An Oifig Buiséid Pharlaiminteach
Parliamentary Budget Office

Tithe an Oireachtais
Houses of the Oireachtas

Working Paper Series

No. 2 of 2024

Medium-Term Macroeconomic Forecasting in Ireland: A VAR Setup with Bayesian and Tree Ensemble Models and Forecast Averaging

Akisato Suzuki

February 2024

Disclaimer

PBO Working Papers present primary research in progress intended to elicit comments and encourage debate. The content is subject to review and revision. The analysis and views contained in this paper are those of the author only, and are not necessarily reflective of the position of the PBO or of the Houses of the Oireachtas generally.

The Information is general in nature. Forward-looking statements involve uncertainties and matters may develop significantly from the Information. The Information does not provide a definitive statement in relation to any specific issue or personal circumstance. It does not constitute advice. You must satisfy yourself as to the suitability and any reliability of the Information that we, The Parliamentary Budget Office and Oireachtas Commission (including its servants, agents and contractors), provide.

We accept no responsibility for, and give no guarantees, undertakings or warranties concerning, the accuracy or suitability or otherwise, of the Information; or that our webpages or the Information or other content will be error free, virus free, or infringement free. We accept no liability (including in respect of intellectual property claims) arising out of any third-party content or any third-party website to which we link or refer. To the fullest extent permitted by applicable law, no liability whatsoever is accepted (i) for any reliance placed on the Information or content on our webpages or (ii) for loss or damage arising out of or in connection with your use of this webpages. See our comprehensive Disclaimer Notice here. In any conflict between this Disclaimer and our comprehensive Disclaimer Notice, the latter will prevail.

Medium-Term Macroeconomic Forecasting in Ireland: A VAR Setup with Bayesian and Tree Ensemble Models and Forecast Averaging

Akisato Suzuki*

Parliamentary Budget Office, Ireland

26 February 2024

Abstract

This paper employs a vector autoregression (VAR) setup to develop self-contained multivariate models for 5-year macroeconomic forecasting in Ireland. Irish macroeconomic variables are often volatile and challenging for predictive models. The paper applies the VAR setup to Bayesian modeling and tree ensemble machine learning (Random Forest, XGBoost, and LightGBM). It compares the predictive performance of these models, the two baseline univariate models (random walk and auto-ARIMA), and a model averaging the forecasts from these individual models, by out-of-sample prediction exercises covering the periods between 2015 and 2022. It finds that (1) different models perform better or worse than others, depending on variables and time periods to be forecasted; (2) the averages of the forecasts from the individual models perform in a stable and consistent manner and seem to be a credible choice for actual forecasting. The paper contributes to the literature on macroeconomic forecasting for Ireland and, more generally, small open economies.

JEL Classification Codes: C520, C530

Keywords: forecast, Bayesian VAR, Random Forest, XGBoost, LightGBM, Ireland, small open economy

---

* Akisato Suzuki, PhD, is an Irish Government Economic and Evaluation Service (IGEES) Policy Analyst / Economist at the Parliamentary Budget Office, the Houses of the Oireachtas. This research was partly supported by library resources at University College Dublin, where the author is affiliated as an Adjunct Research Fellow with the School of Politics and International Relations. The author would like to thank his colleagues at the PBO, Niall Conroy (Fiscal Council in Ireland), Luke Rehill (Department of Finance in Ireland), and Denis Gorea, for their helpful comments on earlier versions of this paper; the analysis and views contained in this paper are those of the author only unless stated otherwise. Email: akisato.suzuki@oireachtas.ie. Address: the Houses of the Oireachtas, Leinster House, Kildare Street, Dublin 2, D02 XR20, Ireland.

1 Introduction

Sound fiscal policy requires multiyear planning. Such planning in turn necessitates an understanding of how much revenue the development of the macroeconomic environment will generate and how much expenditure it will require in coming years. Although the future is always uncertain, an understanding of macroeconomic developments can be facilitated by statistical modeling combined with theoretical knowledge of how the economy works. Indeed, macroeconomic forecasting is an integral part of policy and fiscal analysis, as done by governments, independent fiscal institutes, and international organizations.

This paper utilizes a vector autoregression (VAR) setup to develop self-contained multivariate models for macroeconomic forecasting in a medium term, i.e., over a multiyear horizon. The novel aspect of this paper is to evaluate and compare the medium-term predictive performance of conventional econometric models, machine learning algorithms, and the forecast averaging of the point forecasts of these individual models, under the VAR setup. In particular, it does so against macroeconomic variables in the small open economy of Ireland; Irish macroeconomic variables are known to be often volatile and, therefore, challenging for forecasting models.

The specific models used for the VAR setup are hierarchical Bayesian VAR (Giannone, Lenza, and Primiceri 2015; Kuschnig and Vashold 2021), Random Forest (Breiman 2001), XGBoost (Chen and Guestrin 2016), and LightGBM (Ke et al. 2017). Bayesian VAR is a well-established econometric model for a VAR setup; it has a good track record of out-of-sample forecasting accuracy (Bańbura, Giannone, and Reichlin 2010; Giannone, Lenza, and Primiceri 2015; Gupta and Kabundi 2010; Koop 2013). It has previously been applied in Irish contexts (Carroll 2020; Kenny, Meyler, and Quinn 1998) and in other small open economies (Gupta and Kabundi 2010; Hou, Nguyen, and Zhang 2023).

Random Forest, XGBoost, and LightGBM are all tree-based ensemble machine learning

models. They are nonparametric and enable flexible (i.e., nonlinear) functional forms to be identified by data. Random Forest has probably been one of the most common machine learning algorithms used in macroeconomic forecasting previously (e.g., see Biau and D'Elia 2012; Chu and Qureshi 2023; Qureshi, Chu, and Demers 2021; Yoon 2021). XGBoost and LightGBM, specific implementations of the general gradient boosting algorithm (Friedman 2001), are known in the machine learning community as two of the currently best-performing machine learning algorithms ("XGBoost" 2023; Saha 2023) and have been used in macroeconomic forecasting recently (Qureshi, Chu, and Demers 2021). Previous research has suggested that machine learning models can produce more accurate predictions than benchmark forecasts under certain conditions (Chu and Qureshi 2023; Qureshi, Chu, and Demers 2021; Yoon 2021), including those applied to small open economies (Botha et al. 2023; Marcellino and Sivec 2021).

As the baseline models for comparison with those more advanced and complex models, this paper uses the following two univariate models: the random walk forecasts (Hewamalage, Ackermann, and Bergmeir 2023) and the data-driven auto-ARIMA algorithm (Hyndman and Khandakar 2008). As Hewamalage, Ackermann, and Bergmeir (2023) point out, model complexity does not guarantee better predictive performance; a simple model can outperform a complex model.

Finally, the paper also evaluates the predictive performance of forecast averaging, i.e., taking the averages of the point forecasts from those individual models used in this paper. The forecasting literature has found that simple forecast averaging can perform remarkably well (Genre et al. 2013; Makridakis and Winkler 1983; Stock and Watson 2004). The paper also presents a simulation method to generate the prediction intervals of forecast averages.

