentative training samples. This limitation is particularly acute in forecasting economic indicators such as industrial production, employment rates, and inflation, which are key inputs for policy decisions and economic analysis.

Machine learning models offer a promising alternative due to their ability to integrate and learn from vast amounts of data and their flexibility in modeling complex, non-linear relationships often present in macroeconomic environments. Unlike traditional models that require assumptions about the functional form of relationships between variables, machine learning models can adaptively learn these relationships without pre-selection or fixed weighting, potentially allowing for accurate forecasting. For instance, studies like Serrano & Hoesli (2007) and Choudhary & Haider (2012) show that machine learning models can outperform traditional methods like Vector Auto Regressions and autoregressive processes by responding dynamically to shifts in economic conditions.

Despite the growing use of machine learning in economic analysis, there remains a lack of comprehensive studies that systematically compare the performance of these models across a wide array of macroeconomic variables. Our study aims to fill this gap by utilizing a diverse set of machine learning models to forecast key economic indicators and compare their performance with traditional benchmarks. We build upon the work of Kotchoni et al. (2019) by extending their analysis to incorporate a larger universe of machine learning approaches and a more extensive set of economic indicators, providing new insights into their out-of-sample forecasting performance in a sample that also includes the COVID-19 pandemic.

The recent literature has increasingly recognized the potential of machine learning techniques in enhancing economic forecasting, presenting a formidable challenge to traditional econometric methods. For instance, Kotchoni et al. (2019) demonstrate the superior accuracy of regularized data-rich models in forecasting crucial macroeconomic variables, highlighting the necessity of diverse modeling approaches in data-rich environments to improve forecast accuracy. Similarly, Milunovich (2020) reveal how machine learning and deep learning algorithms could outperform traditional benchmarks like the random walk model in predicting real estate indices. Furthermore, Goulet Coulombe et al. (2022) detail how machine learning excels in capturing nonlinear dynamics – a key feature that traditional models often overlook.

Our study confirms these findings by comparing machine learning models against baseline econometric models. It also expands the analysis by exploring their performance across a larger spectrum of economic indicators, including those impacted by the COVID-19 pandemic. This comprehensive approach allows us to dissect the conditions under which machine learning models excel, particularly in handling nonlinear interactions and large datasets. By doing so, we provide policymakers and economic analysts with deeper insights into the efficacy of various forecasting methods during periods of both stability and significant economic turbulence.

We categorize the models considered in our analysis into five distinct groups to provide a structured comparison: traditional benchmark econometric models; individual machine learning models; ensemble machine learning models, which harness the collective power of multiple learners for enhanced accuracy; and two categories through which we reevaluate the machine learning forecasts using dimension reduction techniques.

Our results can be summarized as follows: An analysis of the period leading up to the COVID-19 pandemic (1960M1-2019M12) reveals that ML models generally surpass traditional econometric models in forecasting real variables like industrial production and employment yet fall short when it comes to nominal and financial variables such as CPI and the S&P 500. Expanding our analysis to include a broader set of variables across real, nominal, and financial categories, we observe a consistent pattern: ML models tend to outperform the baseline in forecasting real variables but not nominal and financial ones. This pattern prompts a deeper exploration into the specific conditions under which ML models excel. By benchmarking against the Auto Regressive Diffusion Indices (ARDI) model, we find that the superior performance of ML models in forecasting real variables cannot be solely attributed to their ability to handle data-rich environments. Instead, as Goulet Coulombe et al. (2022) highlighted, the inherent nonlinearity within ML models emerges as a significant factor. Further investigation into periods of high volatility, such as NBER recessions and a sample that includes the onset of the pandemic in early 2020, indicates that ML models improve more over the baseline in forecasting real variables during turbulent periods than normal times. While reducing dimensionality enhances the forecasting accuracy of certain machine learning models, their performance varies across variables and horizons. Finally, the machine learning model, Adaptive Boosting (AdaBoost), and its diffusion index counterpart produce the most stable forecasts over time for real variables, while the benchmark model is the most stable for forecasting nominal and financial variables.

The remainder of the chapter proceeds as follows. Section 2 describes the forecasting framework and targets, the data used in our forecasting exercise, and the empirical evaluation design, while section 3 presents the forecasting models. In section 4, we discuss the forecasting performance of the key variables. In section 5, we expand our analysis to include new variables to examine whether the pattern across different types of variables uncovered in the previous section holds more broadly and present the forecasting results for these new variables. In section 6, we examine the cause behind the strong performance of the machine learning models in forecasting real variables. After determining that non-linearities play an important role in machine learning models' forecasting ability, we examine their performance during NBER recession periods and a sample that includes post-COVID data in section 7. In section 8, we examine the forecasting stability of the models, and in section 9 we conclude our analysis and discuss future steps for research.

2 Forecasting Environment and Evaluation

This section introduces our general forecasting framework and defines the forecasting targets. Then, we describe the data and the forecast evaluation methodology.

2.1 General Forecasting Framework and Forecasting Targets

In this chapter, we consider the following general framework for our predictive models from Mullainathan & Spiess (2017):

$$\arg\min_{\theta} \sum_{t} L(y_{t+h} - f(X_t;\theta)), \;\; t = 1, \ldots, T. \tag{1}$$

where, y_{t+h} is the target, i.e., the variable we predict h periods into the future and X_t is the N - DImensional vector of predictors. We minimize the quadratic loss function, L, by choosing the parameters θ of the function $f\,(.)$. $f(.)$ models the predictors' space linearly or non-linearly using a flexible functional form. In this chapter, the optimal forecast is the conditional expectation $E(y_{t+h}|X_t)$.

We now describe the targets of our forecasting exercise. Let Y_t denote an economic time series we want to predict. Before forecasting Y_t, we stationarize it by following Stock & Watson (2002) and McCracken & Ng (2016). In this chapter, we make direct forecasts. i.e., we forecast a variable's $h-$period ahead value directly from the current period. If y_t is the stationary transformation of Y_t, we forecast its annualized average over the period $[t+1, t+h]$ given by:

$$y^{(h)}_{t+h} = \left(\frac{1200}{h}\right) \sum_{k=1}^{h} y_{t+k} \tag{2}$$

We multiply the monthly growth by 12 to make it annualized and multiply by 100 to calculate the percentage. For more information, see Kotchoni et al. (2019). We deal with three separate types of series:

1. Specifically, if $y_t \equiv \ln Y_t$ is stationary, we forecast equation (2).

2. If $y_t \equiv \ln Y_t - \ln Y_{t-1}$ is stationary, i.e., if $\ln Y_t$ is integrated of order 1 – $I(1)$, then we forecast:

$$y_{t+h}^{(h)} = \left(\frac{1200}{h}\right) \ln\left(\frac{Y_{t+h}}{Y_t}\right) \tag{3}$$

3. If $y_t \equiv (\ln Y_t - \ln Y_{t-1}) - (\ln Y_{t-1} - \ln Y_{t-2}) = \Delta^2 \ln Y_t$ is stationary, i.e., if $\ln Y_t$ is integrated of order 2 – $I(2)$, then we forecast:

$$y_{t+h}^{(h)} = \left(\frac{1200}{h}\right) \left[\ln\left(\frac{Y_{t+h}}{Y_{t+h-1}}\right) - \ln\left(\frac{Y_t}{Y_{t-1}}\right)\right] \tag{4}$$

2.2 Data

In our study, we utilize data from FRED-MD, an expansive monthly macroeconomic database, to evaluate and compare the performance of the forecasting models outlined in the next section. However, the FRED-MD dataset, post-June 2016, does not include the seven Institute for Supply Management (ISM) manufacturing indices. Economists and policymakers recognize these indices, which serve as measures of the Purchasing Managers' Index (PMI), as a crucial indicator of the health of the US economy (see Kauffman (1999)).

Given the significance of these measures, we sourced these series from the YCharts database and incorporated them into the revised FRED-MD panel. The macroeconomic panel we employ in our study comprises 134 monthly macroeconomic and financial time series from January 1960 to December 2023.

2.3 Forecasting Methodology

2.3.1 Pseudo Out-Of-Sample Forecasting Design

For our forecasting exercise, we adopt a pseudo-out-of-sample approach for the period from January 1970 to December 2019. Our forecast horizons span 1, 3, 6, 9, and 12 months, with $593 - h$ evaluation periods for each forecasting horizon h. We use rolling windows for model estimation, with a window size of $120 - h$ months. For example, the forecast for January 1970 is based on data from January 1960 to December 1969;

similarly, the forecast for February 1970 relies on data from February 1960 to January 1970, and so on. This rolling window approach ensures consistency and eases comparison across different models while dynamically adapting to use the latest available data.

To simplify cross-model comparisons, we uniformly use six lags across all evaluation periods and forecasting horizons. Employing six lags, we use information from the recent past, which is useful in forecasting a variable; at the same time, we do not overwhelm the model with too many parameters that come with higher lags. We determine the hyperparameters for the machine learning models, such as number of neighbors, kernel choice, the maximum number of decision splits, and learning rate, based on conventional heuristics.

2.3.2 Variables of Interest

Our analysis focuses on the four macroeconomic indicators forecasted by Kotchoni et al. (2019). These variables, along with their mnemonics, are industrial production (INDPRO), employment (PAYEMS), consumer price index (CPIAUCSL), and the S&P 500 index (S&P 500). Industrial production and employment are real variables, reflecting actual economic outputs and labor market conditions. In contrast, the CPI represents the price levels in the economy and is a nominal variable. The S&P 500 reflects stock market valuations and hence is a financial variable.

We treat the logarithms of industrial production, employment, and S&P 500 as $I(1)$ variables, indicating that their month-on-month growth rates are stationary. Conversely, Stock & Watson (2002), McCracken & Ng (2016), and Kotchoni et al. (2019), classify the logarithm of the CPI as $I(2)$, which implies that the changes in CPI's growth rate–or the inflation rate's growth rate–are stationary.

2.3.3 Forecast Evaluation Metrics

Following standard economic forecasting practices, we evaluate the accuracy of our point forecasts using the Root Mean Square Prediction Error (RMSPE). The RMSPE is calculated as:

$$\text{RMSPE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{y}_t - y_t)^2}$$

where N is the number of forecast evaluation points for each horizon h, $\hat{y}_t$ denotes the forecasted values, and y_t represents the actual observed values at time t. The RMSPE calculates the square root of the average of the squared differences between the forecasted and actual values, providing a measure of the prediction accuracy of a model. The RMSPE penalizes large forecast errors.

Additionally, we employ the Diebold & Mariano (1995) test – the DM test – to statistically compare the predictive accuracy of our models against the baseline econometric model. This test assesses whether the difference in forecasting errors between two models is statistically significant, providing a robust method to ascertain if one model consistently outperforms another across our forecasting horizons. In the subsequent section, we describe the different models we use to forecast the macroeconomic indicators, focusing on their distinguishing features.

3 Model Universe

We employ 24 time series and machine learning models to forecast macroeconomic indicators, organizing them into five distinct categories (see Table 1): baseline models, individual machine learning models, ensemble machine learning models, individual machine learning models using dimension reduction, and ensemble machine learning models using dimension reduction.

3.1 Baseline Models

We use the autoregressive direct (ARD) model as our benchmark (baseline) model following Stock & Watson (2002) and Kotchoni et al. (2019). ARD model is a univariate forecasting method that predicts a variable's $h-$period forecast using its current and lagged values. The model is mathematically expressed as follows:

$$y_{t+h}^{(h)} = \alpha^{(h)} + \sum_{l=1}^{L} \rho_l^{(h)} y_{t-l+1} + e_{t+h}, \quad t = 1, \ldots, T, \tag{5}$$

for $h \geq 1$ and $L \geq 1$. We use this model as our benchmark because of its simplicity. We standardize our analysis by setting $L = 6$ for all models.

Furthermore, we benchmark financial variables against the Random Walk (RW) model without drift, a convention in finance literature.

3.2 Individual Machine Learning Models

Individual machine learning models are the first of our four groups of machine learning models. This category includes models like k-Nearest Neighbors(k-NN), decision trees, and Support Vector Regressions (SVR) with diverse kernels.

K-Nearest Neighbors (kNN): This nonparametric method does not explicitly assume a specific form for the function $f(x_t)$. Instead, the forecasted outcome $y(x_t)$ is derived as the weighted average of the targets of the k nearest data points to x_t. The optimal value of k depends on the bias-variance trade-off, with a common heuristic being $k = \lfloor \sqrt{N} \rfloor$, where N is the size of the training dataset and $\lfloor x \rfloor$ is the greatest integer $\leq x$. In our analysis, with a training dataset spanning ten years or 120 months, we set $k = 10$ to optimize performance. For distance measurement, we employ the Euclidean metric and introduce two weighting schemes for the target forecast: "kNN (uniform)," where all neighbors are equally weighted, and "kNN (inverse)," where weights are inversely related to their distance, emphasizing nearer neighbors more significantly. By assessing the "neighborhood" of a given data point, kNN captures these spatial dependencies, which are not explicitly modeled in traditional parametric approaches. This enables the algorithm to adaptively respond to the data's intrinsic structure, making it especially effective when economic variables show significant spatial continuity or clustering.

Decision tree regression: This model utilizes a tree-structured approach to forecast future data, adept at identifying nonlinear relationships and interactions between variables without predetermined functional forms. Originating from Quinlan (1986)'s ID3 algorithm, decision trees use a top-down, greedy search to construct decision rules directly from data, optimizing for continuous outputs with strategies like Standard Deviation Reduction. The choice of hyperparameters – limiting to 20 leaf nodes and requiring a minimum of 3 samples per leaf – strikes a balance between simplicity and depth, ensuring interpretability, computational efficiency, and generalizability. Using hyperparameter tuning methods such as grid search, randomized search, or Bayesian optimization facilitates discovering an optimal model configuration that is accurate and resistant to overfitting.

Support Vector Regression (SVR): SVR is a machine learning technique designed specifically for regression analysis. Unlike traditional linear regression, which aims to fit a line through data points, SVR aims to find a function $f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) + b$ that approximates the relationship between input features x and target values y within a certain margin of error ϵ. This function $f(x)$ is created from the training data using kernel functions $K(x, x')$, which enable the model to capture nonlinear relationships by mapping inputs into a higher - DImensional space. We use the following four kernels:

- Linear Kernel: $K(x, x') = \langle x, x' \rangle$
- Polynomial Kernel: $K(x, x') = (\gamma\langle x, x' \rangle + r)^d$ where γ, r, and d are kernel parameters that control the shape of the polynomial.
- Radial Basis Function (RBF) Kernel: $K(x, x') = \exp(-\gamma\|x - x'\|^2)$, , with γ influencing the spread of the RBF kernel.
- Sigmoid Kernel: $K(x, x') = \tanh(\gamma\langle x, x' \rangle + r)$

The selection of the kernel function $K(xx')$ and the tuning parameters, like γ, d, and r, play a role in determining how well the model fits the data. SVR optimizes a set of coefficients α_i and α_i^* along with a bias term b to ensure the model is as flat as possible while fitting within an epsilon tube around the training data. This approach allows SVR to effectively handle high - DImensional datasets where traditional regression models face challenges.

However, SVR has limitations in terms of efficiency. Training SVR models with nonlinear kernels can be computationally intensive. The performance of SVR algorithms is sensitive to the scale of input features. Thus, pre-processing steps such as scaling become essential. Standardizing predictors and outcomes helps prevent any variable from influencing the model due to its magnitude. After estimation, we reverse the scaling process to facilitate model comparison using the Root Mean Squared Prediction Error (RMSPE) metric.

3.3 Ensemble Machine Learning Models

Ensemble methods enhance prediction accuracy and robustness by combining the outputs of multiple base estimators. This approach leverages the strength of various algorithms to offset a single model's weakness. The most popular ensemble methods for regression include Random Forest, Extreme Gradient Boosting (XGBoost), Adaptive Boosting (AdaBoost), and Gradient Boosting, all of which have distinct advantages and implications for the data being forecasted (Dietterich (2000)).

Random Forest: Random Forest works by creating decision trees during training and reporting the average prediction of these trees (Breiman (2001)). This approach is highly effective for datasets with complex interactions and nonlinear relationships as it doesn't rely on the underlying distribution of the data. The key strength of Random Forest lies in its ability to combat overfitting through bagging, a method that combines the results of multiple models to enhance performance (see Breiman (1996)). If Random Forest proves to be the most accurate model, it indicates that the dataset benefits from a model that can handle high - DImensional data and complex interactions between variables.

Gradient Boost: Gradient Boosting builds a model iteratively by minimizing loss through gradient descent (Friedman (2001)). This technique offers a way to develop models that evolve gradually over time by focusing on correcting errors from previous iterations. Gradient Boosting is particularly effective for datasets exhibiting complex interaction patterns where gradual improvements are essential.

Extreme Gradient Boosting (XGBoost): XGBoost is an advanced version of gradient boosting known for its efficiency, flexibility, and portability (Chen & Guestrin

(2016)). It enhances gradient boosting by improving speed and performance while effectively handling sparse data. This approach is particularly suitable for situations where both speed and accuracy are crucial, excelling in cases where precision is a top priority. The triumph of XGBoost in a horse race underscores the dataset's receptiveness to a model that emphasizes enhancements based on errors, underscoring its sensitivity to fine-tuning and regularization to prevent overfitting.

AdaBoost (Adaptive Boosting): AdaBoost is a boosting technique aimed at transforming many weak learners into one strong learner (Freund & Schapire (1997)). It adjusts the weights of misclassified instances to ensure subsequent classifiers pay more attention to them. This method works well for imbalanced datasets or those requiring resilience against noise and outliers. If AdaBoost surpasses all other models in performance, it suggests that the data benefits from iterative instance reweighting, indicating varying degrees of complexity within the data and necessitating adaptive adjustments.

3.4 Machine Learning Models Using Dimension Reduction

Our dataset comprises over 100 macroeconomic variables. Due to the intricate interdependencies between the predictors and the forecasted variables, there is a high risk of overfitting. We integrate diffusion indices (DIs) derived from principal component analysis (PCA) into our forecasting models to address these challenges.

Ma & Zhu (2013) and Kotchoni et al. (2019) identify three key methods to improve out-of-sample forecasting accuracy while mitigating overfitting: sparse modeling, regularization, and dense modeling. We adopt dense modeling via PCA, which assumes that a few principal components can significantly capture the variance in the data. These components, our DIs, condense the dataset's vast information into a manageable form, enhancing model efficiency [1]. DIs retain the information that has the most predictive power and discard the noise and less informative variability that contributes to overfitting.

In categories four and five of table 1, we re-evaluate the forecasts of all the machine learning models using DIs. We select the number of factors for each variable and hori-

[1] Stock & Watson (2002)

zon based on the panel criteria proposed by Bai & Ng (2002). From the recommended number of factors, we pick the smallest number of factors for parsimony. We expect the revised models, identified with a "DI" suffix, to deliver improved forecasting accuracy by efficiently leveraging the condensed yet informative representation of the data.

4 Forecast Results for Key Variables

In this section, we present the results for the forecasting accuracy for industrial production, employment growth, CPI inflation, and S&P 500 index returns, presented in Tables 2 to 5. Our analysis spans the entire out-of-sample period from January 1960 to December 2019. Each table's left panel displays the full out-of-sample forecasting results, while the right panel focuses on performance during NBER recessions (i.e., target observation belongs to a recession episode) which we discuss in section 7.1. The baseline model's RMSPE occupies the first row of each panel, with subsequent rows comparing the relative RMSPE of machine learning models to this baseline. The relative RMSPE of a model is the ratio of its RMSPE to the RMSPE of the baseline. We underline the best model in terms of relative RMSPE (i.e., the minimum relative RMSPE) for each horizon, and the significance levels for the DM tests are displayed using the conventional notation with three, two, and one asterisks.

The forecasting performance for industrial production growth reveals that multiple machine learning models outperform the baseline ARD model at each horizon. Of these models, SVR (RBF) and AdaBoost stand out for their accuracy, aligning with the forecast accuracy of Kotchoni et al. (2019)'s best models. Seven models beat the baseline in the short term at $h = 1$, with AdaBoost demonstrating superior short-term forecasting abilities than the other models. While for longer horizons at $h = 9, 12$, we still find 8, 6 models beating the baseline. However, the DI-enhanced models, notably kNN variants, perform best suggesting a spatial proximity among industrial production data points and their lags.

For employment growth, as shown in Table 3, eight models outperform the baseline in the short term, while only four models outperform the baseline in the long term. Random Forests and AdaBoost initially outperform the baseline, incorporating diffusion indices generally enhancing model performance across various horizons. Medium-term forecasts see SVR (linear) - DI and AdaBoost - DI as front runners. At more extended horizons, kNN (uniform) - DI shows improved RMSPE over the baseline. However, these improvements are not always statistically significant, indicating that while forecast accuracy improves, it does not uniformly exceed baseline performance throughout the period.

As detailed in Table 4, all machine learning models fall short of surpassing the baseline ARD model in forecasting CPI inflation accuracy, underscoring the challenges machine learning models face with nominal variables.