The models are applied to forecast a suite of Irish macroeconomic variables over a 5-year horizon, the maximum time interval between general elections. Ireland is a challenging case for macroeconomic forecasting models, especially medium-term ones. Its macroeconomic variables have often exhibited strong volatility over the last two to

three decades. For example, in the period of 2003–2022, the standard deviation of Ireland's annual real GDP growth was 6.6% while that of the US's was only 1.9%.[1] It is a valuable exercise to see how well forecasting models known to be good, such as those employed in this paper, perform.

The predictive performance of the models is compared, based on out-of-sample predictions covering the periods between 2015 and 2022. The main focus is the comparison between the following two, most contrasting 5-year periods. First, 2015–2019 is a steady growth period including an unusual hike in GDP in 2015[2] and just before the COVID-19 pandemic hit Ireland. Second, 2018–2022 is the latest period available at the time of writing and including an exceptionally volatile period. These two periods have different macroeconomic trends and, therefore, using both for predictive performance evaluation is a hard test. The results from additional analysis using the periods of 2016–2020 and 2017–2021 are also presented in Appendix C.

The results show that none of the models outperforms one another in all variables across the different time periods. Different models seem to perform better or worse, depending on different contexts. Yet, it is also found that the averages of the forecasts from the individual models used perform in a stable and consistent manner; the forecast average almost always performs nearly as well as, and occasionally better than, the best performing individual model per variable forecasted. This finding adds further evidence, from a small open economy perspective, to the literature that has found the utility of simple forecast averaging (Genre et al. 2013; Makridakis and Winkler 1983; Stock and Watson 2004). The true values are never known in actual forecasting and, therefore, it is impossible to select the best model out of all candidate models *ex ante*. Therefore, forecast averaging seems to be a credible choice in practice. These findings suggest that there is value in developing different modeling approaches such as what this paper does and not just using simple univariate models.

[1] These numbers were calculated using data from the Central Statistics Office (2023h) and the Bureau of Economic Analysis (2023).

[2] One major reason was that large multinational corporations relocated their intellectual property to Ireland (OECD 2016).

Two caveats are worth mentioning in advance. First, forecasting does not show what will happen in future for sure. There is always uncertainty about the future, even if to different degrees, depending on what is forecasted. Instead, a forecasting model projects the value of an outcome variable, based on a mathematical deduction or algorithm given the data and assumptions used. Therefore, a forecasting model offers formal methods to evaluate the plausibility (if not trueness) of the prediction. One such method is out-of-sample prediction, as done in this paper.

The second caveat is that the models presented in this paper are not designed to forecast rare but extreme events such as financial crises. For such events, a different modeling approach is necessary. Instead, the models of this paper are designed to forecast regular trends over a medium term. Nonetheless, the paper also presents the prediction intervals for the averages of the point forecasts from the individual models used, which are informative to signal the uncertainty of forecasting and the possibility of extreme events.

The rest of the paper is structured as follows. First, it explains what a VAR setup means. Second, the basics of each model used are explicated. Third, the data employed are introduced. Fourth, the forecast and evaluation methods are presented. Fifth, the results of the predictive performance of each model are discussed. The final section is a conclusion. All computation was done in the statistical programming software, R (R Core Team 2023).

2 VAR Setup

A VAR setup in this paper means that every variable in a data frame is predicted by the lagged values of all variables. The term "setup" is used here to mean that no particular model is *ex ante* attached to it. Assume there are four variables, y_1, y_2, y_3, and y_4. If one uses only one-lagged values of these variables as regressors, then the setup in general terms is:

$$
\begin{aligned}
y_{1,t} &\sim f_1(\theta_1, y_{1,t-1}, y_{2,t-1}, y_{3,t-1}, y_{4,t-1}), \\
y_{2,t} &\sim f_2(\theta_2, y_{1,t-1}, y_{2,t-1}, y_{3,t-1}, y_{4,t-1}), \\
y_{3,t} &\sim f_3(\theta_3, y_{1,t-1}, y_{2,t-1}, y_{3,t-1}, y_{4,t-1}), \\
y_{4,t} &\sim f_4(\theta_4, y_{1,t-1}, y_{2,t-1}, y_{3,t-1}, y_{4,t-1}), \\
\theta_i &\sim g(\Theta),
\end{aligned}
\tag{1}
$$

where t is a time index; $f_i(\cdot)$ is some function defined by the same data-analyzing model, which might be *ex ante* determined parametrically (e.g., via a linear regression model) or *ex post* identified from the data (e.g., via a machine learning algorithm); θ_i is a vector of the parameters for the function; the presence of the index i indicates that the functional forms and parameters may be different across the outcome variables. $\theta_i \sim g(\Theta)$ means that θ_i may be a function of a set of hyperparameters Θ so that θ_i can be correlated with $\theta_j, i \neq j$; if $\theta_i \perp\!\!\!\perp \theta_j \forall i, j$, this part becomes redundant. While the standard VAR model assumes linear functions for all variables, such a parametric assumption is not imposed here. Instead, it is only assumed that the same data-analyzing model is applied to each outcome variable in the VAR system. This looser definition accommodates a nonparametric model that identifies a (most probably nonlinear) function *ex post* based on the data analyzed, such as tree-based machine learning algorithms.

A VAR setup makes a forecasting model multivariate and self-contained. In other words, there is no technical limitation on the number of horizons over which all variables can be forecasted without an external model. This is because the VAR system means that each variable is predicted by the lagged values of all variables. This is not the case for models where the outcome and predictors are clearly separated (e.g., a single ordinary least square model). In such a case, the number of forecast horizons is either limited by the number of the available lags of predictors, or possible to increase only if the predictors are forecasted by an external model.

3 Models

3.1 Bayesian VAR

Unlike the standard VAR, Bayesian VAR is less data-demanding and allows for stable estimation of many parameters through shrinkage (Bańbura, Giannone, and Reichlin 2010; Giannone, Lenza, and Primiceri 2015; Koop 2013; Kuschnig and Vashold 2021). In particular, hierarchical Bayesian VAR relaxes the assumptions on parameters for priors, by drawing prior parameter values from a distribution, rather than using a scalar value (Giannone, Lenza, and Primiceri 2015; Kuschnig and Vashold 2021).[3]

This paper uses a modified version of the Minnesota prior setup (Bańbura, Giannone, and Reichlin 2010; Koop 2013; Litterman 1986), implementing the Normal-inverse-Wishart natural conjugate priors with hierarchical modeling (Giannone, Lenza, and Primiceri 2015; Kuschnig and Vashold 2021).[4] The conjugacy allows the marginal likelihood of the model to be available in the closed form (Giannone, Lenza, and Primiceri 2015, 437). This is useful to compare the predictive performance of different model specifications via the posterior probability of each model.

Hierarchical Bayesian VAR can be expressed as follows:

[3]The meaning of "hierarchical" in this literature is different from what the Bayesian statistics literature usually means, which is a modeling approach to draw parameter values for the multiple priors of the same type (e.g., the mean parameters of normal distributions for varying regression intercepts in panel data) from another prior distribution (the so-called "hyperprior"); see Gelman et al. (2013).

[4]The Minnesota prior is not what the Bayesian statistics literature usually means by "prior," which refers to the prior distribution of a parameter. Instead, the Minnesota prior is a particular parameterization of shrinkage for the regression coefficients in a VAR model.

$$
\begin{aligned}
\underset{NT\times 1}{y} &= \underset{NT\times NK}{(I_N \otimes X)}\ \underset{NK\times 1}{\beta} + \underset{NT\times 1}{u}, \\
u &\sim N(0, \underset{N\times N}{\Sigma} \otimes \underset{T\times T}{I_T}), \\
\beta|\Sigma &\sim N(\underset{NK\times 1}{\mu}, \underset{N\times N}{\Sigma} \otimes \underset{K\times K}{\Omega}), \\
\Sigma &\sim IW(\underset{N\times N}{\Psi}, \nu), \\
\Omega &= \begin{cases} \omega_{\text{intercept}} \text{ if } k = 1, \\ \lambda^2 \frac{1}{p^\alpha} \frac{1}{\psi_{n|p}/(\nu-N-1)} \text{ otherwise,} \end{cases} \\
\lambda &\sim \text{Gamma(mode, sd, min, max)}, \\
\alpha &\sim \text{Gamma(mode, sd, min, max)}, \\
\psi &\sim \text{Inv-Gamma(shape, scale, mode, min, max)},
\end{aligned}
$$

where N is the number of outcome variables and $n = 1, ..., N$; T is the number of time points; P is the number of lags and $p = 1, ..., P$; $K = NP + 1$ is the number of regressors (+1 for the intercept) and $k = 1, ..., K$, where $k = 1$ for the intercept. y is a vectorized version of VAR outcome variables; X is a T times K matrix of regressors; I_N and I_T are identity matrices of the lengths N and T respectively; β is the regression coefficients; and u is the error terms. μ and $\Sigma \otimes \Omega$ are the prior mean and variance-covariance matrix of a matricvariate normal distribution for the regression coefficients.[5] Ψ (which is a diagonal matrix) and ν are the prior scale matrix and degree of freedom parameters of an inverse-Wishart distribution for Σ; ν is fixed such that $\nu - N - 1 = 1$. Ω is a diagonal matrix, whose first diagonal element, $\omega_{\text{intercept}}$, is the prior variance for the constant and the remaining diagonal elements contain the Minnesota prior values calculated per lag order p (as indicated by $n|p$, i.e., $n = 1, ..., N$ for every p). λ controls the overall tightness of the prior on the variance-covariance matrix; α affects the level of shrinkage for greater lags; and ψ_n is the nth diagonal element of Ψ and controls the degree of shrinkage on the lags of regressors other than those of the outcome vari-

[5] A matricvariate normal distribution is a multivariate normal distribution where the variance-covariance matrix is the Kronecker product of two variance-covariance matrices (Gamerman and Moreira 2002, 70).

able (the expected value of Σ_{nn} is $\frac{\psi_n}{\nu-N-1}$ and $\nu - N - 1 = 1$ by setting as above). In Giannone, Lenza, and Primiceri (2015) and Kuschnig and Vashold (2021), λ and α are drawn from Gamma distributions, while ψ is drawn from an inverse Gamma distribution. In this paper, hierarchical Bayesian VAR is implemented by the R `BVAR` package (Kuschnig and Vashold 2021).

3.2 Tree Ensemble Models

Tree ensemble models combine results from multiple decision trees. The decision tree algorithm sorts ("splits" in the decision tree terminology) the observations in data into different bins or "nodes," by the specific threshold value of a predictor that minimizes the errors in the resulting nodes (for a brief summary, see Hastie, Tibshirani, and Friedman 2017, sec. 9.2; James et al. 2021, sec. 8.1). When the algorithm stops, the deepest level of nodes is called terminal nodes or terminal regions. Given its algorithm, the number of splitting in a decision tree model affects the number of observations per node (aka "node size"). With more predictor values and splitting, the observations per node are expected to become more homogeneous in terms of their outcome values. One problem in this algorithm is overfitting. With a sufficient number of predictor values and splitting, it is possible to make all nodes completely homogeneous in terms of their outcome values (an extreme example is one node per observation). Since there is no guarantee that new data will be identical to data used ("training" data in the machine learning terminology), overfitting usually results in worse predictive performance.

Tree-based ensemble models overcome this overfitting problem. Ensemble methods combine results from several "weak learners," i.e., models that by themselves may underfit training data. There are many ways to make individual decision trees weak learners in a tree-based ensemble model. One way is to limit the depth of each tree. Another way is to use a subsample randomly selected from training data. Covering all ways to control the data fitting of a decision tree is beyond the scope of this paper. More generally, the descriptions of tree ensemble algorithms here are only in basic forms; the actual implementation usually involves more complexity and some detail may devi-

ate from these descriptions. Interested readers are encouraged to refer to the documentations of specific algorithms (e.g., Liaw 2022; Microsoft Corporation 2023; XGBoost Developers 2022).

One ensemble method is averaging, as done in Random Forest for regression tasks. Random Forest runs several trees separately using bootstrap samples from training data, and predictors randomly selected at each node, until a specified node size is reached; the predictions from all trees are combined to produce the final prediction (Breiman 2001; Hastie, Tibshirani, and Friedman 2017, chap. 15; James et al. 2021, sec. 8.2.2; Liaw and Wiener 2002). The random forest algorithm for regression proceeds as follows (adapted from Hastie, Tibshirani, and Friedman 2017, 588):

1. Set B, the number of trees to grow, and for every $b = 1, ..., B$, do (a) and (b).

   (a) Generate y_b and x_b, a bootstrap sample of a preset size, from the training data, y (the outcome variable) and x (the predictors).

   (b) Fit a regression tree $f_b(x_b)$, where only a preset number of predictors in x_b will be randomly selected at every node j of the tree (denote the selected predictors as x_b^j), and the tree is grown by splitting every j into two child nodes by the best predictor among x_b^j at j, until the preset minimum node size is reached.

2. The final output is:

$$\hat{y} = \frac{1}{B} \sum_{b=1}^{B} f_b(x_b).$$

Random Forest has been widely used across different disciplines for forecasting tasks including macroeconomic forecasts (Biau and D'Elia 2012; Chu and Qureshi 2023; Qureshi, Chu, and Demers 2021; Yoon 2021). Random Forest is simple to use, as there are only two main parameters – `ntree`, the number of trees to grow, and `mtry`, the number of predictors randomly selected at each node (Liaw and Wiener 2002, 18). Therefore, it is a good baseline machine learning model. This paper implements Random Forest by the

R `randomForest` package (Liaw and Wiener 2002). It was found that the default values for its parameters work well.[6]

Another ensemble method is sequential learning, as done in gradient tree boosting (Friedman 2001; Hastie, Tibshirani, and Friedman 2017, 359–361). Unlike Random Forest, gradient tree boosting is an iterative process, where a new tree is developed to predict errors from the previous iteration (Hastie, Tibshirani, and Friedman 2017, sec. 10.10.2; James et al. 2021, sec. 8.2.3). The iteration goes on until a certain specified number. The results from each tree are then combined to produce the final prediction. A basic gradient tree boosting algorithm using the squared loss function proceeds as follows (adapted from Hastie, Tibshirani, and Friedman 2017, 357, 361, 364–365), where y and x are, as before, the outcome variable and the predictors from training data:

1. Before the initial iteration, set $f_0(x) = \bar{y}$.

2. Per iteration $m = 1, ..., M$:

   (a) Compute the residual for every observation: $u_{i,m} = y_i - f_{m-1}(x_i)$.