The forecasting results of S&P 500 returns are outlined in Table 5. While, in principle, we need real-time data vintages for forecasting financial variables, our study relies on the latest available information at the time of forecasting. Like CPI inflation, machine learning models do not outperform the baseline random walk model, with only kNN (uniform) and AdaBoost yielding better predictions than the baseline at select horizons. The RW baseline ranks first at the longer horizons, while it comes fourth, second, and third at $h = 1, 3, 6$ respectively. Notably, there is no statistically significant difference between the baseline and machine learning forecasts.

Overall, our initial forecast results show the superiority of machine learning models in predicting real variables like industrial production and employment. However, machine learning models appear less promising for forecasting nominal and financial variables such as CPI inflation and S&P 500 returns. While PCA, as a dimension reduction technique, aids in forecasting employment, it does not significantly enhance the prediction of other variables.

5 Do Forecasting Patterns Hold Across a Larger Set of Variables?

5.1 Expanding the Variable Set

In the previous section, we saw that machine learning models are more accurate than univariate time series models in forecasting Industrial Production (INDPRO) and Employment (PAYEMS), which are real variables, i.e., quantities. On the other hand, the baseline econometric models are more accurate than the ML models in forecasting CPI, a price index, and S&P 500, a financial index. To investigate whether there is a pattern that machine learning models can outperform baseline models in predicting real variables and under-perform baseline models in predicting nominal and financial variables, we expand our set of variables to include five variables in each category to investigate if the pattern holds.

In addition to industrial production and employment, our expanded set of real variables includes real personal income (RPI), the unemployment rate (UNRATE), and real personal consumption expenditure (Real PCE). The industrial sector, together with construction, accounts for the bulk of the variation in national output over the course of the business cycle. On the other hand, the three new variables reflect consumer sentiment in the economy. Since consumption contributes to between 60-70% of the GDP, the five real variables we forecast are strong indicators of the economy's health.

We focus on various consumer and producer price indices when expanding our selection of nominal variables. While CPI for all items offers a broad measure of inflation, we also incorporate a less volatile measure of the CPI by excluding volatile food prices. Additionally, we include the Personal Consumption Expenditures Price Index (PCEPI) to complement these measures. Unlike the CPI, which directly assesses consumer out-of-pocket expenses, the PCEPI offers a wider lens on inflation by capturing all goods and services consumed by households, including those paid on their behalf, like healthcare benefits. Finally, the two producer price indices (PPIs) – the PPI for finished consumer goods, and the PPI for crude metals – offer insights into sector-specific inflation pressures that broader indices might not capture. We focus on forecasting the second difference (I(2)) of the logarithm of each price index.

Lastly, our expanded selection of financial variables spans an array of indicators such as treasury rates (1 and 10 years) and exchange rates (US/UK and CA/US foreign exchange rates). The treasury rates provide a spectrum of risk and time preferences in the financial markets, while the foreign exchange rates are critical for assessing international trade dynamics and financial flows.

5.2 Forecast Results

This section presents the forecasting results for the additional variables in tables 6 – 8. For each variable, we show the baseline RMSPE as well as the relative RMSPEs of the five best models for that variable. We present detailed results in the appendix.

Table 6 displays the forecast results for the newly added real variables: real personal income, unemployment rate, and real personal consumption expenditure. We find that the pattern of machine learning models outperforming the baseline continues to hold for all three real variables.

Our analysis of the income forecasts reveals that 12 models beat the baseline at $h = 1$, while seven models are more accurate than the baseline at $h = 12$. In particular, the SVR (RBF) and AdaBoost models excel in the short to medium term (1, 3, and 6 months). For longer forecast horizons (9 and 12 months), kNN methods incorporating diffusion indices emerge as the most accurate.

We next examine the forecasting accuracy of the unemployment rate. At $h = 1$, eight machine learning models including Random Forests, AdaBoost, and the kNN models using diffusion indices consistently outperform the baseline with a relative RMSPE between 0.93 and 0.96. At longer horizons, only the two kNN - DI models consistently outshine the baseline, with their accuracy improving with the horizon from a relative RMSPE decreasing from 0.87 at $h = 3$ to 0.75 at $h = 12$. 12 models perform better than the baseline at $h = 12$, but only two predict significantly differently than the benchmark.

For Real PCE, only four ML models outperform the baseline at $h = 1$, and three perform better at $h = 12$. In the short run (1 and 3 months), the SVR (RBF) model shows remarkable accuracy, whereas, for medium-term forecasts (6 and 9 months), the kNN (uniform) with diffusion indices takes the lead.

The data underscores a consistent trend: machine learning models, particularly those incorporating advanced techniques like diffusion indices, significantly outperform the baseline ARD model across the board for real variables. This finding aligns with our initial hypothesis, affirming the superior predictive power of machine learning models in this context.

Table 7 illustrates the forecasting accuracies of the different ML models for the nominal variables – CPI, CPI less food, PCEPI, PPI for finished consumer goods, and PPI for crude metals. Notably, the baseline model consistently outperforms all machine learning models across all forecast horizons, underscoring the ARD model's robustness in predicting these variables.

Finally, we present the results for the financial variables in table 8. Our analysis reveals that the RW baseline consistently outperforms all machine learning models in full out-of-sample forecasts. For 1-year Treasury rate (GS1), 3 and 2 models forecast better than the RW baseline at $h = 1, 3$ respectively but their performance is not significantly different from the baseline. For both S&P 500 and GS1, AdaBoost shows marginal improvement over the random walk baseline at $h = 1$.

These findings suggest that while the RW model remains a strong predictor for financial variables overall, the AdaBoost model shows promise in specific contexts and horizons.

This section confirmed our hypothesis that machine learning models are more accurate than the baseline AR Direct model in forecasting real variables. On the other hand, the baseline is the best in forecasting nominal and financial variables.

6 Data-Richness vs Non-Linearities

In the previous section, we saw that the machine learning models outperform the baseline when predicting real variables. Now, we will investigate what leads to the better performance of the ML models over the baseline model. Recall that our baseline model, the Auto Regressive Direct forecast, is a linear univariate model. i.e., the baseline uses only the lags of the predicted variable as the predictors, and the predicted variable is a linear combination of the predictors. Our ML models, on the other hand, are in a data-rich space where they use the 134 variables in the FRED-MD database along with their lags. At the same time, all our ML models take advantage of non-linear relationships between the predictors and the predicted variable to improve forecast accuracy. In this section, we try to find why our ML models are good at forecasting the real variables by examining two dimensions: data-richness and non-linearities.

6.1 A New Baseline

To accomplish our goal, we use a new baseline model, the Auto Regressive Diffusion Indices (ARDI), which was first introduced by Stock & Watson (2002). In this model, the diffusion indices are extracted from a set of predictors and then augmented in a direct autoregression. This model is written as:

$$y_{t+h}^{(h)} = \alpha^{(h)} + \sum_{l=1}^{p_y^h} \rho_l^{(h)} y_{t-l+1} + \sum_{l=1}^{p_f^h} \beta_l^{(h)} F_{t-l+1} + e_{t+h}, \quad t = 1, \ldots, T$$

where F_t are $K^{(h)}$ consecutive static factors and the superscript h stands for the value of K when forecasting h periods ahead. We use various information criteria based on penalized likelihood, such as AIC and BIC, to determine the number of factors to be included in the predictive regression for each target variable and forecasting horizon. We use the smallest value these criteria return as the optimal value of $K^{(h)}$. To maintain uniformity in the lag selection across all our models, we use six lags in the ARDI model by setting $p_y^h = p_f^h = 6$. With this, the $h-$step ahead forecast is obtained as:

$$\hat{y}^{(h)}_{T+h|T} = \hat{\alpha}^{(h)} + \sum_{l=1}^{p_y^h} \hat{\rho}^{(h)}_l y_{t-l+1} + \sum_{l=1}^{p_f^h} \hat{\beta}^{(h)}_l F_{t-l+1}$$

6.2 Results

Table 9 displays the results of the top machine learning models compared with ARDI. Even though the ARDI is a data-rich model, our machine learning models continue to beat it for all five of our real variables. For industrial production, we see that AdaBoost's RMSPE is only 0.87 times the RMSPE of ARDI for $h = 1$. For longer horizons, the kNN models have relative RMSPEs less than 1, ranging between 0.83 and 0.72. We also observe that more ML models beat the ARDI baseline than the ARD. At $h = 1$, 12 models outperform the ARDI while only 7 outperformed the ARD. Similarly, at longer horizons of $h = 9, 12$, we find that 13 and 20 models perform better than the ARDI while only 8 and 6 outperformed the ARD.

For employment, 13 models, including Random Forests, XGBoost, AdaBoost, and their DI counterparts, outperform the ARDI baseline at $h = 1$, with random forests having the smallest relative RMSPE of 0.83 compared to ARDI. As the horizons get larger, the number of models outperforming the baseline remains at 13, although the best models change to the SVR (linear) - DI and kNN (uniform) - DI.

A similar pattern repeats for real personal income, unemployment rate, and real PCE. 20 ML models outperform real personal income's baseline forecast at $h = 1$. AdaBoost is the most accurate forecaster for income and unemployment rate, while for real PCE it is the SVR (rbf). As the horizons increase, kNN uniform DI becomes the best predictor for all three variables. This indicates that the kNN uniform DI excels at capturing spatial patterns or dependencies in the data, maintaining temporal stability, handling data complexity, incorporating relevant features, and exhibiting flexibility in modeling the data dynamics over time. At $h = 12$, 21, 13, and 22 models outperform the ARDI baseline for the three variables respectively.

Overall, in this section, we observe four things. First, at least one machine learning (ML) model surpasses the ARDI baseline for all variables at each forecasting horizon.

This suggests that, despite the richness of the data, ML models are adept at leveraging either spatial dependencies or non-linearities in forecasting real variables. Second, at $h = 1$, multiple ML models can predict the value of these variables better than the ARDI. This indicates that the ARDI model might not adequately capture the short-term dynamics or may be too rigid in its assumptions about the data. Other models, including kNN uniform DI, might be more flexible or better suited to capture the short-term fluctuations, potentially by incorporating spatial dependencies or leveraging non-linear relationships in the data. Third, the consistent dominance of kNN models over longer horizons underscores the increasing importance of spatial patterns or dependencies in forecasting. Finally, more ML models beat the ARDI than they did ARD. This indicates that the data-rich linear model, ARDI, has less accuracy than the data-poor linear model, the ARD.

Our analysis echos the finding of Goulet Coulombe et al. (2022) that nonlinearity is of vital importance in forecasting macroeconomic indicators.

7 ML Forecasting in Highly Volatile Environments

In the preceding section, we highlighted the proficiency of machine learning (ML) models in exploiting non-linearities in the data. This section advances that discussion by critically evaluating the performance of these models during significant economic instabilities, notably 1) NBER recessions and 2) the extreme fluctuations during the COVID-19 pandemic. In this section, we aim to shed light on the predictive strength and resilience of ML models across real, nominal, and financial variables during such downturns.

7.1 Performance During NBER Recessions

In this subsection, we analyze the forecasting accuracy and robustness of ML models during NBER recession periods. Focusing on these recession episodes allows us to equip policymakers, financial analysts, and forecasters with actionable insights into which models have consistently outperformed the baseline under economic stress. This analysis is particularly vital when the likelihood of a recession is high, enabling us to develop preemptive strategies for economic forecasting.

We present the results for the real variables in the right-hand side panel of table 6. In the short run, AdaBoost and its DI counterpart emerge as the top performers in predicting these five variables, a phenomenon that can be attributed to AdaBoost's strategy of assigning higher weight to trees with larger errors and lower weight to those with smaller errors. This adaptability makes AdaBoost particularly effective during periods of heightened short-term fluctuations. As we extend our analysis to longer horizons, we observe that Support Vector Regression (SVR) models, especially those utilizing linear and sigmoid kernels, surpass the baseline predictions. For employment growth and unemployment rate, the efficacy of the SVR models, particularly with the linear kernel, in longer-term forecasts demonstrates their strength in capturing the underlying linear trends within the economic indicators. Whereas, for real personal income and real personal consumption expenditure, the success of models with sigmoid kernels hints at their potential to handle non-linear patterns. One notable observation across all five variables is that the performance of most models relative to the baseline improves during recessions compared to the full results. We can also see that the best models have lower relative RMSPE during recessions across all horizons than the full sample. For example, for industrial production, at $h = 12$, the best model in the full POOS is kNN (inverse) - DI with a relative RMSPE of 0.85. The relative RMSPE of this model for the same horizon during recessions is 0.80, and the best model at $h = 12$ during recessions is SVR with a linear kernel with a relative RMSPE of 0.73. We also notice that more models outperform the baseline during the recessions than the full POOS for all five variables. For example, while only six models outperformed the ARD baseline for industrial production at $h = 12$ in the full POOS, seventeen models outperform the baseline at $h = 12$ during the recession periods. Similarly, for employment the corresponding number of models beating the baseline are 4 and 18. For unemployment, the models

For the nominal variables detailed in table 7, we observe that machine learning models, particularly AdaBoost and its DI counterpart, outperform the baseline models in the short run. The machine learning models showed improved performance at $h = 1$ during recessions for all nominal variables compared to the full sample. However, the improvement in prediction accuracy over traditional methods is not statistically significant. As we look towards longer forecasting horizons, the baseline Autoregressive Direct (ARD) model emerges as the most accurate predictor. The ARD model's accuracy could be attributed to the nature of inflation and its expectations, which are typically well anticipated by markets and individuals alike, leading to its variations acting more like exogenous noise in the economic system. Consequently, models that rely on a wealth of data points tend to be overparameterized, resulting in a diminished predictive performance for these nominal series.[2]. This pattern suggests that for nominal variables like inflation, simpler models that can effectively capture long-term trends without overfitting to short-term fluctuations may provide more reliable forecasts.

Finally, the results for the financial variables are detailed in the right-hand side panel of table 8. Our analysis reveals that multiple models outperform the baseline, with the degree of improvement varying across different forecasting horizons. Specifically, for GS1, the SVR model with an RBF kernel shows a notable improvement over the baseline, enhancing predictions by 0.41 and 0.38 percentage points (pp) for horizons of 9 and 12 months, respectively. Similarly, for the US/UK Foreign Exchange rate, the SVR model employing a sigmoid kernel surpasses the baseline's accuracy at a 12-month horizon. For the other three variables–GS10, S&P 500, and CA/US Exchange Rate–although we observe a relative RMSPE of less than 1 for various models at different horizons, the increase in accuracy is not statistically significant. Therefore, while machine learning models can offer marginal improvements in forecasting accuracy for financial variables, these gains are minimal.

The machine learning models continue to show improved performance during recessions compared to the full out-of-sample forecasts for the financial variables. Notably, the relative RMSPE of the best model for the full out-of-sample is less than 1 only for 4 of the 25 variable-horizon combinations. On the other hand, for the recession periods, the best

[2] See Kotchoni et al. (2019)

model's relative RMSPE is less than 1 for 20 of the 25 variable-horizon combinations.

In conclusion, our analysis reveals the differentiated performance of machine learning models across real, nominal, and financial variables during recession periods, highlighting their potential in accurate forecasting under economic stress conditions. The adaptability of models like AdaBoost and SVR is particularly noteworthy, suggesting their utility in addressing the complex dynamics of economic downturns.

During recessions, the machine learning models improve upon both the baseline ARD and their relative performance to baseline for the full POOS period for the real variables. On the other hand, for the other two categories, machine learning models continue to show smaller RMSPE relative to the baseline compared to the full POOS period. However, they improve over the baseline only for a few variables at select horizons during recession periods.

As we transition from the context of NBER recessions to the unprecedented challenges posed by the COVID-19 pandemic, it is crucial to recognize the initial severe disruptions it caused across key economic indicators. Specifically, the five real variables in our study – industrial production, employment, real personal income, unemployment rate, and real personal consumption expenditure – experienced significant spikes at the pandemic's onset which are larger in magnitude than the disruptions during the Great Recession of 2008.[3] Given our findings that machine learning models are adept at capturing extreme fluctuations, this sets a compelling premise for extending our analysis to include the data during COVID and post-COVID up to the end of 2023. By incorporating data from this period, we aim to validate the resilience and forecasting accuracy of these models in the face of such a global crisis. The next section presents the results for the full pseudo-out-of-sample forecasts up to 2023m12. These results underline the critical role of advanced forecasting techniques in navigating through and beyond the economic ramifications of the COVID-19 pandemic.

[3] See Appendix figure 8 for more details.

7.2 Post-Covid Data Analysis

In figure 1, we compare the relative RMSPE of the best models for the pre-pandemic sample up to 2019m12 and the full sample up to 2023m12 for all the real variables. This figure shows that the machine-learning models outperform the baseline more in the full sample than during the pre-pandemic sample.

Further, tables 10 to 14 detail our pseudo-out-of-sample forecast results spanning from January 1960 to December 2023, encompassing the entirety of the COVID-19 pandemic and the subsequent two years. Several critical insights emerge from this analysis:

First, the baseline ARD model's forecasting accuracy declined for all variables across every horizon compared to pre-pandemic performances, with the unemployment rate's 1-month ahead prediction accuracy dropping by 16.82 pp. This decline in accuracy reduced the overall statistical significance of the models' forecasts relative to the baseline during the examined period.

Second, despite these challenges, our analysis demonstrates the resilience of machine learning models, as at least one ML model outperforms the baseline for each variable and forecasting horizon. Notably, the 1-month ahead forecast for real Personal Consumption Expenditures (PCE) stands out, with almost all ML models–excluding Random Forest - DI, Decision Trees, and Decision Tree - DI–surpassing the baseline's accuracy.

These findings highlight a generalized decrease in forecast accuracy across all models, including the baseline, when factoring in the period during and after the COVID-19 pandemic. However, it is evident that certain ML models still managed to exceed baseline performance in accuracy, albeit not consistently across the entire sample. This underscores the potential of machine learning approaches in adapting to and forecasting under the unprecedented economic conditions introduced by the pandemic, suggesting avenues for future research to refine these models for enhanced predictive performance in similarly volatile contexts.

8 Stability of Forecast Results

This section evaluates the stability of forecast performance for real, nominal, and financial variables over time, utilizing a 36-month rolling average of the RMSPE similar to Kotchoni et al. (2019). Our analysis not only assesses model adaptability under changing economic conditions but also highlights the impact of major economic events on forecast accuracy.

Figures 2 – 4 display the 3-year moving average of the RMSPE of select models and the baseline for real, nominal, and financial variables, respectively, at a forecast horizon of $h = 3$ months. By applying a 36-month rolling average to the RMSPE, we can observe how the accuracy of forecasts evolves over time, shedding light on the models' adaptability to changing economic conditions. The models we chose are AdaBoost, AdaBoost - DI, kNN(uniform), kNN (uniform) - DI, and SVR (linear) - DI. We selected these models based on their consistent performance, emerging as the top-performing models across all fifteen variables and various forecasting horizons. For each category, we show the forecasting performance of the original sample period up to 2019 in the left-hand panel. On the right, we also show the extended sample that includes the COVID-19 pandemic.

For industrial production, which is in the first row of figure 2, SVR (linear) - DI is the best performer post-1980s. In the case of real personal income, We find that kNN (uniform) - DI and AdaBoost keep exchanging places for the lowest RMSPE up to 2019. However, post-COVID, we also find that the baseline and AdaBoost - DI models outperform the others. The economic stability of the Great Moderation period from the mid-1980s to 2007 contributed to a decline in the RMSPE of the real activity series, especially employment, real personal income, unemployment rate, and real personal consumption expenditure. Kotchoni et al. (2019) also find that the RMSPE lowered during the Great Moderation period. The relative RMSPE also systematically decreased during and after recessions. Additionally, the increase in the relative RMSPE was larger around the oil price shocks (1973-1974), Great Inflation (1965-1982), Great Recession (2008-2009), and COVID-19 pandemic (2020) than the increase in the relative RMSPE around the 1991 and 2001 recessions. These volatility changes for industrial production and employment align with macroeconomic uncertainty dynamics in Jurado et al. (2015). Also, from the post-COVID plots, we can see that the forecasting performance of all models decreased significantly, with all five variables showing an increased RMSPE. At the end of 2023, the forecast accuracy started improving for all variables except real personal income. This could be because of a second set of fluctuations in income in 2021 due to the checks after an initial disruption in 2020. However, the RMSPE of all models is still above pre-covid levels.

Coming to the nominal variables in figure 3, we observe that the baseline ARD model consistently has the lowest RMSPE of all models throughout the sample period up to 2023.

A slow downward trend in the RMSPE that began in the early 1980s vanished at the beginning of the 1990s, coinciding with the inflation-targeting regime. As suggested by Clark & Davig (2009), Jørgensen & Lansing (2019), etc., improved anchoring of inflation and expectations results in overall lowered volatility, which could have led to better forecasting and lower RMSPE during this period. This downward trend also echoes the 36-month rolling RMSPE plot in Kotchoni et al. (2019). In the early 2000s, the US economy saw a period of sustained high inflation due to increased economic activity worldwide. During this period, the RMSPE of all the nominal variables increased and peaked at the Great Recession of 2007-08.