   (b) Fit a regression tree $f_m(x)$ to predict $u_{i,m}$, resulting in terminal regions $R_{j,m}$, $j = 1, ..., J_m$.

   (c) For every terminal region, compute $\kappa_{j,m}$ that minimizes the loss function, $l(\cdot)$, which is the mean of the residuals in the case of the squared loss function:

$$\begin{aligned}\kappa_{j,m} &= \arg\min_{\kappa_{j,m}} \sum_{x_i \in R_{j,m}} l(y_i, f_{m-1}(x_i) + \kappa_{j,m}) \\ &= \frac{1}{n_{j,m}} \sum_{x_i \in R_{j,m}} (y_i - f_{m-1}(x_i)),\end{aligned}$$

   where $n_{j,m}$ is the number of observations at j, m.

[6] The default value for the two main parameters are `ntree` = 500 and `mtry` = $n_x/3$, where n_x is the number of predictors. Hastie, Tibshirani, and Friedman (2017) note, "In our experience random forests do remarkably well, with very little tuning required."

(d) Update $f_m(x)$ as below:

$$f_m(x) = f_{m-1}(x) + \eta \sum_{j=1}^{J_m} \kappa_{j,m} I(x \in R_{j,m}),$$

where η is a learning rate parameter and $I(\cdot)$ is an indicator function.

3. The final output is: $\hat{y} = f_M(x)$.

Several different implementations of gradient boosting have been used previously in marcoeconomic forecasting (Chu and Qureshi 2023; Qureshi, Chu, and Demers 2021; Richardson, Florenstein Mulder, and Vehbi 2021; Yoon 2021). In the machine learning community nowadays, XGBoost (Chen and Guestrin 2016) and LightGBM (Ke et al. 2017) are known to be particularly good for predictive tasks ("XGBoost" 2023; Saha 2023). This paper implements XGBoost by the R `xgboost` package and LightGBM by the R `lightgbm` package.

3.3 Caveat on Parameter Settings

The parameter settings of the Bayesian VAR and tree ensemble models were determined, based on their out-of-sample predictive performance. However, there is no guarantee that the settings chosen for the out-of-sample prediction exercise in this paper will perform as well for predictive tasks in other time periods. While it is necessary to base the parameter settings on something in the first place, it is important to keep fine-tuning them as more data become available with time.

3.4 Baselines and Forecast Averaging

In addition to the hierarchical Bayesian VAR and tree-based ensemble models discussed so far, baseline forecasts for comparison are produced by random walk (i.e., the forecast at $t + 1$ being a value at t, the so-called "naive" forecast; see Hewamalage, Ackermann, and Bergmeir 2023) and by auto-ARIMA (Hyndman and Khandakar 2008). Hewamalage, Ackermann, and Bergmeir (2023, 796–797) argue that in many practical settings, the naive forecast, which is an extremely simple model, performs well and, there-

fore, should be a benchmark against more complex models. ARIMA is also a relatively straightforward univariate model. The auto-ARIMA algorithm determines the autoregressive and moving average parameters based on data, here using the small sample corrected Akaike Information Criteria.

This paper also calculates the average of the point forecasts across all six models employed per variable forecasted: random walk, auto-ARIMA, Bayesian VAR, Random Forest, XGBoost, and LightGBM. Hereafter, this "composite" model will be referred to as the forecast average. For clarity, the paper refers to each model whose point forecast is used to produce the forecast averages, as an individual model. The forecasting literature has found that simple forecast averaging can perform remarkably well (Genre et al. 2013; Makridakis and Winkler 1983; Stock and Watson 2004).

4 Data

This paper uses the following macroeconomic variables:

- Real gross domestic products (GDP)
- GDP deflator
- Real gross national product (GNP)
- GNP deflator
- Real modified gross national income (GNI*)
- GNI* deflator
- Real modified domestic demand (MDD)
- MDD deflator
- Real consumption
- Consumption deflator
- Harmonised Index of Consumer Prices (HICP)
- HICP-adjusted compensation of employees
- Participation rate (of people aged 15 or over)

- Employment rate (of people aged 15–64)
- Unemployment rate (of people aged 15–74)[7]

GNI* and MDD are Irish-specific measures designed to capture the domestic aspect of the economy, by trying to remove the effect of the activities of multinational enterprises that may not necessarily contribute to the domestic economy (Central Statistics Office 2023d, 2023m). The detail of the data sources is available in Appendix A.[8] All data were downloaded on the 12th of December 2023; as Irish macroeconomic data are often revised (Casey and Smyth 2016), the results here should be seen as conditional on the data retrieved then.

Given the VAR setup, the variable of the shortest time period available determines the first year of the dataset, which turns out to be 1998 – the first recorded year of quarterly data on the labor force on the online database of the Central Statistics Office of Ireland (CSO).[9] As level variables are converted to growth variables and 1-year lags are used (see the subsection, "Forecasting and Evaluation Methods"), the earliest time point of the outcome variables used in the models is 2000. If original data are in quarterly or monthly terms, they are aggregated to annual terms: summation for national account variables and averaging for rate variables. The summary statistics are presented in Table 1, where all variables are on the percentage-change scale.

One might wonder why years rather than quarters are used as the time unit of analysis, as the data of all variables used here, except for GNI*, were available on a quarterly

[7] These age groups for the labor force statistics are standard ones used in the data source, Central Statistics Office. For how these rates are calculated for the COVID-19 period, see the Central Statistics Office (2023e).

[8] Adding more variables was also tried, mainly drawing from the list of variables used in the Irish Fiscal Advisory Council Bayesian VAR model (Carroll 2020). In the current modeling context and the variables of interest to forecast, there was little predictive gain in terms of root mean squared errors in the out-of-sample prediction. There are two possible reasons. First, the bias-variance trade-off may be such that a predictive gain from bias reduction by adding more predictors is outweighed by precision reduction from variance amplification. Second, macroeconomic variables are generally collinear with each other either as a standalone variable or as a linear combination. Therefore, holding the number of observations constant, as the number of variables increases, a diminishing return of predictive gain may follow. For a similar point, see Bańbura, Giannone, and Reichlin (2010).

[9] The CSO also has another, legacy data on the labor force, which covers the years from 1988 up to 1997. However, the data collection methodology was different (Central Statistics Office 1998) and, therefore, is incompatible with the current data.