Finally, two peaks emerge when we look at the finance variables in figure 4. For S&P 500, GS1, GS10, and the CA/US foreign exchange rate, the first spike in RMSPE happens during the 1980-82 recessions following the Iranian revolution and the tightening of monetary policy. The second peak happens for GS10 and the US/UK forex rate during the Great Recession. For all financial variables, the ARD model performs the best. However, during the recession periods, the kNN (uniform) - DI model also had the lowest RMSPE.

Now, we focus on the cumulative forecast errors in figures 5–7. From the forecast errors, we learn about the overall bias of each model. i.e., are the models consistently predicting above or below the observed value?

For all real variables except the change in unemployment rate in figure 5, we see that the benchmark ARD model has the least bias. We find that AdaBoost and AdaBoost - DI produce the least biased forecasts for the unemployment rate.

While we know that the baseline ARD model produces the most accurate predictions for all nominal variables at the aggregate level, it is not always the least biased. In fact, from figure 6, we can see that multiple ML models, such as SVR (linear) - DI, AdaBoost, and AdaBoost - DI, have lower bias than the baseline for CPI, CPI less food, and PCEPI. A similar pattern follows for the financial variables in figure 7, where the RW has the least bias only for the US/UK foreign exchange rate. In contrast, for the other variables, different ML models have the least bias at different times.

In summary, different ML models have the lowest RMSPE over time for real variables, while for nominal and financial variables, the baseline model has the lowest RMSPE. The RMSPE of all models for the real variables increased in the aftermath of the COVID-19 pandemic. Machine learning models showed more bias than the ARD model in forecasting real variables, while the opposite was true for nominal and financial variables.

9 Conclusion

In this chapter, we conducted a comprehensive examination of the forecasting abilities of machine learning (ML) models in comparison with traditional econometric models across a diverse set of macroeconomic variables. Our analysis, spanning a wide range of real, nominal, and financial indicators, provides critical insights into the evolving landscape of economic forecasting.

Firstly, our study reveals that ML models demonstrate superior predictive accuracy for real variables, not only in the complete pre-pandemic sample but also in scenarios involving high volatility, such as economic recessions and the COVID-19 pandemic. While they may fall short in forecasting nominal and financial variables compared to traditional econometric benchmarks, this differential performance suggests that the inherent strengths of ML models – primarily their capacity to uncover non-linear patterns – make them more suited to contexts where such complexities are prevalent.

Additionally, our study underscores the effectiveness of dimension reduction techniques like Principal Component Analysis (PCA) in improving the performance of ML models for certain variables over longer horizons. By distilling information from datasets without introducing unnecessary noise, PCA proves to be a valuable tool for enhancing forecasting accuracy.

During periods of high volatility, ML models consistently outshined benchmark models in forecasting real variables, indicating their resilience and flexibility under challenging environments. The machine learning models show more improvements over the baseline during economic downturns compared to the overall sample.

In conclusion, our chapter contributes to the ongoing debate on the effectiveness of machine learning in economic forecasting. Our exercise points to where these models fit within the spectrum of forecasting tools and under what conditions they are most effective. Looking forward, it suggests a potential avenue for further research into hybrid models that combine the strengths of econometric and machine learning methods to enhance predictive performance across all economic variables. The findings also offer practical implications for policymakers and practitioners in selecting appropriate forecasting models tailored to specific economic indicators and conditions.

Moving forward, we plan to enhance our methodology by exploring dimension reduction techniques, such as regularization methods like Lasso and Ridge, which can improve the model's ability to prevent overfitting and increase prediction accuracy. We also aim to optimize time series models by utilizing model selection criteria such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to identify the model specifications systematically. Furthermore, we need to ensure our results hold after fine-tuning the hyperparameters of our machine-learning models. Through testing and refinement, we seek to strengthen the reliability and accuracy of our forecasting tools so they can effectively perform across diverse economic scenarios.

References

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

Breiman, L. (1996). Bagging predictors. *Machine learning*, 24, 123–140.

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

Chen, T. & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794).

Choudhary, M. A. & Haider, A. (2012). Neural network models for inflation forecasting: an appraisal. *Applied Economics*, 44(20), 2631–2635.

Clark, T. E. & Davig, T. (2009). The relationship between inflation and inflation expectations. *memo to the FOMC, November*, 30.

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accu racy. *Journal of Business and Economic Statistics*, 13(3), 253–263.

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *International workshop on multiple classifier systems* (pp. 1–15).: Springer.

Freund, Y. & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1), 119–139.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, (pp. 1189–1232).

Goulet Coulombe, P., Leroux, M., Stevanovic, D., & Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5), 920–964.

Jørgensen, P. L. & Lansing, K. J. (2019). Anchored inflation expectations and the slope of the phillips curve. *Federal Reserve Bank of San Francisco Working Paper*, 27.

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

Kauffman, R. G. (1999). Indicator qualities of the napm report on business®. *Journal of Supply Chain Management*, 35(1), 29–37.

Kotchoni, R., Leroux, M., & Stevanovic, D. (2019). Macroeconomic forecast accuracy in a data-rich environment. *Journal of Applied Econometrics*, 34(7), 1050–1072.

Ma, Y. & Zhu, L. (2013). A review on dimension reduction. *International Statistical Review*, 81(1), 134–150.

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

Milunovich, G. (2020). Forecasting australia's real house price index: A comparison of time series and machine learning methods. *Journal of Forecasting*, 39(7), 1098–1118.

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

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1, 81–106.

Serrano, C. & Hoesli, M. (2007). Forecasting ereit returns. *Journal of Real Estate Portfolio Management*, 13(4), 293–310.

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

Table 1: List of all forecasting models

Model list	
Baseline models	
ARD	Autoregressive direct
RW	Random walk
Individual machine learning models	
kNN(uniform)	K-nearest neighbor (uniform weighted) regression
kNN(inverse)	K-nearest neighbor (inverse weighted) regression
Decision Tree	Decision tree regression
SVR (linear)	Support vector regression (linear kernel)
SVR (polynomial)	Support vector regression (polynomial kernel)
SVR (rbf)	Support vector regression (rbf kernel)
SVR (sigmoid)	Support vector regression (sigmoid kernel)
Ensemble machine learning models	
Random forest	Random forest regression
XGBoost	XGBoost regression
AdaBoost	AdaBoost regression
Gradient Boost	Gradient boost regression
Individual machine learning models using dimension reduction	
kNN(uniform) - DI	K-nearest neighbor (uniform weighted) regression with diffusion index
kNN(inverse) - DI	K-nearest neighbor (inverse weighted) regression with diffusion index
Decision Tree - DI	Decision tree regression with diffusion index
SVR (linear) - DI	Support vector regression (linear kernel) with diffusion index
SVR (polynomial) - DI	Support vector regression (polynomial kernel) with diffusion index
SVR (rbf) - DI	Support vector regression (rbf kernel) with diffusion index
SVR (sigmoid) - DI	Support vector regression (sigmoid kernel) with diffusion index
Ensemble machine learning models using dimension reduction	
Random Forest - DI	Random forest regression with diffusion index
XGBoost - DI	XGBoost regression with diffusion index
AdaBoost - DI	AdaBoost regression with diffusion index
Gradient Boost - DI	Gradient boost regression with diffusion index

Table 2: Industrial Production growth: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.082	0.061	0.057	0.052	0.048	0.14	0.111	0.101	0.091	0.084
Individual machine learning models										
kNN (uniform)	0.99	0.98	0.93	0.95	0.99	0.98	1.06	1.06	1.09	1.06
kNN (inverse)	0.99	0.98	0.93	0.95	0.99	0.96	1.06	1.06	1.09	1.05
Decision Tree	1.32	1.26	1.20	1.12	1.11	1.23	1.07	1.04	0.89	0.85*
SVR (linear)	1.16	1.15	1.05	1.04	1.11	1.00	1.00	0.87	0.83*	0.73***
SVR (polynomial)	1.03	1.14	1.19	1.15	1.11	1.06	1.16	1.09	1.09	1.07
SVR (rbf)	1.01	1.00	0.93	0.94	0.99	1.04	1.08	1.00	0.96	0.94
SVR (sigmoid)	1.05	1.03	0.98	0.99	1.07	1.01	0.92	0.92	0.88	0.82***
Ensemble machine learning models										
Random Forest	0.97	1.03	1.03	1.02	1.00	0.93	0.90	0.96	0.86	0.78**
XGBoost	1.01	1.00	0.98	1.01	1.01	0.99	0.92	0.88	0.89	0.77*
AdaBoost	0.95**	0.95	1.00	0.99	0.96	0.95	0.89	0.99	0.89	0.76**
Gradient Boost	1.25	1.22	1.19	1.10	1.10	1.13	1.03	1.03	0.84*	0.81*
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	0.97*	0.94	0.87*	0.84	0.85	0.93*	0.90	0.88	0.82*	0.81**
kNN (inverse) - DI	0.96*	0.94	0.88*	0.84	0.85	0.92**	0.90	0.88	0.82*	0.80**
Decision Tree - DI	1.22	1.22	1.25	1.23	1.21	1.07	0.95	0.95	1.00	0.97
SVR (linear) - DI	1.12	1.39	1.33	1.23	1.35	1.05	1.11	1.00	0.99	1.00
SVR (polynomial) - DI	1.91	1.50	1.07	1.14	1.22	2.88	2.06	1.23	0.82	0.86
SVR (rbf) - DI	1.06	1.11	1.01	0.97	1.01	1.17	1.30	1.18	1.03	0.95
SVR (sigmoid) - DI	1.27	1.26	1.19	1.35	1.36	1.27	1.09	1.01	1.22	1.14
Ensemble machine learning models using dimension reduction										
Random Forest - DI	1.21	1.25	1.20	1.22	1.20	1.08	0.97	0.97	0.98	0.97
XGBoost - DI	1.13	1.01	0.99	1.02	1.09	1.22	1.01	0.98	0.86	0.83
AdaBoost - DI	0.97*	0.95	0.98	1.00	1.04	0.95	0.89	0.90	0.81*	0.80*
Gradient Boost - DI	1.22	1.23	1.19	1.19	1.18	1.15	0.98	1.00	0.95	0.93

Table 3: Employment: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					NBER recession periods				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.018	0.014	0.015	0.015	0.016	0.024	0.025	0.026	0.026	0.026
Individual machine learning models										
kNN (uniform)	1.04	1.05	1.03	1.04	1.05	1.28	1.25	1.23	1.19	1.12
kNN (inverse)	1.03	1.04	1.02	1.03	1.04	1.25	1.24	1.23	1.19	1.11
Decision Tree	1.24	1.24	1.32	1.24	1.15	1.32	1.19	1.28	1.00	0.86
SVR (linear)	1.11	1.10	1.07	1.05	1.05	1.13	1.02	0.88	0.72***	0.59***
SVR (polynomial)	1.22	1.49	1.49	1.40	1.33	1.73	1.64	1.39	1.28	1.19
SVR (rbf)	1.07	1.12	1.07	1.06	1.06	1.43	1.39	1.16	0.97	0.86
SVR (sigmoid)	1.07	1.08	1.09	1.04	1.05	1.14	1.01	0.96	0.78**	0.66***
Ensemble machine learning models										
Random Forest	0.92**	0.97	1.08	1.04	1.04	0.93	1.00	1.15	0.92	0.81*
XGBoost	0.95	0.93	0.99	0.98	1.01	1.00	1.02	1.08	0.94	0.83
AdaBoost	0.92**	0.92	1.06	1.07	1.05	0.99	1.03	1.16	0.97	0.83
Gradient Boost	1.20	1.14	1.30	1.20	1.16	1.25	1.08	1.29	0.96	0.89
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	0.98	0.97	0.94	0.94	0.93	1.12	1.12	1.03	0.89	0.79
kNN (inverse) - DI	0.98	0.97	0.94	0.94	0.93	1.11	1.11	1.03	0.89	0.79
Decision Tree - DI	1.21	1.15	1.23	1.19	1.21	1.06	1.12	1.13	1.01	0.95
SVR (linear) - DI	0.99	0.87**	0.90	0.96	1.05	0.94	0.82*	0.80**	0.73***	0.57***
SVR (polynomial) - DI	3.05	4.10	4.18	3.79	3.43	6.04	6.70	5.73	3.01	1.00
SVR (rbf) - DI	1.07	1.13	1.08	1.07	1.08	1.51	1.49	1.25	1.00	0.87
SVR (sigmoid) - DI	3.62	3.55	2.89	2.61	2.41	6.24	4.80	3.26	2.01	1.77
Ensemble machine learning models using dimension reduction										
Random Forest - DI	1.23	1.15	1.22	1.18	1.19	1.05	1.09	1.12	1.00	0.93
XGBoost - DI	0.97	0.91	0.92	0.96	0.98	0.93	0.93	0.96	0.88*	0.81*
AdaBoost - DI	0.92**	0.89*	0.94	0.96	0.98	0.90	0.91	0.99	0.93	0.83*
Gradient Boost - DI	1.17	1.10	1.21	1.18	1.18	1.07	1.07	1.10	1.00	0.94

Table 4: CPI inflation: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.031	0.027	0.023	0.021	0.021	0.052	0.048	0.036	0.032	0.03
Individual machine learning models										
kNN (uniform)	1.11	1.27	1.45	1.53	1.53	0.94	1.14	1.64	1.39	1.38
kNN (inverse)	1.12	1.27	1.45	1.52	1.53	0.94	1.14	1.64	1.38	1.37
Decision Tree	1.39	1.67	1.66	1.65	1.66	1.25	1.65	1.50	1.27	1.35
SVR (linear)	1.29	1.38	1.58	1.58	1.54	1.06	1.07	1.42	1.45	1.49
SVR (polynomial)	1.10	1.33	1.49	1.50	1.50	0.99	1.31	1.53	1.32	1.36
SVR (rbf)	1.10	1.23	1.39	1.47	1.47	0.94	1.11	1.57	1.34	1.35
SVR (sigmoid)	1.17	1.31	1.48	1.57	1.52	1.00	1.10	1.53	1.42	1.38
Ensemble machine learning models										
Random Forest	1.12	1.27	1.36	1.34	1.33	1.02	1.32	1.44	1.16	1.22
XGBoost	1.21	1.25	1.29	1.31	1.32	1.07	1.17	1.33	1.20	1.24
AdaBoost	1.08	1.21	1.26	1.28	1.30	0.94	1.21	1.41	1.18	1.18
Gradient Boost	1.34	1.61	1.54	1.57	1.63	1.17	1.67	1.55	1.30	1.35
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	1.12	1.28	1.43	1.47	1.46	0.96	1.17	1.61	1.29	1.25
kNN (inverse) - DI	1.12	1.28	1.43	1.47	1.46	0.96	1.17	1.61	1.29	1.24
Decision Tree - DI	1.34	1.37	1.49	1.63	1.60	1.11	1.12	1.40	1.36	1.21
SVR (linear) - DI	1.38	1.63	2.09	2.19	2.08	1.26	1.51	2.29	2.08	1.87
SVR (polynomial) - DI	2.11	2.44	2.76	3.14	3.08	1.59	1.56	2.37	3.09	5.00
SVR (rbf) - DI	1.14	1.29	1.53	1.63	1.62	0.96	1.14	1.66	1.41	1.35
SVR (sigmoid) - DI	1.38	1.60	1.88	2.04	1.96	1.20	1.47	2.02	1.90	1.68
Ensemble machine learning models using dimension reduction										
Random Forest - DI	1.35	1.38	1.50	1.66	1.58	1.13	1.12	1.41	1.39	1.20
XGBoost - DI	1.21	1.19	1.32	1.38	1.38	1.13	1.08	1.35	1.22	1.14
AdaBoost - DI	1.06	1.14	1.23	1.28	1.30	<u>0.88</u>	1.01	1.34	1.16	1.18
Gradient Boost - DI	1.31	1.35	1.40	1.62	1.54	1.14	1.10	1.36	1.30	1.11

Table 5: S&P 500: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
RW (RMSPE)	0.438	0.296	0.228	0.193	0.172	0.735	0.489	0.341	0.281	0.241
Individual machine learning models										
kNN (uniform)	1.01	0.99	0.99	1.00	1.00	1.00	1.02	1.10	1.16	1.22
kNN (inverse)	1.01	0.99	0.99	1.00	1.00	1.00	1.02	1.10	1.16	1.21
Decision Tree	1.32	1.37	1.30	1.27	1.15	1.20	1.32	1.12	1.11	1.02
SVR (linear)	1.22	1.30	1.32	1.34	1.35	0.96	1.08	0.98	0.96	1.08
SVR (polynomial)	1.02	1.19	1.37	1.26	1.20	1.07	1.23	1.07	1.18	1.25
SVR (rbf)	0.99	1.01	1.02	1.06	1.08	1.00	1.03	1.05	1.14	1.22
SVR (sigmoid)	1.03	1.10	1.14	1.19	1.27	0.94	1.02	0.99	1.03	1.15
Ensemble machine learning models										
Random Forest	1.02	1.11	1.09	1.12	1.09	1.01	1.17	1.01	1.03	1.05
XGBoost	1.07	1.09	1.11	1.14	1.12	1.05	1.14	0.99	1.06	1.09
AdaBoost	0.98	1.03	1.04	1.03	1.06	0.97	1.06	0.98	0.99	1.06
Gradient Boost	1.25	1.34	1.24	1.26	1.15	1.16	1.31	1.07	1.11	1.03
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	1.02	1.06	1.05	1.03	1.04	1.02	1.04	0.97	1.06	1.15
kNN (inverse) - DI	1.02	1.06	1.06	1.03	1.04	1.01	1.04	0.97	1.06	1.15
Decision Tree - DI	1.29	1.44	1.43	1.33	1.29	1.07	1.30	1.19	1.11	1.21
SVR (linear) - DI	1.05	1.23	1.39	1.54	1.68	1.02	1.13	1.14	1.16	1.28
SVR (polynomial) - DI	1.12	1.10	1.16	1.23	1.38	1.25	1.23	1.05	1.02	1.13
SVR (rbf) - DI	1.02	1.07	1.08	1.12	1.15	1.04	1.10	1.16	1.29	1.34
SVR (sigmoid) - DI	1.10	1.15	1.25	1.28	1.30	1.13	1.22	1.15	1.20	1.35
Ensemble machine learning models using dimension reduction										
Random Forest - DI	1.29	1.43	1.41	1.29	1.28	1.03	1.30	1.18	1.12	1.20
XGBoost - DI	1.14	1.18	1.28	1.22	1.20	1.05	1.22	1.17	1.11	1.21
AdaBoost - DI	0.99	1.05	1.13	1.10	1.11	0.97	1.05	1.04	1.00	1.13
Gradient Boost - DI	1.22	1.36	1.35	1.25	1.28	1.02	1.25	1.16	1.11	1.20

Table 6: Relative RMSPE of the best models for real variables (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample						**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**		**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Industrial Production											
Baseline - ARD	0.082	0.061	0.057	0.052	0.047	Baseline - ARD	0.14	0.111	0.101	0.091	0.084
kNN (inverse)	0.99	1.00	0.92	0.90	0.93	SVR (linear)	1.00	1.00	0.87	0.83*	0.73***
AdaBoost	0.95***	0.94	1.00	1.00	0.96	AdaBoost	0.94	0.88*	0.99	0.88	0.76**
kNN (uniform) - DI	1.02	0.96	0.87*	0.89*	0.94	Random Forest	0.95	0.89	0.96	0.86	0.78**
kNN (inverse) - DI	1.02	0.96	0.88*	0.89*	0.94	Gradient Boost	1.12	1.02	1.03	0.84*	0.82*
AdaBoost - DI	0.97*	0.96	0.99	1.00	1.05	AdaBoost - DI	0.96	0.89	0.90	0.81*	0.79*
Employment											
Baseline - ARD	0.018	0.014	0.015	0.015	0.016	Baseline - ARD	0.024	0.025	0.026	0.026	0.026
Random Forest	0.91**	0.98	1.08	1.04	1.04	SVR (linear)	1.13	1.02	0.88	0.72***	0.59***
AdaBoost	0.92**	0.92	1.06	1.08	1.06	SVR (sigmoid)	1.14	1.01	0.96	0.78**	0.66***
kNN (uniform) - DI	0.98	0.91	0.87	0.89	0.92	AdaBoost	1.00	1.03	1.15	0.96	0.82*
SVR (linear) - DI	0.99	0.87**	0.90	0.96	1.05	SVR (linear) - DI	0.94	0.82*	0.80**	0.73***	0.57***
AdaBoost - DI	0.93**	0.90*	0.93	0.96	0.99	AdaBoost - DI	0.89*	0.92	0.99	0.92	0.84*
Real Personal Income											
Baseline - ARD	0.075	0.038	0.027	0.024	0.022	Baseline - ARD	0.105	0.061	0.045	0.039	0.037
SVR (rbf)	0.92**	0.95	0.94*	0.96	0.98	SVR (rbf)	0.91	0.94	0.92**	0.95	0.96
SVR (sigmoid)	0.94*	0.97	0.96	1.03	1.08	SVR (sigmoid)	0.90	0.88**	0.80**	0.83	0.84*
AdaBoost	0.92**	0.95	0.89***	0.92	0.93	AdaBoost	0.92	0.96	0.84**	0.87	0.88***
SVR (rbf) - DI	0.93*	0.99	1.00	1.02	1.05	SVR (linear)	1.07	1.18	0.93	0.80	0.84*
AdaBoost - DI	0.99	0.95	0.95	0.94	0.95	AdaBoost - DI	0.86*	0.95	0.92*	0.89	0.87
Unemployment Rate											
Baseline - ARD	2.011	1.329	1.212	1.151	1.117	Baseline - ARD	2.711	2.204	1.999	1.846	1.672
kNN (inverse)	0.98	0.95	0.88	0.87*	0.87*	SVR (linear)	1.01	0.91	0.83*	0.71***	0.56***
AdaBoost	0.92***	0.89**	0.92	0.91	0.90	AdaBoost	0.89***	0.92	1.03	0.86	0.74**
kNN (uniform) - DI	1.01	0.90*	0.84*	0.87*	0.89**	XGBoost	0.85**	0.88**	0.99	0.87	0.75**
kNN (inverse) - DI	1.01	0.90*	0.84*	0.86*	0.88**	XGBoost - DI	1.01	0.78**	0.82	0.73**	0.68**
AdaBoost - DI	0.96**	0.87**	0.83**	0.88	0.94	AdaBoost - DI	0.94	0.83**	0.86	0.79**	0.75**
Real Personal Consumption Expenditure											
Baseline - ARD	0.06	0.03	0.022	0.02	0.019	Baseline - ARD	0.086	0.053	0.04	0.038	0.037
SVR (rbf)	0.97**	0.92***	0.95	0.99	1.01	SVR (linear)	1.13	0.93	0.86	0.77**	0.88
SVR (sigmoid)	1.05	0.96	0.99	1.01	1.04	SVR (sigmoid)	1.01	0.81*	0.88*	0.80*	0.86
Random Forest	1.00	0.96	0.96	0.96	1.00	Random Forest	0.92*	0.82**	0.81*	0.85	0.94
AdaBoost	0.98	0.94**	0.93*	0.99	1.00	AdaBoost	0.93	0.85***	0.88*	0.92	0.94
AdaBoost - DI	0.98	0.94*	0.95	0.99	1.01	AdaBoost - DI	0.91*	0.89**	0.89	0.91	1.00

Table 7: Relative RMSPE of the best models for nominal variables (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample						NBER recession periods				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**		**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
CPI: All Items											
Baseline - ARD	0.031	0.027	0.023	0.021	0.021	Baseline - ARD	0.052	0.048	0.036	0.032	0.03
kNN (uniform)	1.12	1.27	1.44	1.51	1.52	kNN (uniform)	0.96	1.14	1.59	1.35	1.36
Random Forest	1.12	1.27	1.37	1.33	1.33	Random Forest	1.02	1.32	1.46	1.15	1.20
AdaBoost	1.07	1.20	1.25	1.29	1.31	AdaBoost	0.92	1.20	1.39	1.19	1.20
AdaBoost - DI	1.05	1.14	1.23	1.28	1.29	AdaBoost - DI	0.89	1.01	1.33	1.16	1.17
Gradient Boost - DI	1.32	1.36	1.41	1.62	1.52	Gradient Boost - DI	1.12	1.10	1.35	1.30	1.09
CPI: All Items Less Food											
Baseline - ARD	0.034	0.03	0.026	0.025	0.024	Baseline - ARD	0.057	0.055	0.043	0.039	0.038
XGBoost	1.14	1.22	1.25	1.28	1.32	XGBoost	1.05	1.18	1.26	0.97	1.04
AdaBoost	1.06	1.19	1.24	1.26	1.29	AdaBoost	0.96	1.19	1.32	1.02	0.98
Decision Tree - DI	1.29	1.34	1.47	1.59	1.55	Decision Tree - DI	0.95	1.16	1.45	1.24	1.19
XGBoost - DI	1.11	1.14	1.29	1.32	1.36	XGBoost - DI	0.98	1.04	1.25	1.06	0.97
AdaBoost - DI	1.02	1.10	1.19	1.24	1.26	AdaBoost - DI	0.89	0.99	1.21	1.01	0.99
PCEPI											
Baseline - ARD	0.022	0.019	0.017	0.016	0.016	Baseline - ARD	0.036	0.034	0.028	0.027	0.026
SVR (rbf)	1.07	1.22	1.35	1.38	1.39	SVR (rbf)	0.95	1.08	1.37	1.15	1.19
Random Forest	1.13	1.28	1.34	1.29	1.29	Random Forest	1.09	1.29	1.25	1.03	1.06
AdaBoost	1.06	1.19	1.24	1.24	1.27	AdaBoost	0.98	1.12	1.23	1.02	1.07
AdaBoost - DI	1.04	1.12	1.22	1.24	1.27	AdaBoost - DI	0.93	0.97	1.21	1.01	1.10
Gradient Boost - DI	1.25	1.26	1.49	1.62	1.64	Gradient Boost - DI	1.07	0.93	1.45	1.24	1.31
PPI: Finished Consumer Goods											
Baseline - ARD	0.093	0.07	0.06	0.055	0.053	Baseline - ARD	0.148	0.118	0.092	0.08	0.065
Random Forest	1.15	1.25	1.42	1.47	1.44	Random Forest	1.04	1.18	1.51	1.33	1.51
AdaBoost	1.11	1.23	1.36	1.42	1.45	AdaBoost	0.96	1.12	1.50	1.37	1.57
SVR (rbf) - DI	1.22	1.40	1.67	1.78	1.83	SVR (rbf) - DI	0.98	1.19	1.69	1.65	1.92
XGBoost - DI	1.25	1.35	1.41	1.50	1.50	XGBoost - DI	1.15	1.34	1.40	1.41	1.50
AdaBoost - DI	1.11	1.22	1.36	1.43	1.45	AdaBoost - DI	0.97	1.12	1.46	1.38	1.58
PPI: Crude Metals											
Baseline - ARD	0.457	0.321	0.284	0.267	0.256	Baseline - ARD	0.681	0.552	0.51	0.458	0.362
SVR (rbf)	1.16	1.39	1.56	1.64	1.68	SVR (rbf)	0.91	1.21	1.50	1.55	1.66
Random Forest	1.21	1.30	1.51	1.56	1.54	Random Forest	0.98	1.20	1.50	1.38	1.41
AdaBoost	1.14	1.29	1.43	1.48	1.50	AdaBoost	0.91	1.22	1.42	1.40	1.43
SVR (rbf) - DI	1.20	1.43	1.65	1.77	1.83	SVR (rbf) - DI	0.88	1.22	1.56	1.63	1.80
AdaBoost - DI	1.15	1.27	1.38	1.45	1.49	AdaBoost - DI	0.90	1.18	1.41	1.43	1.49

Table 8: Relative RMSPE of the best models for financial variables (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample						NBER recession periods				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**		**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
S&P 500											
Baseline - RW	0.438	0.296	0.228	0.193	0.172	Baseline - RW	0.735	0.489	0.341	0.281	0.241
kNN (uniform)	1.01	0.99	0.99	1.00	1.00	SVR (linear)	0.96	1.08	0.98	0.96	1.08
kNN (inverse)	1.01	0.99	0.99	1.00	1.00	SVR (sigmoid)	0.94	1.02	0.99	1.03	1.15
SVR (rbf)	0.99	1.01	1.02	1.06	1.08	AdaBoost	0.97	1.06	0.98	0.99	1.06
AdaBoost	0.98	1.03	1.04	1.03	1.06	kNN (uniform)-DI	1.02	1.04	0.97	1.06	1.15
AdaBoost-DI	0.99	1.05	1.13	1.10	1.11	AdaBoost-DI	0.97	1.05	1.04	1.00	1.13
1-Year Treasury Rate											
Baseline - RW	5.222	3.623	2.522	1.972	1.759	Baseline - RW	10.813	7.287	4.084	2.936	2.544
SVR (rbf)	1.01	1.01	1.01	1.05	1.12	SVR (rbf)	1.00	1.00	0.98	0.94	0.91
SVR (sigmoid)	1.08	1.11	1.11	1.22	1.33	SVR (sigmoid)	1.01	0.97	1.04	1.05	0.98
AdaBoost	0.98	1.04	1.06	1.17	1.30	AdaBoost	0.99	0.99	0.99	1.26	1.21
SVR (sigmoid) - DI	1.20	1.28	1.37	1.39	1.40	SVR (sigmoid) - DI	1.10	1.11	0.98	1.16	1.02
AdaBoost - DI	1.03	1.03	1.02	1.06	1.17	AdaBoost - DI	0.98	1.01	0.94	0.97	1.02
10-Year Treasury Rate											
Baseline - RW	3.518	2.366	1.684	1.361	1.209	Baseline - RW	6.065	3.918	2.179	1.683	1.47
SVR (polynomial)	1.04	1.10	1.11	1.14	1.19	SVR (polynomial)	1.08	1.11	1.01	1.02	1.01
SVR (rbf)	1.03	1.05	1.08	1.14	1.18	SVR (rbf)	1.03	1.05	1.07	1.05	0.99
SVR (sigmoid)	1.11	1.13	1.20	1.29	1.35	SVR (sigmoid)	1.14	1.14	1.15	1.11	0.99
AdaBoost	1.01	1.08	1.13	1.25	1.34	AdaBoost	1.03	1.08	1.05	1.42	1.28
kNN (uniform) - DI	1.03	1.11	1.13	1.15	1.18	kNN (uniform) - DI	1.00	1.08	1.07	1.02	1.02
US/UK Foreign Exchange Rate											
Baseline - RW	0.278	0.195	0.147	0.121	0.104	Baseline - RW	0.33	0.256	0.207	0.176	0.151
SVR (rbf)	1.03	1.06	1.07	1.07	1.04	SVR (sigmoid)	1.20	1.12	1.12	1.02	0.88*
kNN (uniform) - DI	1.03	1.03	1.04	1.04	1.08	kNN (uniform) - DI	0.99	1.00	0.98	0.97	0.97
kNN (inverse) - DI	1.03	1.04	1.05	1.04	1.08	kNN (inverse) - DI	0.99	0.99	0.98	0.97	0.97
SVR (rbf) - DI	1.06	1.09	1.10	1.11	1.16	XGBoost - DI	0.98	1.04	1.06	1.12	1.08
AdaBoost - DI	1.00	1.11	1.15	1.16	1.16	AdaBoost - DI	0.89**	1.06	1.03	1.07	1.04
CA/US Foreign Exchange Rate											
Baseline - RW	0.174	0.117	0.087	0.073	0.064	Baseline - RW	0.236	0.165	0.118	0.095	0.081
Random Forest	1.07	1.21	1.26	1.25	1.26	Random Forest	1.15	1.35	1.11	0.93	0.93
AdaBoost	1.01	1.13	1.20	1.20	1.29	AdaBoost	1.00	1.25	1.00	1.01	0.98
SVR (rbf)	1.02	1.05	1.09	1.10	1.11	XGBoost - DI	1.13	1.32	1.13	1.10	0.93
kNN (inverse) - DI	1.05	1.11	1.14	1.15	1.21	kNN (inverse) - DI	0.98	1.00	1.13	1.16	1.07
AdaBoost - DI	1.01	1.12	1.22	1.18	1.18	AdaBoost - DI	0.93	1.15	1.22	1.21	1.09

Table 9: Real Variables: RMSPE relative to ARDI (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample						**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**		**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Industrial Production											
Baseline - ARDI	0.089	0.065	0.058	0.057	0.061	Baseline - ARDI	0.134	0.087	0.075	0.067	0.058
kNN (uniform)	0.92**	0.93	0.89	0.83	0.72	kNN (uniform)	1.00	1.26	1.28	1.24	1.36
kNN (inverse)	0.91**	0.93	0.89	0.83	0.72	kNN (inverse)	1.00	1.26	1.28	1.24	1.36
AdaBoost	0.87***	0.88**	0.97	0.91	0.75	Random Forest	0.99	1.13	1.29	1.17	1.13
kNN (uniform) - DI	0.94	0.90	0.85	0.82	0.73	AdaBoost	0.98	1.12	1.34	1.20	1.11
kNN (inverse) - DI	0.94	0.90	0.85	0.82	0.73	AdaBoost - DI	1.00	1.13	1.22	1.11	1.14
Employment											
Baseline - ARDI	0.019	0.015	0.015	0.016	0.017	Baseline - ARDI	0.026	0.02	0.019	0.016	0.012
Random Forest	0.83***	0.97	1.05	1.02	0.97	Random Forest	0.87	1.23	1.51	1.53	1.71
XGBoost	0.86***	0.91	0.96	0.97	0.94	XGBoost	0.92	1.23	1.44	1.57	1.77
AdaBoost	0.84***	0.91	1.03	1.06	0.98	SVR (linear) - DI	0.86**	0.99	1.06	1.22	1.21
kNN (uniform) - DI	0.90**	0.90	0.84	0.88	0.85	XGBoost - DI	0.86*	1.12	1.28	1.48	1.73
SVR (linear) - DI	0.90***	0.85**	0.88	0.95	0.97	AdaBoost - DI	0.83**	1.11	1.32	1.54	1.79
Real Personal Income											
Baseline - ARDI	0.091	0.044	0.029	0.027	0.028	Baseline - ARDI	0.128	0.065	0.04	0.032	0.03
SVR (rbf)	0.76***	0.82***	0.86***	0.84	0.78	SVR (linear)	0.89	1.10	1.06	0.97	1.05
SVR (sigmoid)	0.77***	0.84***	0.89**	0.90	0.85	SVR (sigmoid)	0.75**	0.83	0.91	1.00	1.05
AdaBoost	0.76***	0.82***	0.82***	0.80*	0.73	AdaBoost	0.76**	0.90	0.95	1.05	1.11
SVR (rbf) - DI	0.77***	0.85***	0.92	0.89	0.83	SVR (rbf) - DI	0.74***	0.91	1.10	1.26	1.35
AdaBoost - DI	0.82***	0.83***	0.88**	0.82	0.75	AdaBoost - DI	0.71***	0.89	1.04	1.08	1.10
Unemployment Rate											
Baseline - ARDI	2.154	1.311	1.143	1.118	1.161	Baseline - ARDI	2.827	1.761	1.228	1.025	0.861
kNN (inverse)	0.91***	0.96	0.94	0.90	0.84	SVR (linear)	0.97	1.14	1.35	1.27	1.09
SVR (rbf)	0.91**	0.96	0.97	0.96	0.90	Random Forest	0.82*	1.14	1.67	1.52	1.37
AdaBoost	0.86***	0.90	0.97	0.94	0.86	XGBoost	0.82*	1.10	1.62	1.57	1.45
kNN (inverse) - DI	0.95	0.91	0.89	0.89	0.85	AdaBoost	0.85*	1.16	1.68	1.55	1.43
AdaBoost - DI	0.90***	0.89**	0.88	0.90**	0.91	XGBoost - DI	0.97	0.98	1.33	1.31	1.32
Real Personal Consumption Expenditure											
Baseline - ARDI	0.07	0.034	0.025	0.026	0.025	Baseline - ARDI	0.11	0.054	0.035	0.033	0.032
kNN (uniform)	0.85***	0.86*	0.88	0.78	0.80	SVR (linear)	0.88	0.91	0.97	0.88	1.02
SVR (rbf)	0.84***	0.82**	0.82	0.76	0.77	SVR (rbf)	0.72***	0.89	1.09	1.15	1.19
Random Forest	0.87***	0.86**	0.83	0.74	0.76	SVR (sigmoid)	0.78**	0.80*	0.99	0.92	0.99
AdaBoost	0.84***	0.84**	0.80*	0.76	0.76	Random Forest	0.72***	0.81	0.92	0.98	1.10
AdaBoost - DI	0.84***	0.84**	0.82	0.76	0.77	AdaBoost - DI	0.71***	0.87	1.01	1.04	1.16

Table 10: Industrial Production Growth: relative RMSPE (sample period: 1960m1-2023m12)

	Full out-of-sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.17	0.098	0.07	0.06	0.054	0.245	0.139	0.109	0.095	0.086
Individual machine learning models										
kNN (uniform)	0.69	0.82	0.85	0.86	0.86	0.99	0.99	0.97	0.94	0.96
kNN (inverse)	0.69	0.82	0.85	0.86	0.86	0.99	0.99	0.97	0.94	0.96
Decision Tree	0.92	1.15	1.10	1.05	1.03	1.06	1.05	1.01	0.91	0.85*
SVR (linear)	0.88	1.18	1.09	1.08	1.06	1.01	1.00	0.91	0.85*	0.77***
SVR (polynomial)	0.73	0.88	1.06	1.05	1.01	1.03	1.12	1.08	1.09	1.07
SVR (rbf)	0.68	0.79	0.84	0.88	0.92	1.02	1.05	1.01	0.98	0.95
SVR (sigmoid)	0.74	0.85	0.91	0.93	0.99	1.01	0.96	0.94	0.89	0.86**
Ensemble machine learning models										
Random Forest	0.75	0.96	0.97	0.96	0.93	0.96**	0.94	0.98	0.88	0.80**
XGBoost	0.87	1.06	1.07	0.97	0.94	1.02	0.95	0.91	0.90	0.81*
AdaBoost	0.68	0.99	0.93	0.93	0.89	0.98	0.93	0.99	0.91	0.78**
Gradient Boost	0.83	1.09	1.08	1.06	1.01	0.95	1.03	1.00	0.89	0.85*
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	0.70	0.81	0.81	0.84*	0.87	1.02	1.04	1.00	0.97	0.93
kNN (inverse) - DI	0.70	0.81	0.81	0.84*	0.87	1.02	1.04	1.00	0.97	0.93
Decision Tree - DI	0.85	1.07	1.15	1.15	1.17	0.96	0.96	1.03	0.95	0.90
SVR (linear) - DI	0.81	1.04	1.09	1.12	1.22	1.06	1.06	0.98	1.01	1.02
SVR (polynomial) - DI	1.06	1.20	1.26	1.18	1.11	1.67	1.87	0.98	0.83	0.89
SVR (rbf) - DI	0.71	0.85	0.90	0.91	0.93	1.07	1.21	1.17	1.04	0.97
SVR (sigmoid) - DI	0.79	1.00	1.04	1.20	1.24	1.10	1.15	1.04	1.25	1.17
Ensemble machine learning models using dimension reduction										
Random Forest - DI	0.85	1.07	1.14	1.14	1.15	0.96	0.96	1.03	0.95	0.89
XGBoost - DI	0.94	0.94	0.96	1.01	1.06	1.10	1.02	0.99	0.93	0.91
AdaBoost - DI	0.68	0.78	0.93	0.94	1.04	0.98	0.94	0.95	0.86	0.84*
Gradient Boost - DI	0.87	1.11	1.14	1.19	1.13	1.05	1.01	1.06	0.94	0.87

Table 11: Employment: relative RMSPE (sample period: 1960m1-2023m12)

	Full out-of-sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.694	0.091	0.068	0.069	0.031	0.204	0.077	0.044	0.036	0.032
Individual machine learning models										
kNN (uniform)	0.11	0.50	0.45	0.36	0.73	1.00	1.04	1.08	1.05	1.03
kNN (inverse)	0.11	0.50	0.45	0.36	0.73	1.00	1.04	1.08	1.05	1.02
Decision Tree	0.16	0.82	0.64	0.46	0.80	1.01	1.01	1.12	0.99	0.93
SVR (linear)	0.15	0.83	0.61	0.51	0.88	1.00	1.01	0.97	0.86***	0.76***
SVR (polynomial)	0.12	0.50	0.50	0.41	0.83	1.01	1.09	1.13	1.15	1.12
SVR (rbf)	0.11	0.47	0.42	0.35	0.72	1.01	1.04	1.05	0.98	0.90
SVR (sigmoid)	0.12	0.50	0.43	0.35	0.74	1.00	1.00	0.99	0.88***	0.79***
Ensemble machine learning models										
Random Forest	0.13	0.67	0.50	0.40	0.76	1.00	1.00	1.05	0.95	0.88*
XGBoost	0.17	0.82	0.66	0.37	0.76	1.00	1.00	1.03	0.96	0.89
AdaBoost	0.11	0.78	0.52	0.39	0.72	1.00	1.00	1.05	0.97	0.88*
Gradient Boost	0.14	0.84	0.58	0.43	0.82	1.00	1.01	1.11	0.98	0.92
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	0.11	0.50	0.43	0.34	0.68	1.00	1.01	1.01	0.94	0.89*
kNN (inverse) - DI	0.11	0.50	0.44	0.34	0.68	1.00	1.01	1.01	0.94	0.89*
Decision Tree - DI	0.13	0.70	0.52	0.51	0.96	1.00	1.01	1.05	1.01	0.98
SVR (linear) - DI	0.12	0.50	0.44	0.38	0.81	1.00	0.98**	0.95**	0.87***	0.76***
SVR (polynomial) - DI	0.14	0.72	0.89	0.76	1.54	1.18	2.12	2.97	2.17	1.00
SVR (rbf) - DI	0.11	0.47	0.43	0.36	0.75	1.01	1.06	1.09	1.01	0.92
SVR (sigmoid) - DI	0.15	0.69	0.70	0.63	1.32	1.26	1.66	1.91	1.52	1.37
Ensemble machine learning models using dimension reduction										
Random Forest - DI	0.13	0.63	0.52	0.51	0.98	1.00	1.01	1.05	1.00	0.98
XGBoost - DI	0.18	0.74	0.54	0.44	0.97	1.00	0.99	1.01	0.97	0.92
AdaBoost - DI	0.11	0.49	0.49	0.45	1.00	1.00*	0.99	1.01	0.98	0.90
Gradient Boost - DI	0.12	0.68	0.57	0.53	1.08	1.00	1.00	1.04	1.01	0.98