	Mean	SD	Min	Max
GDP	0.06	0.06	-0.05	0.24
GDP Deflator	0.02	0.03	-0.05	0.08
GNP	0.05	0.05	-0.08	0.14
GNP Deflator	0.03	0.04	-0.05	0.09
GNI*	0.03	0.05	-0.10	0.14
GNI* Deflator	0.03	0.04	-0.05	0.11
MDD	0.03	0.05	-0.11	0.11
MDD Deflator	0.02	0.03	-0.06	0.07
Consumption	0.03	0.05	-0.10	0.11
Consumption Deflator	0.02	0.02	-0.06	0.06
HICP	0.02	0.02	-0.02	0.08
Compensation of Employees	0.04	0.05	-0.08	0.10
Participation Rate	0.00	0.02	-0.03	0.05
Employment Rate	0.01	0.03	-0.09	0.05
Unemployment Rate	0.00	0.24	-0.28	0.87

Table 1: Summary statistics. All variables are on the percentage change scale. SD: standard deviation.

basis at the time of the analysis. It might be argued that aggregating these variables to annual terms causes some information to be lost. However, statistical optimization using quarterly data can result in sub-optimal predictive accuracy in annual terms. This can happen because the errors at the annual level are not a monotonic function of those at the quarterly level.

For example, assume that the true quarterly growth values in a year are $\{0.02, -0.01, 0.03, 0.01\}$. Also assume that there are two models using quarterly data as the unit of analysis. One model produces the predicted values, $\{0.05, -0.03, 0.05, -0.01\}$; the other generates the predicted values, $\{0.01, 0.00, 0.02, 0.00\}$. The root mean squared error (RMSE) at the quarterly level is 0.023 for the first model and 0.01 for the second model; the second model is better. However, the RMSE at the annual level (here, calculated using the mean of the quarterly growth values) is 0.0025 for the first model and 0.005 for the second model; the first model is better. In other words, minimizing the RMSE of predicted values from a model using quarterly data does not necessarily lead to the minimization of the RMSE of the aggregated predicted values.

One disadvantage of using annual rather than quarterly data is that if forecasting is to

be done in real time, there can be the situation where quarterly data are available up to the 2nd or 3rd quarter of the current year so that the annual figure of a variable in that year might be more accurately forecasted by a model using quarterly data than one using annual data. In such a case, one might want to first predict an annual figure for the current year by a short-term forecasting model and plug them into the VAR data frame for a medium-term forecasting model.

5 Forecasting and Evaluation Methods

For predictive tasks, it is important to find the measures of underlying concepts that behave in a similar fashion. National account variables and inflation measures generally increase over time in levels, while rate variables related to the labor force (participation, employment, and unemployment rates) usually exhibit a cyclical move in levels. Meanwhile, once these variables are all measured on the percentage-change scale, it is reasonable to theoretically expect that they behave in a similar fashion, for example, that a proportional change in GDP is followed by a proportional change in the unemployment rate.[10]

Forecasts are produced up to five years ahead (i.e., horizon = 5), based on 1-year lags. One lag is chosen, based on the results from the otherwise equivalent Bayesian VAR model showing that the posterior probability of the model is greater given one lag than given two lags (assuming the uniform prior model probability). For Bayesian VAR, the `predict` function from the R `BVAR` package is used to generate forecasts over a 5-year horizon. For the three machine learning algorithms, 1-year ahead forecasts are generated first, and the algorithms are run again including these forecasted values in the training data. This is repeated five times to generate total 5-year ahead forecasts.[11]

[10]The variables used here can take only positive values in levels. The use of growth rates rather than first differences make sure that *ex post* calculated forecasted values in levels do not go negative. Coulombe et al. (2021) point out that in machine-learning time-series forecasting, transforming data into growth rates is only one option, and other transformations may generate greater predictive accuracy, depending on variables and algorithms. This paper leaves this point for future research.

[11]In parametric models, there is no difference, in expectation, between generating forecasts with and without reestimating the parameters and stacking forecasted values onto the original

The predictive performance of each model is evaluated by the RMSE for each variable. RMSEs penalize a larger error more than absolute mean errors. Since a single extreme event usually inflicts a greater pressure on the government and society than a series of less extreme events (even if the sum of the effects of the latter is equal to the effect of the former), RMSEs are more suitable in evaluating the predictive performance of macroeconomic forecasting models.

The paper does not compare its out-of-sample predictive performance with the official forecast for the same period that was done in real time by the Department of Finance. This is because Irish macroeconomic data are often subject to revisions (Casey and Smyth 2016; Conroy and Casey 2019). It is possible that the data used at the time of official forecasting were a different version from what the models of this paper use. Therefore, direct comparison is impossible.

6 Results

This section focuses on the performance measures, the RMSEs, for the period of 2015–2019 (based on the data from 1999 to 2014) and for the period of 2018–2022 (based on the data from 1999 to 2017); the predicted values of each model for these periods are displayed in Figures 7 and 8 in Appendix B. The RMSEs for the forecast period of 2015–2019 are presented in Figure 1, while those for the forecast period of 2018–2022 are displayed in Figure 2. The upper panels are the RMSEs on the original scale; the lower panels are those on the standardized scale, calculated as the RMSEs on the original scale divided by the standard deviation of the corresponding variable within the in-sample data used. For brevity of presentation, the results from additional analysis using the periods of 2016–2020 and 2017–2021 are displayed in Appendix C, without

observed data. This is because these forecasted values are a function of the parametric model fitted into the data and, therefore, consistent with the originally estimated parameters of the model. In the tree ensemble models used here (which are nonparametric), this is generally not the case. For extreme example, if splitting is done such that every node has one observation, it follows that a forecasted value ends up with the node that did not exist in the previous model (unless the previous model observed the same combination of the outcome and predictor values). This means that the updated model identifies a different functional form from the original one and, therefore, is likely to generate different forecasts in the next forecast horizon.

hampering the discussion of key findings.

Three key findings are as follows. First, none of the models outperforms one another in every variable and time period. Instead, different models perform better or worse than others, depending on different variables and time periods. One striking example is the random walk forecasts for the unemployment rate. In the period of 2015–2019, the random walk model is significantly more accurate than any remaining ones. This is because of the similar values being persistent after the last in-sample year. Meanwhile, in the period of 2018–2022, its performance is the worst, because the out-of-sample years exhibit a highly volatile trend (as it includes the COVID-19 period). In other words, the predictive performance can remarkably differ across different time periods, even when the same model is used for the same variable. The overall finding is consistent with the conventional wisdom: There is no single model that does everything well.

Second, no model (except for random walk for 2015–2019) performs well in forecasting the unemployment rate, as the RMSEs for the growth rate of the unemployment rate on the original scale are much larger than those of the other variables. This may be because the empirical variation (measured by the standard deviation) in the percentage change of the unemployment rate is much greater than those of the remaining variables in the in-sample data used. According to Table 1, the standard deviation of the growth rate of the unemployment rate is 0.24, while those of the other variables are all around or less than 0.06. In terms of the RMSEs on the standardized scale, forecasting performance for the unemployment rate is actually as good as many other variables. In other words, although the size of the forecast errors for the unemployment rate looks large on the original scale, the models are performing well given the large variation in the data used.