Table 12: Real Personal Income: relative RMSPE (sample period: 1960m1-2023m12)

	Full out-of-sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.162	0.063	0.043	0.034	0.028	0.184	0.075	0.05	0.041	0.037
Individual machine learning models										
kNN (uniform)	0.97	1.01	0.89	0.89	1.00	1.00	1.00	1.02	1.06	1.05
kNN (inverse)	0.97	1.01	0.89	0.89	1.00	1.00	1.00	1.02	1.06	1.05
Decision Tree	1.15	1.28	1.11	1.09	1.16	1.16	1.02	1.08	0.97	1.03
SVR (linear)	1.07	1.23	1.11	1.18	1.29	1.09	1.14	0.98	0.89**	0.88*
SVR (polynomial)	0.96	1.00	0.94	1.00	1.10	1.05	1.09	0.98	1.04	1.03
SVR (rbf)	0.95	0.99	0.86	0.87	0.97	1.00	0.97	0.93**	0.97	0.96
SVR (sigmoid)	0.95	0.97	0.90	0.97	1.11	1.00	0.93*	0.87*	0.86	0.86*
Ensemble machine learning models										
Random Forest	1.04	1.07	0.98	0.95	1.12	1.05	0.98	0.91	0.91	0.97
XGBoost	1.41	1.34	0.95	0.90	1.09	1.04	1.02	0.98	0.94	0.93
AdaBoost	1.02	1.03	0.87	0.85	0.94	0.99	0.97	0.88*	0.90*	0.88***
Gradient Boost	1.41	1.20	1.03	1.10	1.37	1.06	1.05	0.98	1.01	1.02
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	0.96	1.00	0.88	0.89	1.00	1.00	1.02	1.02	1.06	1.05
kNN (inverse) - DI	0.96	1.00	0.89	0.89	1.00	1.01	1.02	1.02	1.05	1.05
Decision Tree - DI	1.17	1.09	0.96	0.97	1.25	0.98	0.97	1.01	0.90	0.97
SVR (linear) - DI	0.99	1.03	1.02	1.08	1.21	1.02	0.92	1.06	1.08	1.03
SVR (polynomial) - DI	1.05	1.46	1.56	1.67	2.11	1.19	1.65	1.52	1.12	1.00
SVR (rbf) - DI	0.98	1.01	0.89	0.90	1.02	0.99	0.98	0.98	1.04	1.06
SVR (sigmoid) - DI	1.06	1.11	1.09	1.03	1.18	1.00	0.96	1.04	1.04	0.99
Ensemble machine learning models using dimension reduction										
Random Forest - DI	1.16	1.07	0.97	0.97	1.24	0.98	0.96	1.01	0.92	0.99
XGBoost - DI	1.18	1.11	0.86	0.91	1.11	1.07	0.96	1.00	0.96	0.98
AdaBoost - DI	1.01	0.96	0.85	0.88	1.10	0.97	0.93**	0.93*	0.92	0.92
Gradient Boost - DI	1.28	1.21	0.93	1.01	1.26	1.10	0.89	0.97	0.92	1.00

Table 13: Unemployment Rate: relative RMSPE (sample period: 1960m1-2023m12)

	Full out-of-sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	18.839	6.249	3.762	2.464	2.183	14.913	5.716	3.345	2.578	2.14
Individual machine learning models										
kNN (uniform)	0.30	0.54	0.59	0.69	0.65	0.97	1.00	1.00	0.96	0.94***
kNN (inverse)	0.30	0.54	0.59	0.69	0.65	0.97	1.00	1.00	0.96	0.94***
Decision Tree	0.38	0.88	0.74	0.85	0.83	0.98	1.03	1.03	0.96	0.85**
SVR (linear)	0.40	0.90	0.82	1.03	0.85	0.97	0.99	0.96	0.87***	0.76***
SVR (polynomial)	0.32	0.54	0.63	0.75	0.72	0.98	1.04	1.03	1.04	1.04
SVR (rbf)	0.30	0.51	0.58	0.70	0.69	0.98	1.01	1.01	0.96	0.92**
SVR (sigmoid)	0.32	0.55	0.58	0.70	0.69	0.97	0.98*	0.98	0.89**	0.82***
Ensemble machine learning models										
Random Forest	0.34	0.72	0.69	0.77	0.76	0.97	0.99	1.01	0.93*	0.84***
XGBoost	0.50	0.93	0.89	0.81	0.76	0.97	0.98*	0.99	0.93	0.86**
AdaBoost	0.29	0.79	0.68	0.72	0.77	0.97	0.99	1.01	0.92*	0.86***
Gradient Boost	0.36	0.91	0.74	0.85	0.84	0.96	1.02	1.03	0.95	0.87**
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	0.30	0.56	0.60	0.69	0.67	0.98	1.00	1.00	0.97	0.95
kNN (inverse) - DI	0.30	0.56	0.60	0.69	0.67	0.98	1.00	1.00	0.97	0.95
Decision Tree - DI	0.36	0.71	0.81	1.05	0.97	0.94	0.99	0.97	0.95	0.91
SVR (linear) - DI	0.36	0.61	0.65	0.84	0.82	0.96	0.99	0.98	0.92***	0.87***
SVR (polynomial) - DI	0.34	0.52	0.65	0.80	0.75	0.99	0.87	1.05	0.88*	0.90
SVR (rbf) - DI	0.30	0.52	0.59	0.72	0.72	0.98	1.03	1.03	1.00	0.97
SVR (sigmoid) - DI	0.35	0.61	0.60	0.84	0.86	1.02	1.21	1.10	1.04	1.01
Ensemble machine learning models using dimension reduction										
Random Forest - DI	0.35	0.68	0.78	1.02	0.96	0.94	0.99	0.97	0.95	0.90
XGBoost - DI	0.51	0.79	0.81	0.95	1.00	0.99	0.96**	0.92**	0.88*	0.82**
AdaBoost - DI	0.29	0.50	0.71	0.98	0.93	0.97	0.97**	0.96	0.92*	0.86**
Gradient Boost - DI	0.37	0.72	0.82	1.08	1.01	1.02	0.99	0.96	0.96	0.89*

Table 14: Real PCE: relative RMSPE (sample period: 1960m1-2023m12)

	Full out-of-sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.276	0.064	0.048	0.03	0.043	0.202	0.108	0.061	0.049	0.043
Individual machine learning models										
kNN (uniform)	0.37*	0.90	0.73	0.94	0.58	1.01	0.99	1.02	1.03	1.08
kNN (inverse)	0.37*	0.90	0.73	0.94	0.58	1.01	0.99	1.02	1.03	1.08
Decision Tree	0.56	1.35	0.94	1.21	0.75	1.06	0.99	0.96	0.96	0.96
SVR (linear)	0.49*	1.46	1.04	1.26	0.69	1.05	0.99	0.97	0.85**	0.91
SVR (polynomial)	0.38*	0.89	0.76	0.98	0.61	1.03	1.01	1.05	1.07	1.09
SVR (rbf)	0.36*	0.85	0.71	0.94	0.58	1.01	0.98*	0.99	1.00	1.02
SVR (sigmoid)	0.38*	0.92	0.72	0.94	0.58	1.02	0.97	0.97	0.88*	0.90*
Ensemble machine learning models										
Random Forest	0.44*	1.11	0.82	1.04	0.62	0.95	0.97	0.93	0.91	0.94
XGBoost	0.46*	1.36	1.12	1.03	0.65	1.00	0.99	0.97	0.99	1.02
AdaBoost	0.39*	1.13	0.77	0.99	0.57	1.01	0.97*	0.95	0.96	0.97
Gradient Boost	0.46*	1.23	0.97	1.13	0.76	0.98	0.98	0.96	0.93	0.98
Individual machine learning models using dimension reduction										
kNN (uniform) - DI	0.37*	0.94	0.75	0.95	0.58	0.99	0.98*	1.00	1.03	1.05
kNN (inverse) - DI	0.37*	0.94	0.75	0.95	0.57	0.99	0.98*	1.00	1.03	1.05
Decision Tree - DI	0.50	1.16	1.04	1.29	0.74	1.06	1.00	1.04	1.05	1.07
SVR (linear) - DI	0.46*	1.07	0.87	1.16	0.69	1.04	0.97	0.99	0.97	1.00
SVR (polynomial) - DI	0.41*	0.93	0.90	1.24	0.70	1.08	0.97	1.03	1.03	1.09
SVR (rbf) - DI	0.37*	0.87	0.72	0.94	0.59	1.01	1.00	1.02	1.03	1.05
SVR (sigmoid) - DI	0.40*	0.82	0.68	1.17	0.72	1.15	0.97	0.95	1.00	1.04
Ensemble machine learning models using dimension reduction										
Random Forest - DI	0.50	1.18	1.01	1.25	0.74	1.06	1.01	1.02	1.05	1.06
XGBoost - DI	0.52*	1.25	0.86	1.16	0.68	1.07	1.00	0.99	1.01	1.05
AdaBoost - DI	0.36*	0.85	0.78	1.11	0.64	1.00	0.97*	0.95	0.96	1.01
Gradient Boost - DI	0.46*	1.33	1.07	1.48	0.84	1.06	1.05	0.97	1.04	1.08

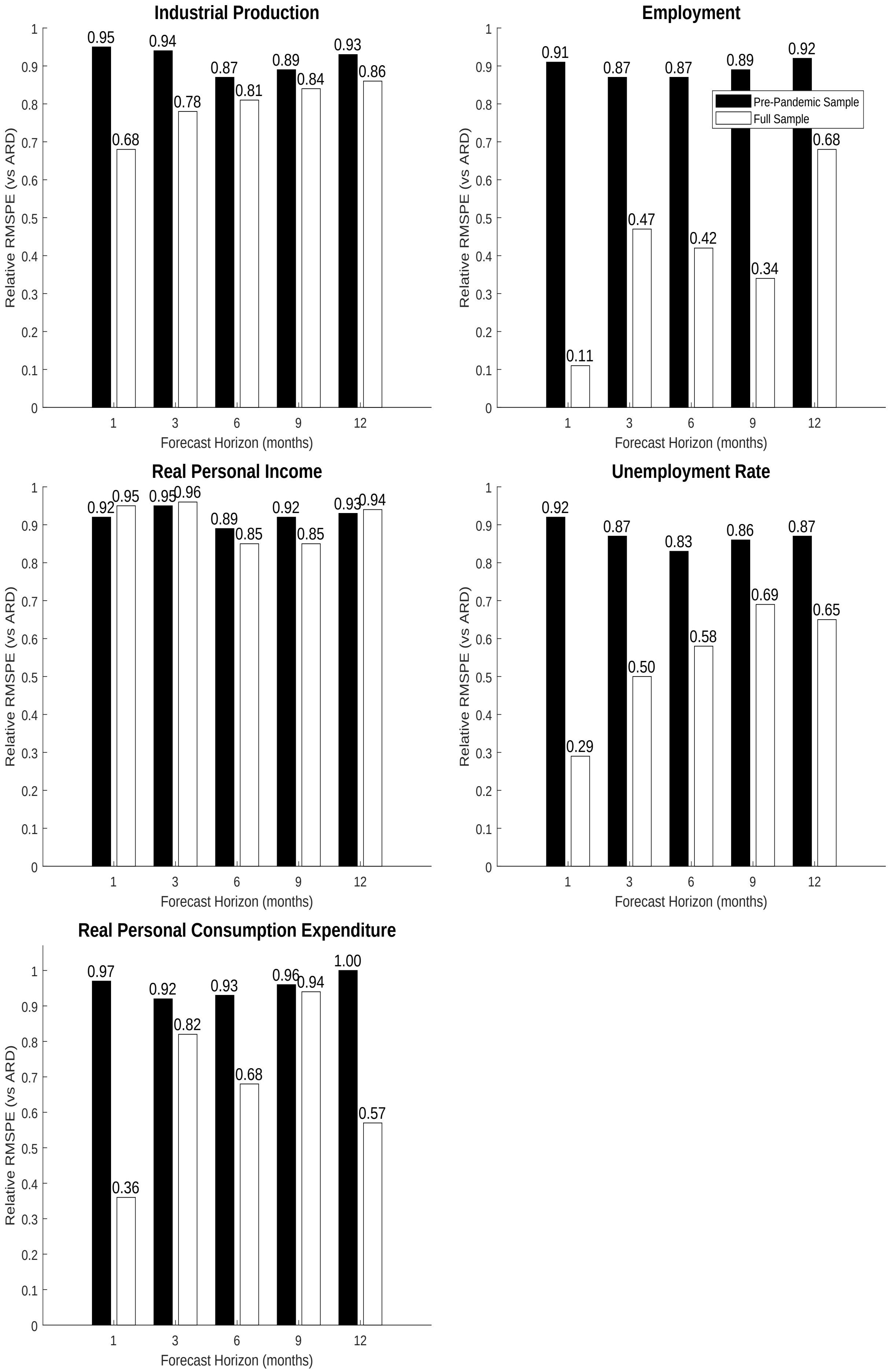
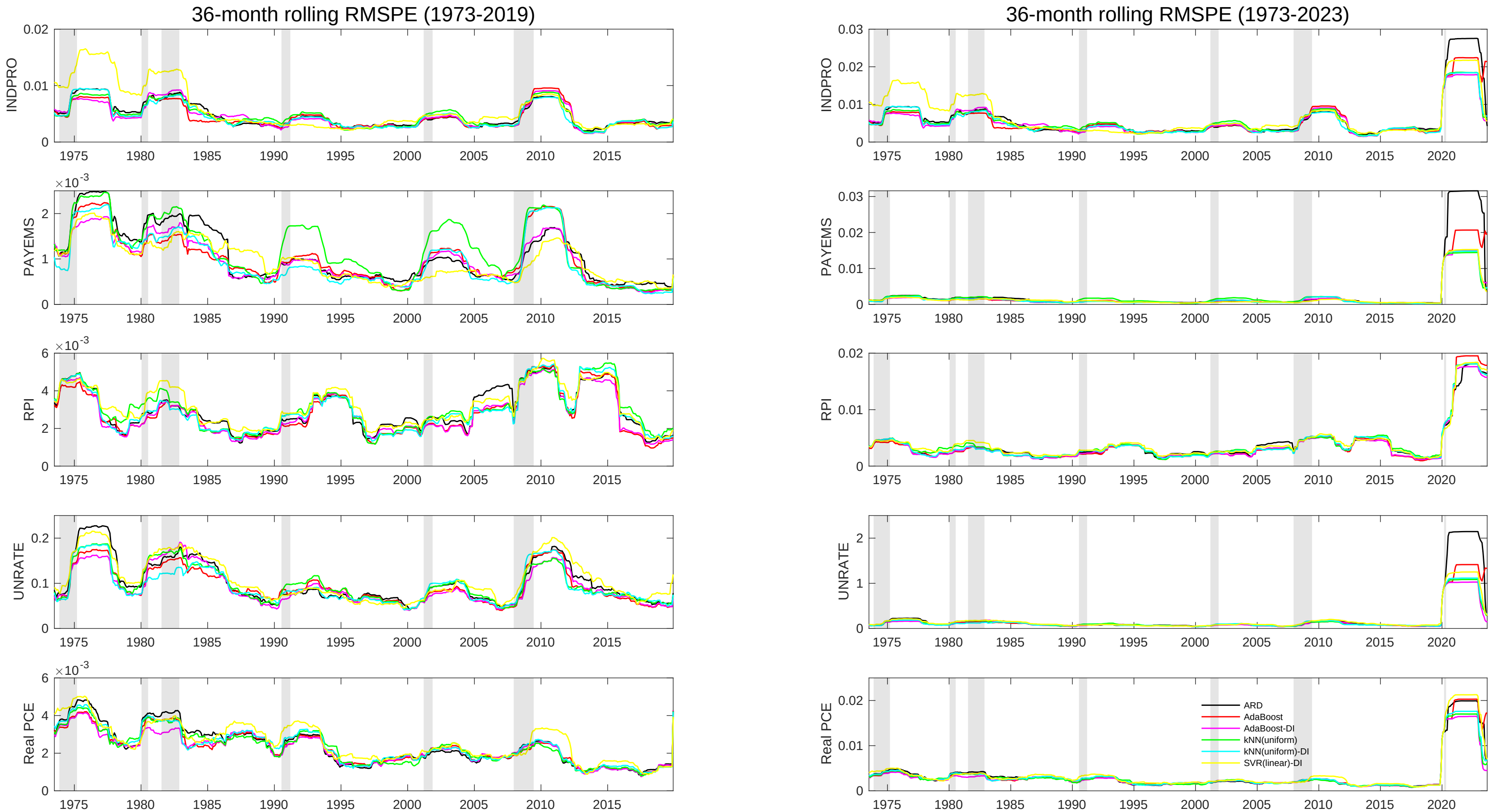


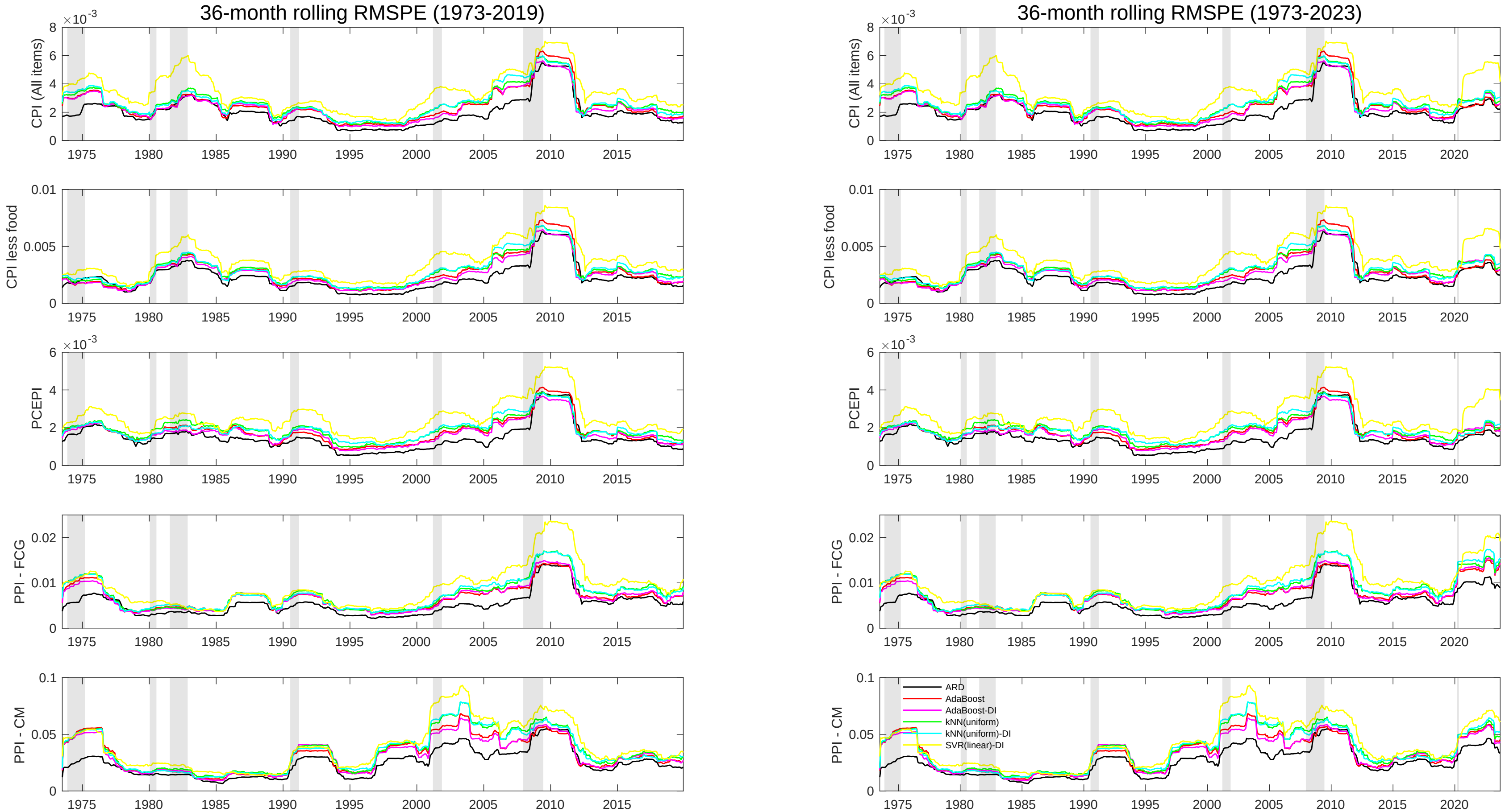
Figure 1: Relative RMSPE Pre-Post Covid

Figure 2: **RMSPE over time for real variables**



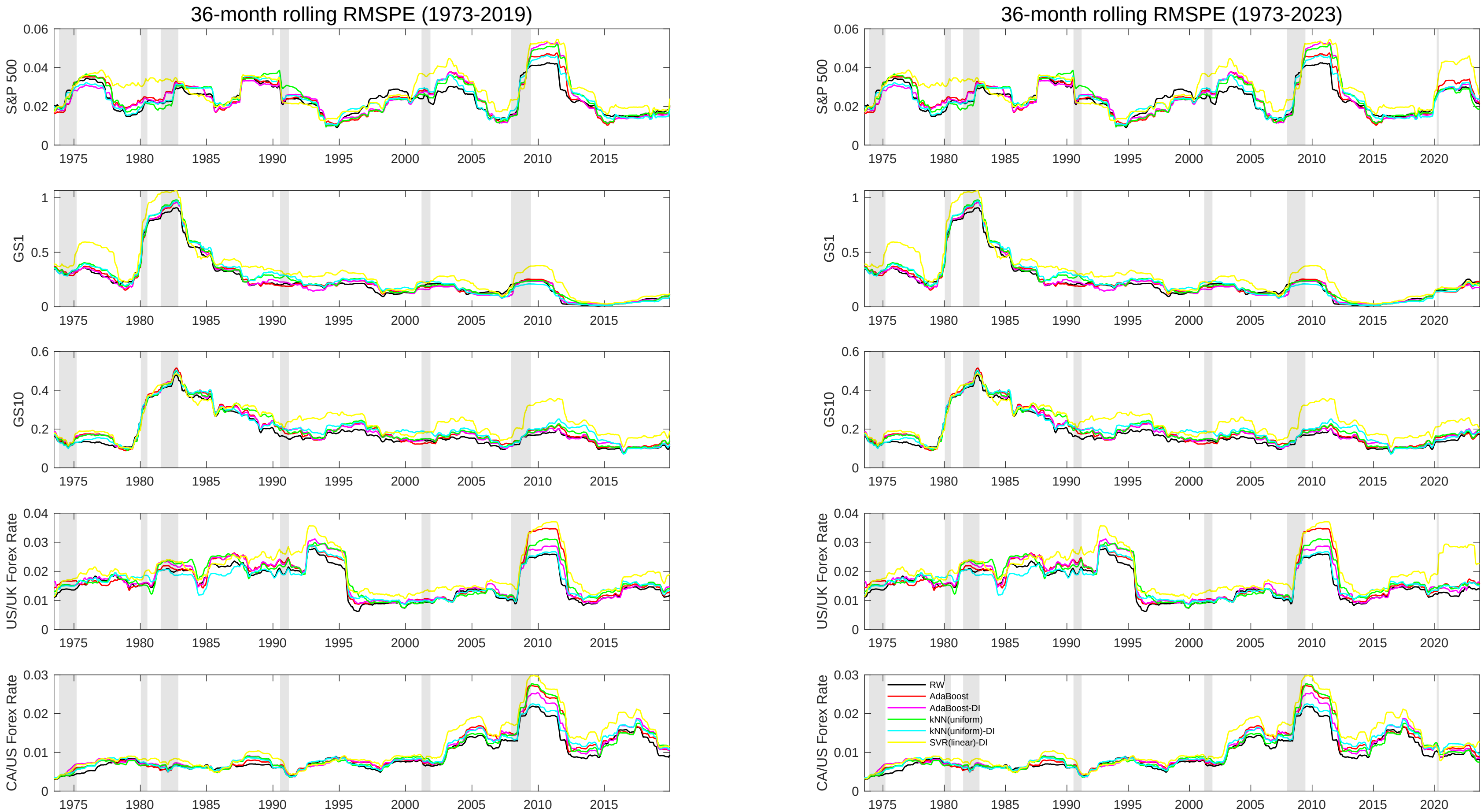
Notes: The figure shows the 3-year moving average of the RMSPE of the real variables for selected models at $h = 3$. The left panel covers the original sample between 1973-2019, while the right panel includes obersvations till 2023.

Figure 3: **RMSPE over time for nominal variables**



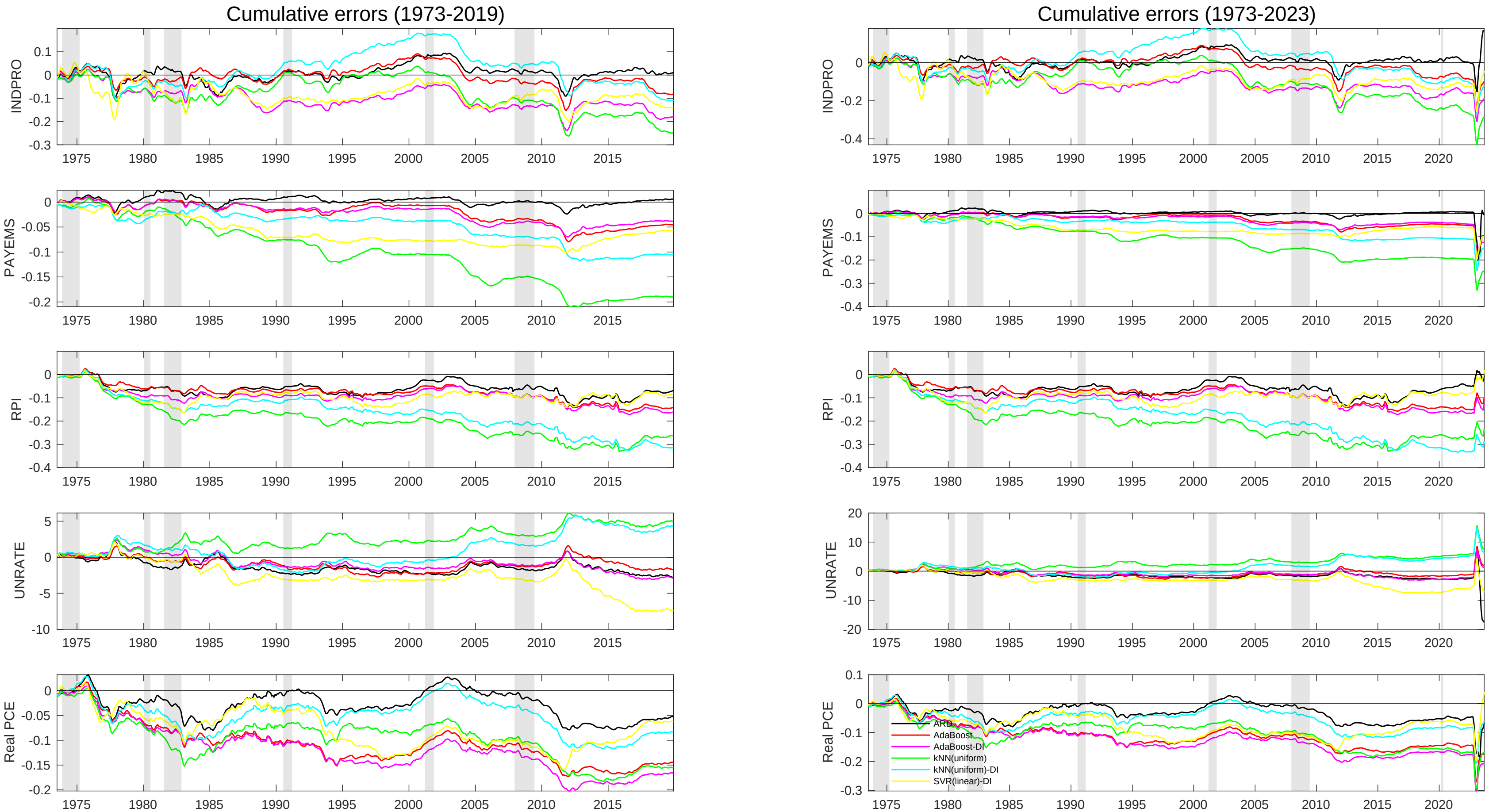
Notes: The figure shows the 3-year moving average of the RMSPE of the nominal variables for selected models at $h = 3$. The left panel covers the original sample between 1973-2019, while the right panel includes observations till 2023.

Figure 4: **RMSPE over time for financial variables**



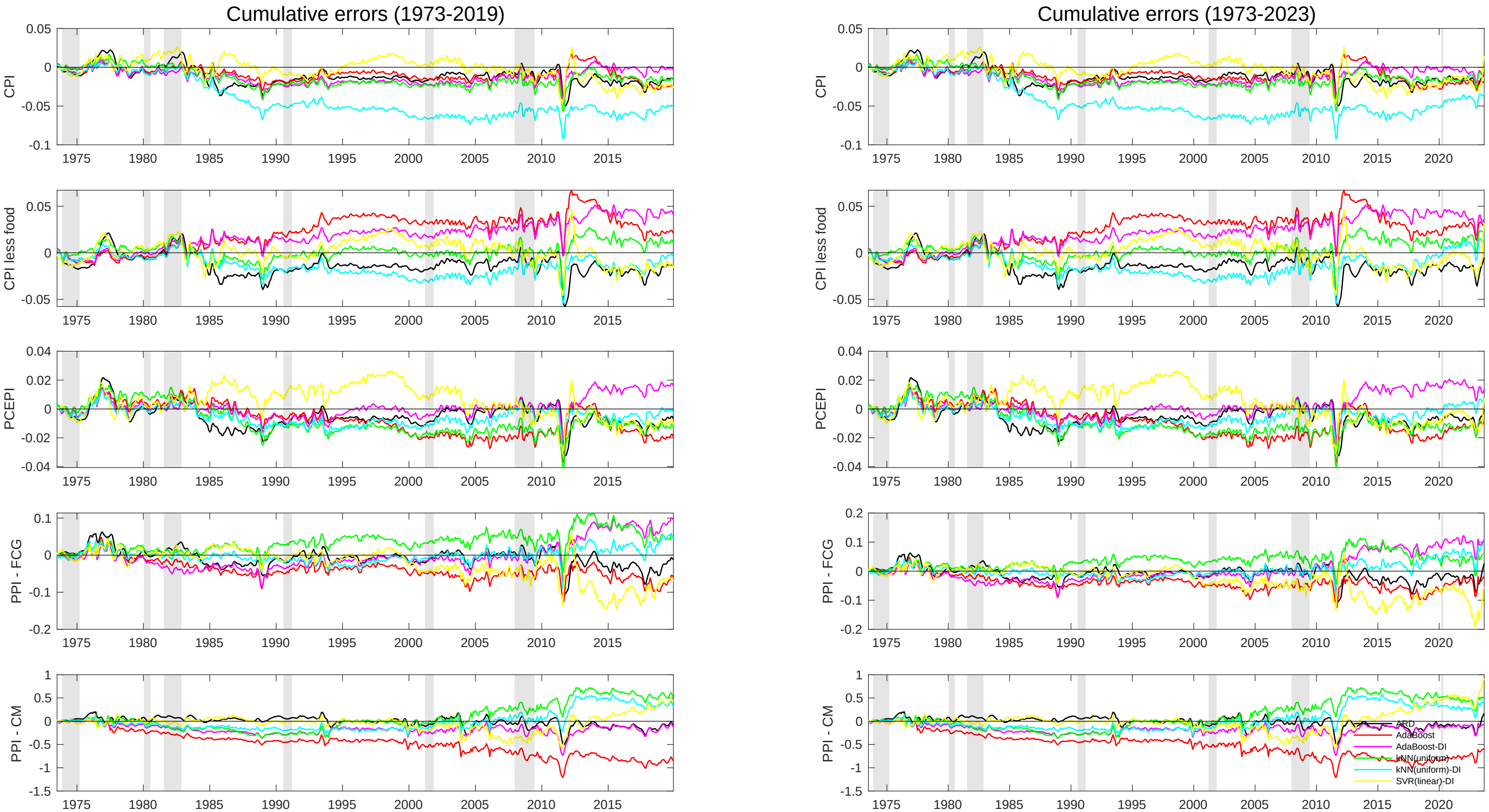
Notes: The figure shows the 3-year moving average of the RMSPE of the real variables for selected models at $h = 3$. The left panel covers the original sample between 1973-2019, while the right panel includes observations till 2023.

Figure 5: **Cumulative errors over time for real variables**



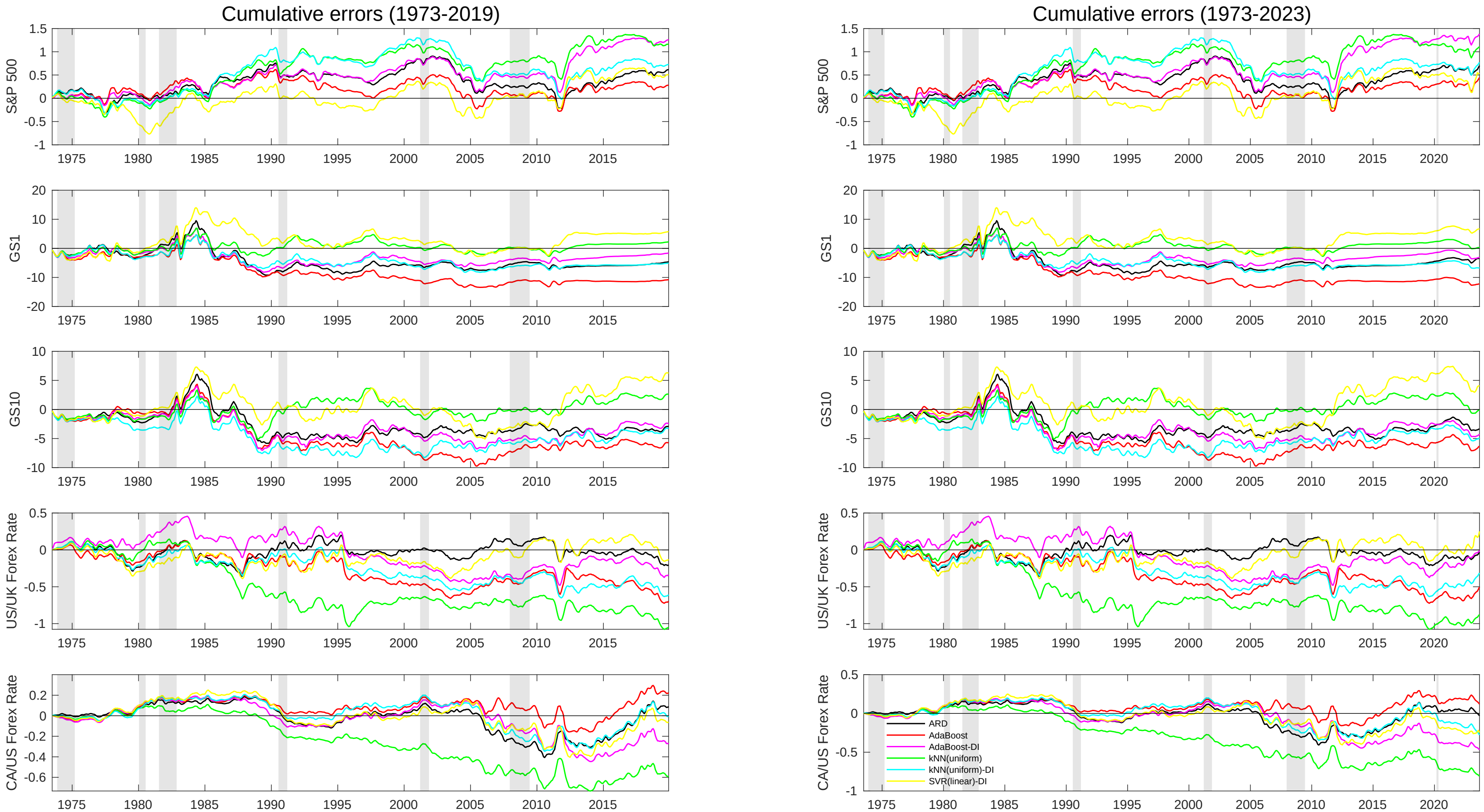
Notes: The figure shows the cumulative errors of the real variables for selected models at $h = 3$. The left panel covers the original sample between 1973-2019, while the right panel includes obersvations till 2023.

Figure 6: **Cumulative errors over time for nominal variables**



Notes: The figure shows the cumulative errors of the nominal variables for selected models at $h = 3$. The left panel covers the original sample between 1973-2019, while the right panel includes observations till 2023.

Figure 7: **Cumulative errors over time for financial variables**



Notes: The figure shows the cumulative errors of the real variables for selected models at $h = 3$. The left panel covers the original sample between 1973-2019, while the right panel includes observations till 2023.

Table 15: Real Personal Income: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.075	0.038	0.027	0.024	0.022	0.105	0.061	0.045	0.039	0.037
Individual machine learning models										
kNN (uniform)	0.95	0.99	0.94*	0.95	0.97	0.96	0.97	0.98*	1.02	1.02
kNN (inverse)	0.96	0.99	0.94*	0.95	0.97	0.96	0.97	0.98**	1.02	1.02
Decision Tree	1.29	1.33	1.21	1.18	1.22	1.25	1.25	0.99	0.92	1.13
SVR (linear)	1.15	1.20	1.19	1.20	1.18	1.07	1.18	0.93	0.80	0.84*
SVR (polynomial)	0.96	1.01	1.09	1.18	1.16	1.07	1.11	0.99	1.05	1.04
SVR (rbf)	0.92**	0.95	0.94*	0.96	0.98	0.91	0.94	0.92**	0.95	0.96
SVR (sigmoid)	0.94*	0.97	0.96	1.04	1.08	0.90	0.88**	0.80**	0.83	0.84*
Ensemble machine learning models										
Random Forest	0.98	1.01	1.00	1.01	1.04	0.99	0.98	0.91	0.90	1.00
XGBoost	1.07	1.06	0.97	0.99	1.00	1.12	1.04	0.94	0.92	0.93
AdaBoost	0.92**	0.94*	0.90***	0.92	0.93	0.92	0.95	0.85**	0.87	0.88***
Gradient Boost	1.18	1.21	1.18	1.16	1.22	1.13	1.08	0.98	0.91	1.11
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	0.96	0.98	0.94	0.92*	0.90*	0.93	0.94	0.92	0.92	0.88
kNN (inverse)-DI	0.96	0.98	0.94	0.92*	0.90*	0.93	0.94	0.92	0.92	0.88
Decision Tree-DI	1.12	1.18	1.17	1.11	1.12	1.09	1.06	0.97	0.86	0.87
SVR (linear)-DI	0.96	1.09	1.16	1.21	1.24	0.96	1.00	1.02	1.05	1.02
SVR (polynomial)-DI	1.46	1.85	1.90	2.09	2.09	1.35	1.48	1.25	1.16	1.00
SVR (rbf)-DI	0.93*	0.99	1.00	1.02	1.05	0.90	0.97	0.97	1.04	1.07
SVR (sigmoid)-DI	1.20	1.21	1.17	1.16	1.19	1.26	1.08	0.97	1.02	1.00
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.13	1.20	1.17	1.11	1.11	1.08	1.06	0.98	0.84	0.88
XGBoost-DI	1.09	1.04	1.00	1.03	1.01	1.14	1.04	0.97	0.96	0.93
AdaBoost-DI	0.98	0.95*	0.95	0.94	0.95	0.86*	0.94	0.91**	0.88	0.88
Gradient Boost-DI	1.14	1.15	1.10	1.08	1.14	1.18	1.07	0.98	0.88	0.88