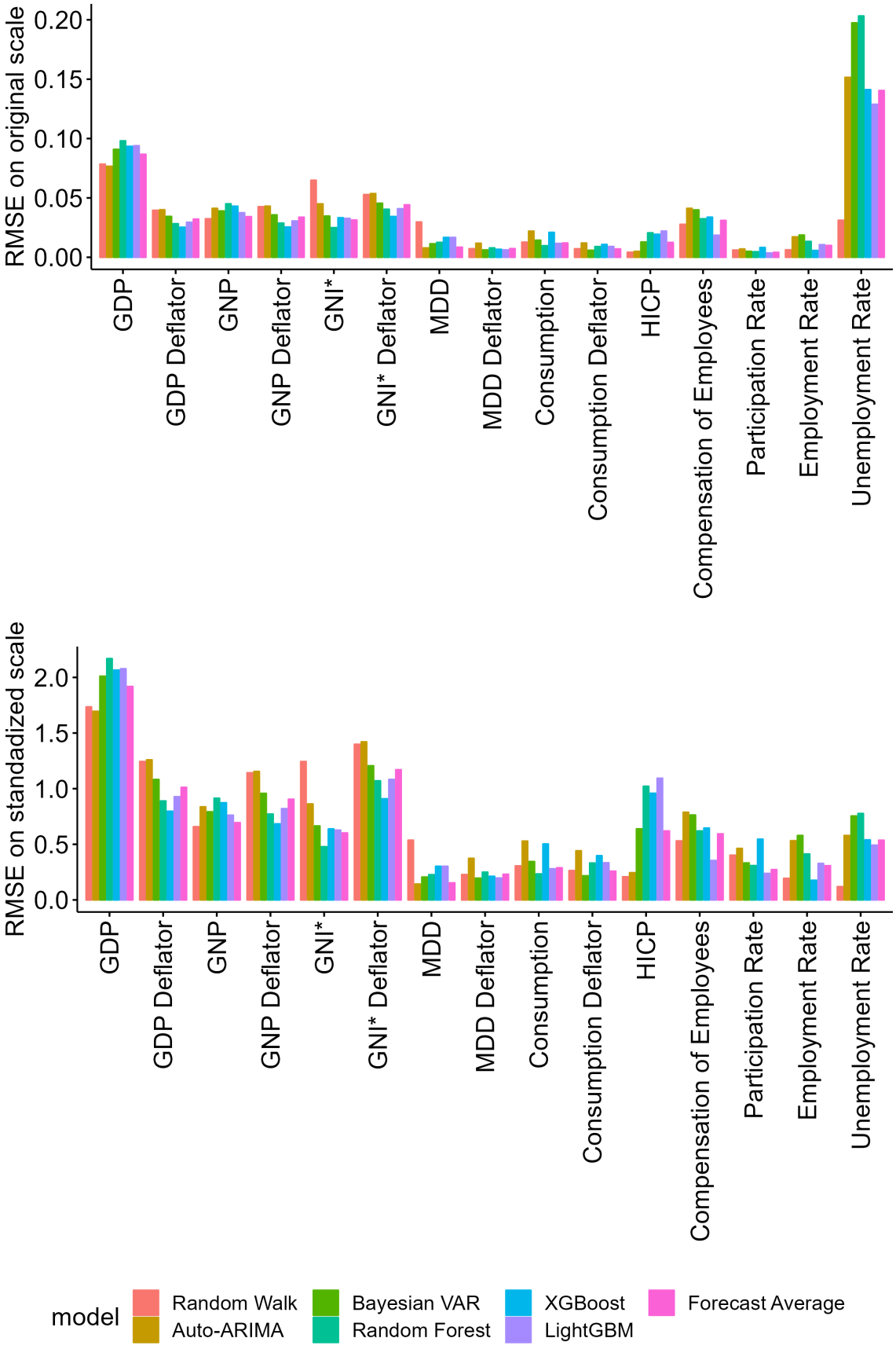

Figure 1: RMSEs over the 5-year horizon of 2015–2019.

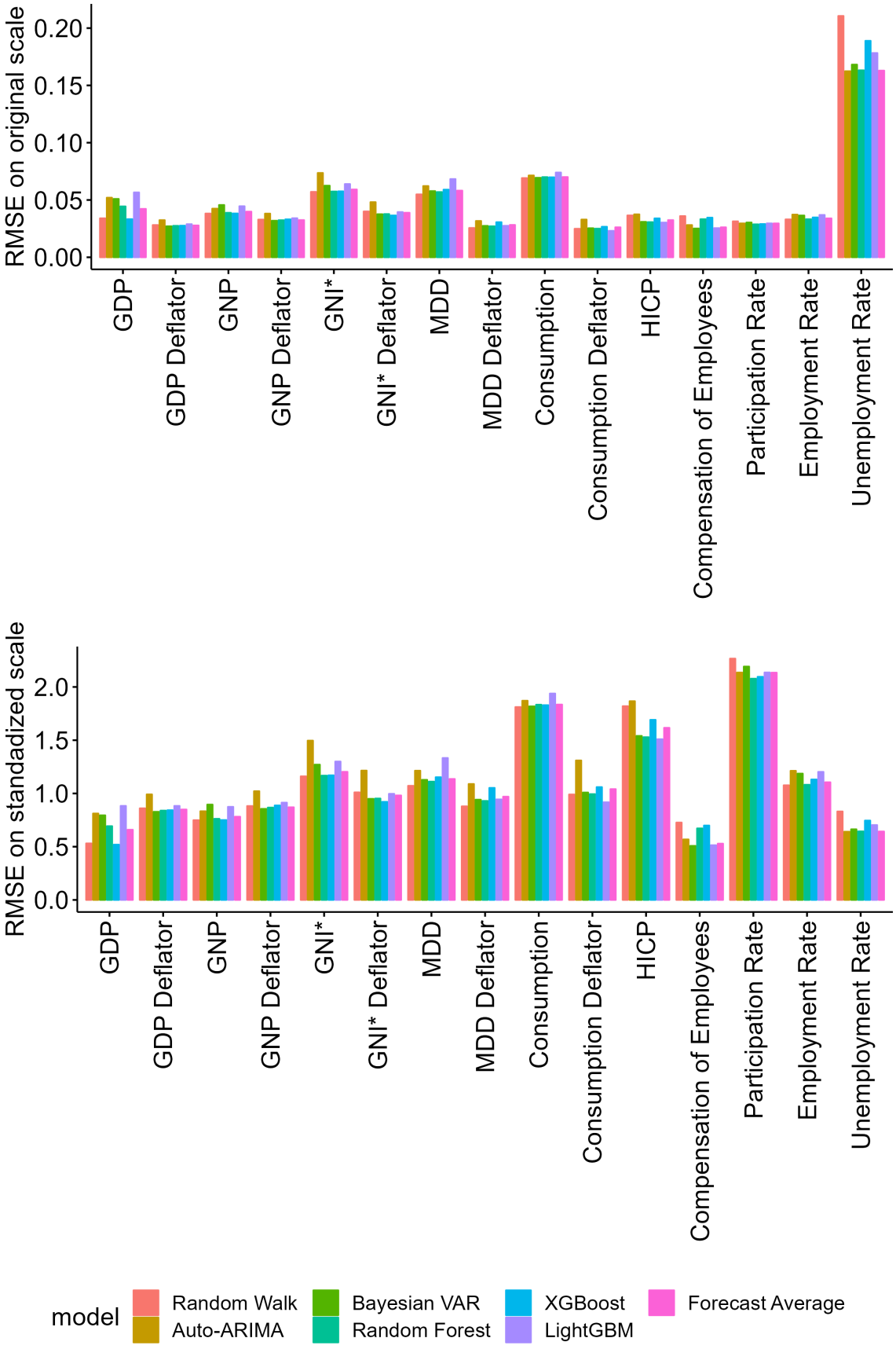

Figure 2: RMSEs over the 5-year horizon of 2018–2022.

Third, the stable performance of the forecast average is noteworthy. Figure 3 and Figure 4 present the relative performance of each of the models over the best performing model per variable and time period, as the difference in their RMSEs. In each figure, the largest number of each y-axis is set to be the greatest RMSE difference per variable and time period. The lower a bar is, the closer a model is to the best performer; no bar means that the best performer is itself. The differences in the RMSEs of the forecast averages are comparatively small across all variables, except for the unemployment rate in 2015–2019. On a few occasions, the forecast average is the best performer, as seen in Figures 11 and 12 in Appendix C.

This consistent performance is noteworthy. The true values are unknown in actual forecasting tasks. Even if a model performed very well for a certain variable in an out-of-sample prediction exercise, its performance could turn out different in the future, for example, because of a structural change (e.g., compare the RMSEs of the random walk for the unemployment rate in 2015–2019 and 2018–2022). In other words, it is impossible to select the best model (and parameter setting) out of all possible candidates *ex ante*. For these reasons, rather than relying on an individual model that appears to perform well in a certain context, using the forecast averages seems to be a credible choice in practice. This point, together with the first key finding, suggests that there is a gain in developing and utilizing different models.

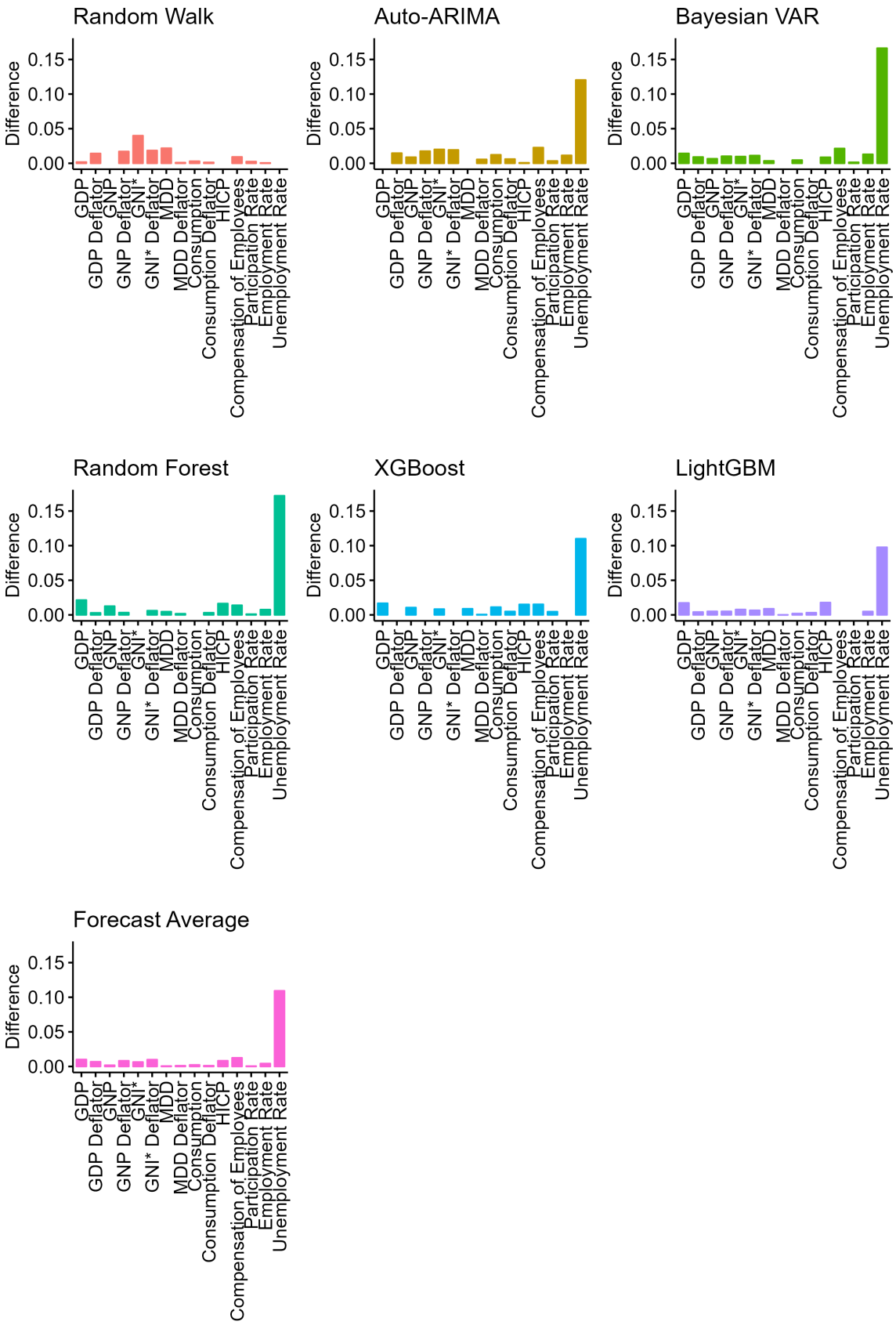

Figure 3: Difference between each of the models and the best performer in RMSEs over the 5-year horizon of 2015–2019. A zero difference means that the model of concern is the best performer.