Table 16: Unemployment Rate: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	2.011	1.33	1.216	1.158	1.126	2.711	2.204	1.999	1.846	1.672
Individual machine learning models										
kNN (uniform)	0.97	0.92	0.89	0.92	0.93	1.03	1.06	1.08	1.09	1.05
kNN (inverse)	0.97	0.92	0.89	0.92	0.93	1.02	1.06	1.08	1.09	1.04
Decision Tree	1.29	1.18	1.24	1.11	1.07	1.15	1.04	1.03	0.91	0.74*
SVR (linear)	1.11	1.05	0.95	0.94	0.98	1.01	0.91	0.83*	0.71***	0.56***
SVR (polynomial)	1.02	1.14	1.18	1.12	1.07	1.15	1.23	1.11	1.10	1.07
SVR (rbf)	0.98	0.95	0.92	0.93	0.94	1.05	1.11	1.03	0.93	0.86**
SVR (sigmoid)	1.00	0.94	0.93	0.91	0.91	0.95	0.89*	0.91	0.78**	0.68***
Ensemble machine learning models										
Random Forest	0.95**	0.96	1.01	0.97	0.95	0.84***	0.92	1.02	0.83**	0.71**
XGBoost	1.01	0.92	0.98	0.92	0.92	0.85**	0.88**	0.99	0.87	0.75**
AdaBoost	0.93***	0.89**	0.92	0.90	0.90	0.90**	0.93	1.02	0.86	0.74**
Gradient Boost	1.25	1.17	1.22	1.07	1.05	1.07	1.02	1.02	0.88	0.77**
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	0.96**	0.87**	0.78**	0.76**	0.75***	0.91*	0.89*	0.84**	0.76**	0.70***
kNN (inverse)-DI	0.96**	0.87**	0.78**	0.76**	0.75***	0.91*	0.89*	0.84**	0.76**	0.70***
Decision Tree-DI	1.24	1.12	1.09	1.05	1.14	1.14	1.01	1.05	0.88	0.87
SVR (linear)-DI	1.18	1.07	0.98	1.02	1.06	1.13	0.98	0.85**	0.80***	0.71***
SVR (polynomial)-DI	1.18	0.99	1.10	1.10	1.16	1.62	1.16	1.29	0.73*	0.81
SVR (rbf)-DI	1.05	1.01	0.97	0.97	0.98	1.09	1.18	1.09	0.99	0.94
SVR (sigmoid)-DI	1.19	1.06	1.07	1.03	1.10	1.31	1.05	1.11	1.03	0.99
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.23	1.11	1.07	1.05	1.14	1.15	0.98	1.05	0.85	0.88
XGBoost-DI	1.08	0.95	0.89	0.96	1.03	1.01	0.78**	0.82	0.73**	0.68**
AdaBoost-DI	0.96**	0.87**	0.83*	0.89	0.94	0.93	0.83**	0.87	0.79**	0.75**
Gradient Boost-DI	1.24	1.14	1.04	1.03	1.10	1.25	1.03	0.99	0.86	0.85

Table 17: Real Personal Consumption Expenditures: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.06	0.03	0.022	0.02	0.019	0.086	0.053	0.04	0.038	0.037
Individual machine learning models										
kNN (uniform)	1.04	1.01	1.05	1.06	1.06	0.99	1.00	1.14	1.13	1.10
kNN (inverse)	1.04	1.01	1.05	1.06	1.05	0.99	1.00	1.15	1.13	1.09
Decision Tree	1.32	1.21	1.17	1.12	1.14	1.12	0.94	0.85	0.89	0.96
SVR (linear)	1.17	1.08	1.13	1.11	1.15	1.13	0.93	0.86	0.77**	0.88
SVR (polynomial)	1.00	1.03	1.14	1.12	1.13	1.02	1.04	1.10	1.11	1.11
SVR (rbf)	0.97**	0.92***	0.95	0.99	1.01	0.92**	0.90*	0.96	1.00	1.03
SVR (sigmoid)	1.05	0.96	0.99	1.01	1.04	1.01	0.81*	0.88*	0.80*	0.86
Ensemble machine learning models										
Random Forest	1.01	0.96	0.96	0.95	0.99	0.92	0.82**	0.81*	0.85	0.95
XGBoost	1.01	1.02	0.95	1.01	1.03	0.93	0.91	0.91	0.96	1.01
AdaBoost	0.98	0.95*	0.93	0.99	1.00	0.92*	0.85***	0.88*	0.91	0.94
Gradient Boost	1.25	1.19	1.14	1.11	1.13	1.06	0.88	0.83	0.88	0.97
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	0.99	0.93**	0.91*	0.89*	0.90	0.93	0.87**	0.90	0.89	0.95
kNN (inverse)-DI	1.00	0.93**	0.92*	0.89*	0.90	0.93	0.86**	0.90	0.90	0.95
Decision Tree-DI	1.26	1.21	1.16	1.21	1.20	1.19	1.08	0.97	1.00	1.10
SVR (linear)-DI	1.11	1.09	1.16	1.23	1.23	1.12	0.94	0.99	0.94	1.00
SVR (polynomial)-DI	1.24	1.23	1.16	1.18	1.25	1.42	1.25	1.02	1.03	1.10
SVR (rbf)-DI	1.02	1.01	1.00	1.01	1.03	0.95	1.02	1.05	1.05	1.06
SVR (sigmoid)-DI	1.08	1.02	1.10	1.05	1.09	0.94	0.87	1.05	0.94	1.03
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.28	1.21	1.16	1.20	1.20	1.22	1.09	0.94	0.98	1.11
XGBoost-DI	1.10	1.03	0.98	1.03	1.06	1.15	1.05	0.97	1.01	1.08
AdaBoost-DI	0.97	0.94**	0.94	0.98	1.01	0.91**	0.90*	0.88	0.91	1.01
Gradient Boost-DI	1.20	1.21	1.15	1.16	1.18	1.09	1.13	0.94	1.00	1.10

Table 18: CPI (All items less food) inflation: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.034	0.03	0.026	0.025	0.024	0.057	0.055	0.043	0.039	0.038
Individual machine learning models										
kNN (uniform)	1.10	1.24	1.39	1.47	1.50	0.93	1.10	1.48	1.19	1.16
kNN (inverse)	1.10	1.24	1.39	1.46	1.49	0.93	1.11	1.47	1.18	1.16
Decision Tree	1.44	1.64	1.65	1.60	1.57	1.16	1.63	1.47	1.25	1.12
SVR (linear)	1.24	1.37	1.54	1.54	1.51	1.07	1.07	1.25	1.19	1.17
SVR (polynomial)	1.06	1.29	1.43	1.44	1.48	1.00	1.27	1.35	1.10	1.12
SVR (rbf)	1.06	1.19	1.33	1.40	1.42	0.95	1.08	1.41	1.15	1.14
SVR (sigmoid)	1.13	1.28	1.42	1.51	1.48	1.05	1.08	1.37	1.17	1.14
Ensemble machine learning models										
Random Forest	1.11	1.27	1.33	1.26	1.30	1.03	1.34	1.36	1.01	1.02
XGBoost	1.14	1.22	1.25	1.28	1.32	1.05	1.18	1.26	0.97	1.04
AdaBoost	1.06	1.19	1.24	1.26	1.29	0.96	1.18	1.33	1.01	0.97
Gradient Boost	1.34	1.50	1.53	1.44	1.50	1.20	1.50	1.49	1.18	1.14
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.08	1.23	1.37	1.40	1.41	0.96	1.13	1.44	1.07	1.01
kNN (inverse)-DI	1.09	1.23	1.37	1.40	1.41	0.96	1.14	1.44	1.07	1.01
Decision Tree-DI	1.27	1.33	1.46	1.58	1.55	0.92	1.18	1.44	1.15	1.16
SVR (linear)-DI	1.35	1.59	1.99	2.06	2.03	1.26	1.48	2.12	1.72	1.54
SVR (polynomial)-DI	1.60	1.73	2.08	2.27	2.37	1.84	1.77	2.39	2.03	1.76
SVR (rbf)-DI	1.11	1.25	1.46	1.56	1.57	0.98	1.11	1.49	1.21	1.18
SVR (sigmoid)-DI	1.24	1.48	1.82	1.95	1.97	1.06	1.44	2.01	1.76	1.45
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.29	1.31	1.46	1.59	1.55	0.95	1.14	1.43	1.18	1.16
XGBoost-DI	1.11	1.14	1.29	1.32	1.34	0.98	1.04	1.25	1.06	0.97
AdaBoost-DI	1.02	1.11	1.19	1.23	1.26	0.89	0.98	1.22	1.00	1.00
Gradient Boost-DI	1.20	1.28	1.36	1.55	1.52	0.99	1.14	1.36	1.22	1.14

Table 19: PCEPI inflation: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.022	0.019	0.017	0.016	0.016	0.036	0.034	0.028	0.027	0.026
Individual machine learning models										
kNN (uniform)	1.13	1.29	1.43	1.46	1.47	1.02	1.15	1.49	1.25	1.25
kNN (inverse)	1.13	1.30	1.44	1.46	1.47	1.03	1.16	1.49	1.25	1.25
Decision Tree	1.46	1.60	1.64	1.61	1.68	1.26	1.50	1.36	1.16	1.34
SVR (linear)	1.26	1.40	1.54	1.52	1.53	1.02	1.07	1.21	1.22	1.28
SVR (polynomial)	1.07	1.30	1.43	1.40	1.42	1.00	1.24	1.32	1.15	1.20
SVR (rbf)	1.07	1.22	1.35	1.38	1.39	0.95	1.08	1.37	1.15	1.19
SVR (sigmoid)	1.12	1.28	1.43	1.45	1.44	0.98	1.06	1.27	1.14	1.18
Ensemble machine learning models										
Random Forest	1.13	1.27	1.35	1.29	1.29	1.08	1.27	1.27	1.01	1.05
XGBoost	1.17	1.23	1.26	1.27	1.32	1.03	1.12	1.22	1.02	1.07
AdaBoost	1.06	1.20	1.24	1.24	1.26	0.96	1.15	1.24	1.02	1.07
Gradient Boost	1.40	1.55	1.55	1.57	1.56	1.13	1.47	1.28	1.12	1.33
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.12	1.28	1.40	1.38	1.38	1.01	1.14	1.39	1.09	1.10
kNN (inverse)-DI	1.12	1.28	1.40	1.38	1.38	1.02	1.15	1.39	1.09	1.10
Decision Tree-DI	1.35	1.34	1.57	1.68	1.59	1.10	1.06	1.43	1.36	1.27
SVR (linear)-DI	1.42	1.62	1.98	1.97	1.91	1.37	1.42	1.95	1.66	1.60
SVR (polynomial)-DI	1.44	1.64	1.95	2.05	2.04	1.35	1.34	1.91	2.11	2.45
SVR (rbf)-DI	1.11	1.26	1.47	1.52	1.53	0.96	1.09	1.43	1.21	1.20
SVR (sigmoid)-DI	1.27	1.53	1.79	1.82	1.79	1.21	1.49	1.80	1.34	1.42
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.36	1.33	1.58	1.66	1.60	1.11	1.06	1.44	1.32	1.27
XGBoost-DI	1.09	1.16	1.29	1.35	1.34	0.98	1.04	1.29	1.15	1.11
AdaBoost-DI	1.04	1.11	1.22	1.24	1.28	0.93	0.97	1.20	1.03	1.11
Gradient Boost-DI	1.26	1.25	1.49	1.60	1.63	1.08	0.93	1.43	1.23	1.33

Table 20: Producer Price Index (Finished Consumer Goods) inflation: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.093	0.07	0.06	0.055	0.053	0.148	0.118	0.092	0.08	0.065
Individual machine learning models										
kNN (uniform)	1.20	1.41	1.60	1.70	1.75	1.01	1.25	1.71	1.66	1.92
kNN (inverse)	1.20	1.41	1.60	1.70	1.75	1.01	1.25	1.71	1.65	1.92
Decision Tree	1.47	1.51	1.66	1.87	1.76	1.21	1.26	1.58	1.45	1.55
SVR (linear)	1.36	1.58	1.78	1.84	1.84	1.17	1.25	1.59	1.71	1.97
SVR (polynomial)	1.15	1.42	1.60	1.64	1.69	1.03	1.32	1.60	1.59	1.82
SVR (rbf)	1.16	1.35	1.53	1.62	1.67	1.01	1.21	1.65	1.59	1.83
SVR (sigmoid)	1.22	1.44	1.65	1.76	1.76	1.04	1.22	1.60	1.64	1.88
Ensemble machine learning models										
Random Forest	1.16	1.26	1.40	1.48	1.45	1.03	1.21	1.47	1.33	1.49
XGBoost	1.21	1.31	1.42	1.48	1.49	1.02	1.23	1.47	1.38	1.60
AdaBoost	1.12	1.23	1.37	1.43	1.46	0.97	1.13	1.51	1.40	1.57
Gradient Boost	1.42	1.43	1.66	1.76	1.78	1.12	1.23	1.61	1.37	1.61
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.19	1.40	1.60	1.65	1.68	1.00	1.26	1.69	1.54	1.71
kNN (inverse)-DI	1.20	1.40	1.60	1.65	1.68	1.01	1.26	1.69	1.54	1.71
Decision Tree-DI	1.41	1.54	1.71	1.78	1.82	1.15	1.30	1.68	1.47	1.65
SVR (linear)-DI	1.53	1.68	2.11	2.20	2.20	1.37	1.48	2.05	1.98	2.31
SVR (polynomial)-DI	2.04	2.39	3.28	3.47	3.22	1.43	1.84	2.77	4.21	5.70
SVR (rbf)-DI	1.22	1.40	1.66	1.78	1.82	0.98	1.19	1.69	1.65	1.92
SVR (sigmoid)-DI	1.45	1.76	2.06	2.22	2.11	1.27	1.66	2.18	2.31	2.14
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.42	1.53	1.73	1.78	1.83	1.14	1.31	1.65	1.46	1.63
XGBoost-DI	1.25	1.35	1.41	1.50	1.50	1.15	1.34	1.40	1.41	1.50
AdaBoost-DI	1.10	1.23	1.34	1.42	1.44	<u>0.95</u>	1.12	1.45	1.35	1.56
Gradient Boost-DI	1.32	1.50	1.58	1.69	1.72	1.23	1.29	1.54	1.48	1.55

Table 21: Producer Price Index (Crude Materials) inflation: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
ARD (RMSPE)	0.457	0.321	0.285	0.268	0.258	0.681	0.552	0.51	0.458	0.362
Individual machine learning models										
kNN (uniform)	1.20	1.44	1.61	1.70	1.76	0.99	1.27	1.53	1.52	1.64
kNN (inverse)	1.20	1.44	1.61	1.71	1.76	0.99	1.27	1.54	1.52	1.63
Decision Tree	1.58	1.67	2.01	2.03	1.91	1.17	1.35	1.90	1.87	1.58
SVR (linear)	1.47	1.75	2.00	2.07	2.05	1.26	1.41	1.49	1.52	1.61
SVR (polynomial)	1.17	1.42	1.64	1.72	1.74	0.96	1.29	1.47	1.49	1.56
SVR (rbf)	1.16	1.39	1.55	1.63	1.68	0.91	1.21	1.50	1.55	1.66
SVR (sigmoid)	1.25	1.53	1.70	1.80	1.83	0.98	1.31	1.49	1.57	1.67
Ensemble machine learning models										
Random Forest	1.22	1.30	1.49	1.57	1.56	0.97	1.18	1.49	1.39	1.42
XGBoost	1.34	1.42	1.56	1.69	1.67	1.19	1.46	1.52	1.43	1.58
AdaBoost	1.14	1.30	1.43	1.48	1.52	0.91	1.22	1.41	1.40	1.45
Gradient Boost	1.56	1.59	1.97	2.06	2.05	1.24	1.35	1.98	2.27	2.48
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.20	1.44	1.60	1.61	1.61	0.91	1.26	1.48	1.46	1.37
kNN (inverse)-DI	1.20	1.44	1.60	1.61	1.61	0.92	1.26	1.48	1.46	1.36
Decision Tree-DI	1.48	1.57	1.80	1.76	1.78	1.46	1.38	1.51	1.50	1.63
SVR (linear)-DI	1.37	1.64	2.00	2.14	2.19	1.19	1.36	1.74	1.82	1.99
SVR (polynomial)-DI	3.24	4.00	4.10	4.48	4.75	1.41	1.91	2.87	6.01	8.17
SVR (rbf)-DI	1.20	1.43	1.65	1.76	1.82	0.88	1.22	1.56	1.63	1.80
SVR (sigmoid)-DI	1.57	1.84	2.11	2.14	2.28	1.18	1.56	1.82	1.91	2.34
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.48	1.59	1.79	1.77	1.80	1.40	1.39	1.50	1.51	1.66
XGBoost-DI	1.60	1.69	1.69	1.83	1.68	1.85	1.71	1.51	1.69	1.62
AdaBoost-DI	1.15	1.27	1.38	1.45	1.48	0.90	1.17	1.40	1.43	1.49
Gradient Boost-DI	1.57	1.61	1.69	1.80	1.78	1.37	1.47	1.47	1.46	1.69

Table 22: 1-Year Treasury Rate: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
RW (RMSPE)	5.347	3.687	2.529	1.944	1.719	10.852	7.337	4.112	2.952	2.553
Individual machine learning models										
kNN (uniform)	0.98	1.00	0.99	1.04	1.06	0.94*	0.93	0.86**	0.96	0.99
kNN (inverse)	0.98	1.00	0.99	1.05	1.06	0.94	0.93	0.87**	0.97	0.99
Decision Tree	1.32	1.65	1.46	1.40	1.62	1.19	1.56	1.50	1.43	1.49
SVR (linear)	1.15	1.20	1.21	1.36	1.60	1.10	1.02	1.00	1.21	1.05
SVR (polynomial)	1.05	1.18	1.32	1.37	1.38	1.05	1.10	0.99	1.02	1.06
SVR (rbf)	1.01	1.02	1.02	1.06	1.13	1.00	1.00	0.98	0.94	0.90
SVR (sigmoid)	1.06	1.09	1.11	1.22	1.37	1.01	0.95	1.02	1.04	0.94
Ensemble machine learning models										
Random Forest	1.08	1.21	1.18	1.26	1.31	1.10	1.15	1.21	1.33	1.20
XGBoost	1.12	1.20	1.26	1.29	1.46	1.20	1.22	1.36	1.40	1.39
AdaBoost	0.98	1.09	1.06	1.20	1.35	1.00	1.01	0.97	1.28	1.26
Gradient Boost	1.33	1.47	1.34	1.37	1.55	1.26	1.30	1.38	1.43	1.43
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.00	1.02	1.01	1.04	1.10	0.98	0.98	0.92***	0.92**	0.95
kNN (inverse)-DI	1.00	1.02	1.01	1.04	1.10	0.98	0.98	0.92***	0.91***	0.95
Decision Tree-DI	1.21	1.33	1.30	1.32	1.41	1.01	1.17	1.18	1.25	1.22
SVR (linear)-DI	1.14	1.23	1.50	1.51	1.68	1.05	1.05	1.42	1.25	1.24
SVR (polynomial)-DI	1.83	2.77	2.72	1.86	1.89	1.74	3.66	3.10	1.33	1.14
SVR (rbf)-DI	1.01	1.11	1.17	1.15	1.20	1.00	1.06	1.09	1.04	1.05
SVR (sigmoid)-DI	1.18	1.25	1.40	1.45	1.47	1.07	1.06	0.97	1.16	1.09
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.22	1.33	1.28	1.33	1.40	1.00	1.17	1.18	1.25	1.23
XGBoost-DI	1.17	1.16	1.12	1.13	1.26	1.18	1.22	1.18	1.15	1.21
AdaBoost-DI	1.02	1.04	1.02	1.08	1.14	0.99	1.01	0.94*	1.03	1.00
Gradient Boost-DI	1.18	1.28	1.24	1.27	1.35	1.14	1.26	1.09	1.25	1.23

Table 23: 10-Year Treasury Rate: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
RW (RMSPE)	3.557	2.389	1.694	1.362	1.206	6.076	3.911	2.187	1.682	1.46
Individual machine learning models										
kNN (uniform)	1.01	1.04	1.07	1.11	1.12	0.97	0.98	0.96*	0.98	1.00
kNN (inverse)	1.01	1.04	1.07	1.11	1.12	0.97	0.98	0.96	0.98	1.00
Decision Tree	1.38	1.42	1.43	1.51	1.63	1.31	1.22	1.26	1.56	1.66
SVR (linear)	1.24	1.27	1.30	1.42	1.54	1.19	1.17	1.21	1.31	1.14
SVR (polynomial)	1.04	1.09	1.10	1.13	1.18	1.07	1.11	1.01	1.02	1.01
SVR (rbf)	1.03	1.05	1.08	1.14	1.18	1.03	1.04	1.06	1.03	0.98
SVR (sigmoid)	1.09	1.10	1.16	1.27	1.34	1.12	1.10	1.11	1.10	0.97
Ensemble machine learning models										
Random Forest	1.05	1.15	1.17	1.27	1.38	1.08	1.12	1.14	1.41	1.35
XGBoost	1.14	1.24	1.21	1.31	1.42	1.25	1.35	1.21	1.48	1.45
AdaBoost	1.00	1.09	1.12	1.24	1.36	1.03	1.09	1.06	1.43	1.26
Gradient Boost	1.33	1.44	1.36	1.50	1.59	1.28	1.43	1.20	1.57	1.65
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.01	1.09	1.10	1.13	1.19	0.98	1.05	1.01	0.99	1.05
kNN (inverse)-DI	1.01	1.09	1.10	1.13	1.19	0.98	1.04	1.01	0.99	1.05
Decision Tree-DI	1.23	1.39	1.40	1.44	1.44	1.11	1.31	1.25	1.16	1.16
SVR (linear)-DI	1.15	1.26	1.42	1.43	1.56	1.00	1.06	1.57	1.23	1.13
SVR (polynomial)-DI	1.33	1.46	1.47	1.21	1.26	1.56	1.69	1.90	1.02	1.08
SVR (rbf)-DI	1.05	1.12	1.17	1.17	1.24	1.01	1.12	1.14	1.12	1.08
SVR (sigmoid)-DI	1.17	1.22	1.26	1.27	1.34	1.08	1.15	1.07	1.09	1.01
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.24	1.42	1.41	1.43	1.46	1.11	1.35	1.29	1.17	1.20
XGBoost-DI	1.11	1.23	1.22	1.24	1.29	1.14	1.28	1.27	1.16	1.20
AdaBoost-DI	1.04	1.08	1.12	1.15	1.24	1.11	1.05	1.05	1.02	1.15
Gradient Boost-DI	1.28	1.33	1.32	1.39	1.41	1.29	1.35	1.26	1.13	1.20

Table 24: US/UK Foreign Exchange Rate: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
RW (RMSPE)	0.281	0.197	0.149	0.122	0.106	0.328	0.257	0.21	0.178	0.153
Individual machine learning models										
kNN (uniform)	1.05	1.06	1.09	1.10	1.09	1.03	1.07	1.09	1.07	1.00
kNN (inverse)	1.05	1.06	1.09	1.10	1.10	1.03	1.07	1.09	1.06	1.00
Decision Tree	1.37	1.43	1.48	1.38	1.35	1.31	1.33	1.20	1.10	1.25
SVR (linear)	1.30	1.34	1.36	1.40	1.35	1.22	1.32	1.23	1.17	0.99
SVR (polynomial)	1.02	1.10	1.12	1.09	1.10	1.07	1.21	1.04	1.11	1.07
SVR (rbf)	1.03	1.06	1.07	1.07	1.05	1.01	1.03	1.05	1.04	0.96
SVR (sigmoid)	1.13	1.14	1.16	1.16	1.16	1.18	1.13	1.12	0.99	0.88
Ensemble machine learning models										
Random Forest	1.04	1.19	1.27	1.22	1.17	0.96	1.23	1.10	1.14	1.11
XGBoost	1.05	1.18	1.27	1.24	1.22	0.94	1.23	1.10	1.06	1.10
AdaBoost	1.01	1.11	1.21	1.22	1.21	0.98	1.14	1.07	1.14	1.14
Gradient Boost	1.34	1.38	1.43	1.34	1.31	1.27	1.31	1.17	1.10	1.23
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.05	1.10	1.11	1.10	1.08	1.00	1.04	0.99	0.96	0.88**
kNN (inverse)-DI	1.05	1.10	1.11	1.10	1.07	0.99	1.05	0.99	0.96	0.87**
Decision Tree-DI	1.29	1.39	1.35	1.40	1.37	1.23	1.34	1.09	1.15	1.09
SVR (linear)-DI	1.17	1.26	1.33	1.31	1.44	1.18	1.35	1.13	0.95	0.93
SVR (polynomial)-DI	1.17	1.23	1.33	1.24	1.17	1.31	1.36	1.32	1.05	1.02
SVR (rbf)-DI	1.06	1.08	1.11	1.12	1.18	1.03	1.03	1.02	1.04	1.04
SVR (sigmoid)-DI	1.22	1.20	1.23	1.27	1.30	1.22	1.15	0.98	0.99	0.95
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.30	1.39	1.37	1.42	1.36	1.22	1.28	1.13	1.18	1.09
XGBoost-DI	1.11	1.18	1.25	1.26	1.23	0.94	1.11	1.09	1.10	1.06
AdaBoost-DI	1.00	1.08	1.15	1.17	1.13	0.93	1.00	1.02	1.06	1.01
Gradient Boost-DI	1.28	1.34	1.37	1.34	1.32	1.10	1.24	1.12	1.08	1.09

Table 25: Canada/US Foreign Exchange Rate: relative RMSPE (sample period: 1960m1-2019m12)

	Pre-Pandemic Sample					**NBER recession periods**				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model										
RW (RMSPE)	0.173	0.117	0.088	0.073	0.064	0.228	0.161	0.118	0.095	0.082
Individual machine learning models										
kNN (uniform)	1.03	1.04	1.09	1.10	1.10	1.02	1.06	1.16	1.20	1.10
kNN (inverse)	1.03	1.04	1.09	1.10	1.10	1.02	1.06	1.15	1.19	1.11
Decision Tree	1.38	1.42	1.43	1.37	1.35	1.28	1.51	1.13	1.00	0.97
SVR (linear)	1.26	1.37	1.53	1.51	1.42	1.38	1.62	1.57	1.61	1.50
SVR (polynomial)	1.12	1.32	1.32	1.14	1.17	1.43	1.57	1.02	1.06	1.03
SVR (rbf)	1.02	1.05	1.09	1.10	1.11	1.03	1.05	1.07	1.09	1.05
SVR (sigmoid)	1.17	1.16	1.22	1.28	1.26	1.26	1.17	1.05	1.12	1.12
Ensemble machine learning models										
Random Forest	1.07	1.20	1.26	1.22	1.25	1.13	1.37	1.13	0.94	0.94
XGBoost	1.07	1.23	1.24	1.24	1.22	1.12	1.39	1.14	1.13	0.94
AdaBoost	1.01	1.14	1.23	1.21	1.29	0.98	1.31	1.10	1.02	0.98
Gradient Boost	1.28	1.41	1.37	1.34	1.37	1.22	1.55	1.12	0.99	0.97
Individual machine learning models using dimension reduction										
kNN (uniform)-DI	1.03	1.11	1.15	1.11	1.09	1.07	1.14	1.07	1.09	1.00
kNN (inverse)-DI	1.03	1.11	1.16	1.11	1.09	1.07	1.14	1.07	1.08	1.00
Decision Tree-DI	1.39	1.39	1.43	1.40	1.39	1.19	1.33	1.31	1.49	1.20
SVR (linear)-DI	1.16	1.34	1.52	1.51	1.47	1.08	1.46	1.24	1.21	1.19
SVR (polynomial)-DI	1.27	1.49	1.90	1.58	1.47	1.62	2.03	1.70	1.45	1.44
SVR (rbf)-DI	1.07	1.14	1.15	1.17	1.19	1.03	1.08	1.13	1.13	1.06
SVR (sigmoid)-DI	1.16	1.36	1.37	1.42	1.42	1.17	1.44	1.25	1.08	0.97
Ensemble machine learning models using dimension reduction										
Random Forest-DI	1.37	1.41	1.46	1.41	1.38	1.12	1.33	1.32	1.48	1.21
XGBoost-DI	1.17	1.26	1.35	1.22	1.20	1.15	1.33	1.04	1.08	0.94
AdaBoost-DI	1.00	1.11	1.22	1.19	1.17	0.93	1.16	1.11	1.13	1.00
Gradient Boost-DI	1.31	1.40	1.45	1.35	1.35	1.21	1.41	1.28	1.39	1.27

Table 26: Industrial Production growth: RMSPE relative to ARDI (sample period: 1960m1-2019m12)

	Full out-of-sample				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model					
ARDI (RMSPE)	0.089	0.065	0.058	0.057	0.061
Individual machine learning models					
kNN (uniform)	0.92**	0.93	0.89	0.83	0.72
kNN (inverse)	0.91**	0.93	0.89	0.83	0.72
Decision Tree	1.23	1.17	1.15	1.03	0.86
SVR (linear)	1.07	1.07	1.02	0.96	0.86
SVR (polynomial)	0.95	1.06	1.16	1.05	0.86
SVR (rbf)	0.93	0.93	0.90	0.86	0.77
SVR (sigmoid)	0.97	0.96	0.96	0.91	0.83
Ensemble machine learning models					
Random Forest	0.90***	0.96	1.00	0.94	0.77
XGBoost	0.94*	0.93	0.96	0.93	0.78
AdaBoost	0.87***	0.88**	0.97	0.91	0.75
Gradient Boost	1.16	1.14	1.16	1.01	0.85
Individual machine learning models using dimension reduction					
kNN (uniform)-DI	0.94	0.90	0.85	0.82	0.73
kNN (inverse)-DI	0.94	0.90	0.85	0.82	0.73
Decision Tree-DI	1.13	1.15	1.17	1.11	0.93
SVR (linear)-DI	1.04	1.30	1.29	1.12	1.04
SVR (polynomial)-DI	1.77	1.40	1.04	1.04	0.95
SVR (rbf)-DI	0.98	1.03	0.98	0.89	0.78
SVR (sigmoid)-DI	1.17	1.18	1.16	1.23	1.05
Ensemble machine learning models using dimension reduction					
Random Forest-DI	1.12	1.16	1.17	1.12	0.93
XGBoost-DI	1.05	0.94	0.96	0.94	0.84
AdaBoost-DI	0.89***	0.90**	0.96	0.92	0.81
Gradient Boost-DI	1.13	1.13	1.16	1.09	0.91

Table 27: Employment: RMSPE relative to ARDI (sample period: 1960m1-2019m12)

	Full out-of-sample				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model					
ARDI (RMSPE)	0.019	0.015	0.015	0.016	0.017
Individual machine learning models					
kNN (uniform)	0.98	1.11	1.08	1.09	1.01
kNN (inverse)	0.98	1.11	1.08	1.09	1.01
Decision Tree	1.13	1.20	1.30	1.21	1.12
SVR (linear)	1.02	1.08	1.04	1.03	0.98
SVR (polynomial)	1.11	1.46	1.45	1.38	1.23
SVR (rbf)	0.98	1.10	1.04	1.05	0.98
SVR (sigmoid)	0.97	1.06	1.06	1.02	0.97
Ensemble machine learning models					
Random Forest	0.83***	0.97	1.05	1.02	0.97
XGBoost	0.86***	0.91	0.96	0.97	0.94
AdaBoost	0.84***	0.91	1.03	1.06	0.98
Gradient Boost	1.10	1.12	1.26	1.19	1.10
Individual machine learning models using dimension reduction					
kNN (uniform)-DI	0.90**	0.90	0.84	0.88	0.85
kNN (inverse)-DI	0.90***	0.89	0.85	0.88	0.85
Decision Tree-DI	1.11	1.12	1.19	1.16	1.11
SVR (linear)-DI	0.90***	0.85**	0.88	0.95	0.97
SVR (polynomial)-DI	2.78	4.02	4.06	3.74	3.18
SVR (rbf)-DI	0.98	1.11	1.05	1.06	1.00
SVR (sigmoid)-DI	3.30	3.48	2.81	2.57	2.23
Ensemble machine learning models using dimension reduction					
Random Forest-DI	1.12	1.13	1.18	1.17	1.11
XGBoost-DI	0.89***	0.89	0.90	0.95	0.91
AdaBoost-DI	0.85***	0.88*	0.91	0.95	0.92
Gradient Boost-DI	1.06	1.09	1.16	1.16	1.10

Table 28: Real Personal Income: RMSPE relative to ARDI (sample period: 1960m1-2019m12)

	Full out-of-sample				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model					
ARDI (RMSPE)	0.091	0.044	0.029	0.027	0.028
Individual machine learning models					
kNN (uniform)	0.80***	0.89*	0.94	0.91	0.82
kNN (inverse)	0.80***	0.89*	0.94	0.91	0.82
Decision Tree	1.03	1.18	1.14	1.03	0.98
SVR (linear)	0.95	1.05	1.09	1.04	0.93
SVR (polynomial)	0.79***	0.88**	1.00	1.03	0.92
SVR (rbf)	0.76***	0.82***	0.86***	0.84	0.78
SVR (sigmoid)	0.77***	0.84***	0.89**	0.90	0.85
Ensemble machine learning models					
Random Forest	0.82***	0.88**	0.92	0.88	0.83
XGBoost	0.89*	0.92	0.90*	0.86	0.79
AdaBoost	0.76***	0.82***	0.82***	0.80*	0.73
Gradient Boost	0.98	1.05	1.10	1.01	0.97
Individual machine learning models using dimension reduction					
kNN (uniform)-DI	0.78***	0.86**	0.89	0.87	0.82
kNN (inverse)-DI	0.78***	0.86**	0.89	0.87	0.82
Decision Tree-DI	0.93	1.04	1.08	0.97	0.89
SVR (linear)-DI	0.80***	0.94	1.06	1.05	0.98
SVR (polynomial)-DI	1.21	1.60	1.75	1.82	1.65
SVR (rbf)-DI	0.77***	0.85***	0.92	0.89	0.83
SVR (sigmoid)-DI	0.99	1.05	1.08	1.01	0.94
Ensemble machine learning models using dimension reduction					
Random Forest-DI	0.94	1.04	1.08	0.96	0.88
XGBoost-DI	0.90*	0.90*	0.92	0.90	0.80
AdaBoost-DI	0.82***	0.83***	0.88**	0.82	0.75
Gradient Boost-DI	0.95	0.98	1.02	0.94	0.90

Table 29: Unemployment Rate: RMSPE relative to ARDI (sample period: 1960m1-2019m12)

	Full out-of-sample				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model					
ARDI (RMSPE)	2.154	1.311	1.143	1.118	1.161
Individual machine learning models					
kNN (uniform)	0.91***	0.96	0.94	0.90	0.84
kNN (inverse)	0.91***	0.96	0.94	0.90	0.84
Decision Tree	1.20	1.22	1.28	1.14	1.07
SVR (linear)	1.03	1.07	1.01	0.97	0.94
SVR (polynomial)	0.95	1.16	1.25	1.15	1.03
SVR (rbf)	0.91**	0.96	0.97	0.96	0.90
SVR (sigmoid)	0.93*	0.95	0.98	0.94	0.87
Ensemble machine learning models					
Random Forest	0.89***	0.97	1.08	1.01	0.92
XGBoost	0.94*	0.94	1.04	0.95	0.90
AdaBoost	0.86***	0.90	0.97	0.94	0.86
Gradient Boost	1.16	1.16	1.29	1.10	1.04
Individual machine learning models using dimension reduction					
kNN (uniform)-DI	0.95	0.91	0.89	0.89	0.85
kNN (inverse)-DI	0.95	0.91	0.89	0.89	0.85
Decision Tree-DI	1.15	1.15	1.15	1.08	1.09
SVR (linear)-DI	1.10	1.08	1.04	1.06	1.02
SVR (polynomial)-DI	1.10	1.00	1.17	1.13	1.12
SVR (rbf)-DI	0.98	1.03	1.03	1.00	0.94
SVR (sigmoid)-DI	1.11	1.07	1.14	1.06	1.06
Ensemble machine learning models using dimension reduction					
Random Forest-DI	1.15	1.12	1.14	1.08	1.09
XGBoost-DI	1.00	0.96	0.95	0.99	0.99
AdaBoost-DI	0.90***	0.89**	0.88	0.90**	0.91
Gradient Boost-DI	1.16	1.14	1.10	1.06	1.06

Table 30: Real PCE: RMSPE relative to ARDI (sample period: 1960m1-2019m12)

	Full out-of-sample				
Model	**h=1**	**h=3**	**h=6**	**h=9**	**h=12**
Baseline Model					
ARDI (RMSPE)	0.07	0.034	0.025	0.026	0.025
Individual machine learning models					
kNN (uniform)	0.85***	0.86*	0.88	0.78	0.80
kNN (inverse)	0.85***	0.86*	0.88	0.78	0.80
Decision Tree	1.14	1.09	1.00	0.86	0.88
SVR (linear)	1.01	0.97	0.98	0.85	0.88
SVR (polynomial)	0.86***	0.92	0.98	0.86	0.86
SVR (rbf)	0.84***	0.82**	0.82	0.76	0.77
SVR (sigmoid)	0.91***	0.86**	0.85	0.78	0.79
Ensemble machine learning models					
Random Forest	0.87***	0.86**	0.83	0.74	0.76
XGBoost	0.88***	0.91	0.82	0.78	0.79
AdaBoost	0.84***	0.84**	0.80*	0.76	0.76
Gradient Boost	1.07	1.06	0.98	0.85	0.87
Individual machine learning models using dimension reduction					
kNN (uniform)-DI	0.89***	0.88	0.85	0.78	0.79
kNN (inverse)-DI	0.89***	0.88	0.85	0.78	0.79
Decision Tree-DI	1.09	1.07	1.00	0.92	0.91
SVR (linear)-DI	0.96	0.98	1.00	0.94	0.94
SVR (polynomial)-DI	1.07	1.10	1.00	0.91	0.95
SVR (rbf)-DI	0.89***	0.91	0.86	0.77	0.78
SVR (sigmoid)-DI	0.94*	0.91	0.95	0.81	0.83
Ensemble machine learning models using dimension reduction					
Random Forest-DI	1.11	1.08	1.00	0.92	0.92
XGBoost-DI	0.95	0.92	0.85	0.79	0.81
AdaBoost-DI	0.84***	0.84**	0.82	0.76	0.77
Gradient Boost-DI	1.03	1.07	1.00	0.89	0.90

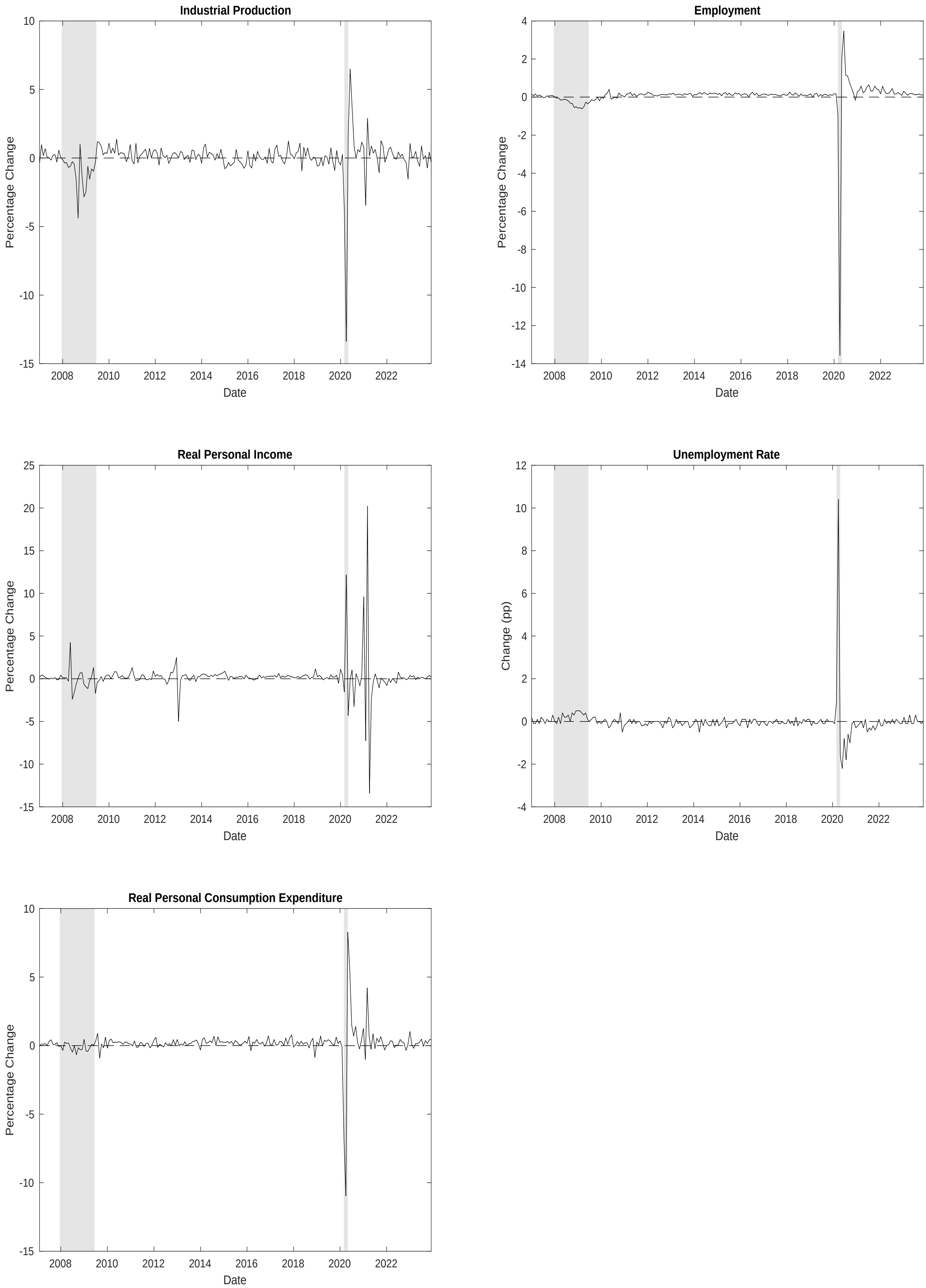


Figure 8: Real variables extreme observations