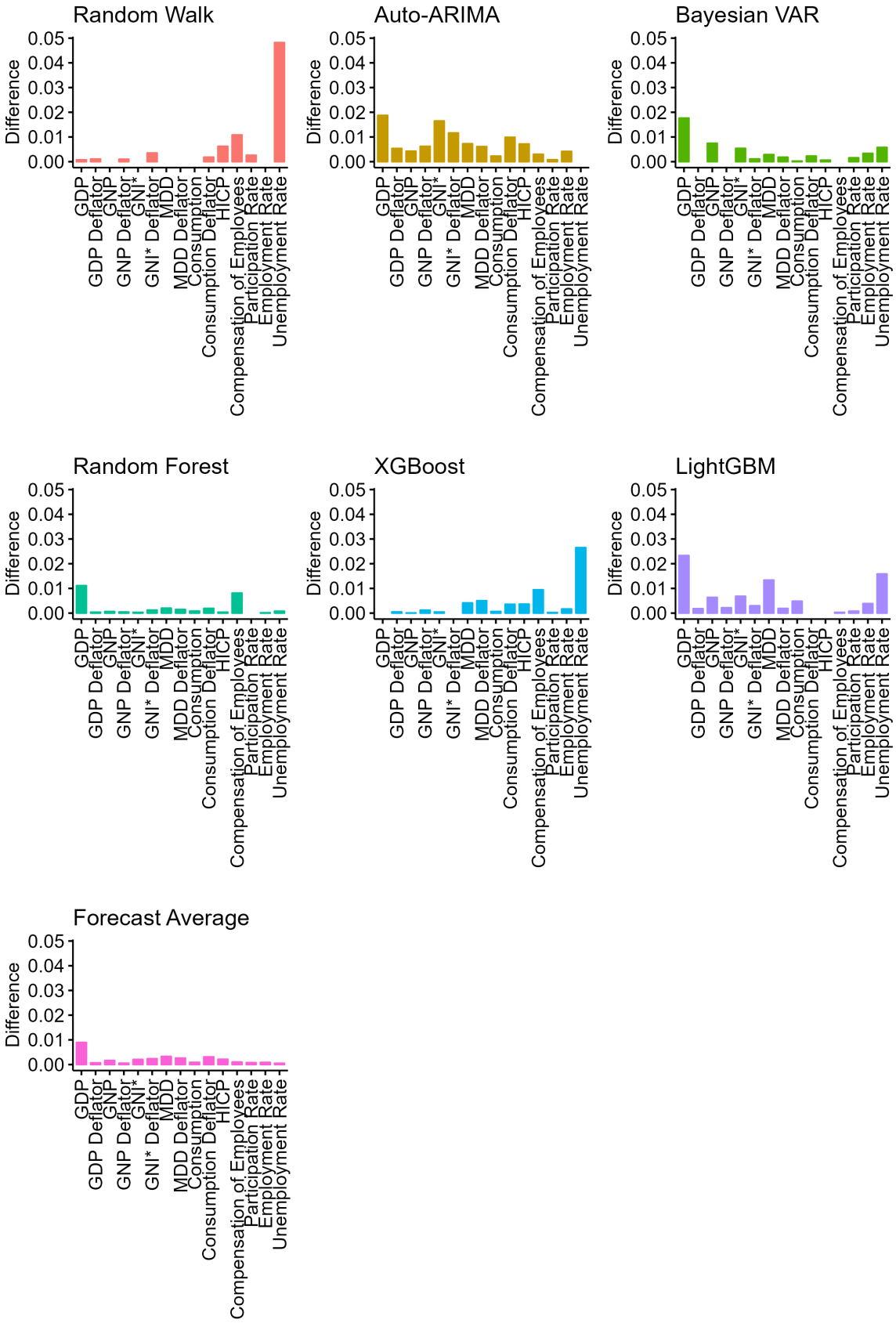

Figure 4: Difference between each of the models and the best performer in RMSEs over the 5-year horizon of 2018–2022. A zero difference means that the model of concern is the best performer.

Given its robust performance, this paper also considers how to generate a prediction interval for a forecast average. It is generally incorrect to use the average of the upper and lower bounds of the prediction intervals across different individual models as the prediction interval of the forecast average (Nowotarski and Weron 2015, 797).

Here, a prediction interval is generated based on a simulation approach using the following normal distribution:

$$\begin{aligned} \hat{y}_{t+h} &\sim N(\mu, \sigma), \\ \mu &= \frac{1}{n}\sum_{i=1}^{n} \hat{y}_{t+h,i}, \end{aligned} \tag{2}$$

where μ is the mean parameter; σ is the standard deviation parameter; n is the total number of individual models used; $\hat{y}_{t+h,i}$ is the point forecast at a time point in the forecast horizon $t+h$ from an individual model i.

σ is calibrated empirically per variable by the following algorithm. The algorithm begins with a small σ value and generates the prediction intervals over the forecast horizon, $t+h, ..., t+H$, where H is the final forecast horizon. Then, it calculates the rate at which the mean values of the upper and lower bounds of the intervals cover the in-sample data used. If the rate is between the preset credible level plus / minus some margin (here, set as 0.90 ± 0.02), the σ value is accepted. If the rate is below the credible level minus the margin, the σ value is rejected and a slightly greater σ value is tried. If the rate is above the credible level plus the margin, the σ value is rejected and the algorithm ends. If a σ value is accepted, the algorithm computes the mean squared difference (MSD) between the mean upper and lower bound values of the intervals and the in-sample data.[12] Among those σ values accepted, the algorithm identifies the one that minimizes the MSD. Such a σ value is used for the σ parameter in equation 2.

[12] Formally, $MSD = \frac{1}{2}\left(\frac{1}{H}\sum_{h=1}^{H}\left((y_{t+h} - \bar{b}_u)^2 + (y_{t+h} - \bar{b}_l)^2\right)\right)$, where $\bar{b}_u$ is the mean upper bound value over the forecast horizon, $\bar{b}_l$ is the mean lower bound value over the forecast horizon, and y is in-sample data.

Once the normal distribution in equation 2 is specified, any quantile values are easy to compute. Figures 5 and 6 present the results with the 90% prediction intervals generated based on the above setup, for the periods of 2015–2019 and 2018–2022 respectively. The intervals are generally wide across variables and forecast horizons. This is unsurprising, given the volatile nature of the empirical data used (see Figures 7 and 8 in Appendix B).

The average coverage rate of the intervals is 95% in the out-of-sample period of 2015–2019 and 80% in that of 2018–2022. The coverage rate is somewhat worse in the latter period, as it includes the COVID-19 and War in Ukraine years, which experienced extraordinary macroeconomic changes.[13] There is always a trade-off between the coverage rate and width (i.e., informativeness) of prediction intervals, especially when variables to forecast are anticipated to exhibit extreme values from time to time. As mentioned in the introduction, the models of this paper are designed to forecast regular trends over a medium term. Nonetheless, it is also undesirable to make a prediction interval too narrow to signal any possibility of extreme events. In the current setup, although the coverage rate for the volatile period of 2018–2022 is less than theoretically expected from a perspective of frequentist statistics, the differences between the actual extreme values and the bounds are generally not significantly large. Thus, the aforementioned calibration algorithm for the σ parameter seems to achieve a good balance between the coverage rate and width of prediction intervals, at least in the current out-of-sample prediction exercise.

[13] Interestingly, when only the COVID-19 period was included as for the periods of 2016–2020 and 2017–2021, the average coverage rates were better: 0.91 and 0.87 respectively. In other words, the presence of both the COVID-19 period and the War in Ukraine period seems to be the main cause of the worse coverage rate of the intervals for 2018–2022.

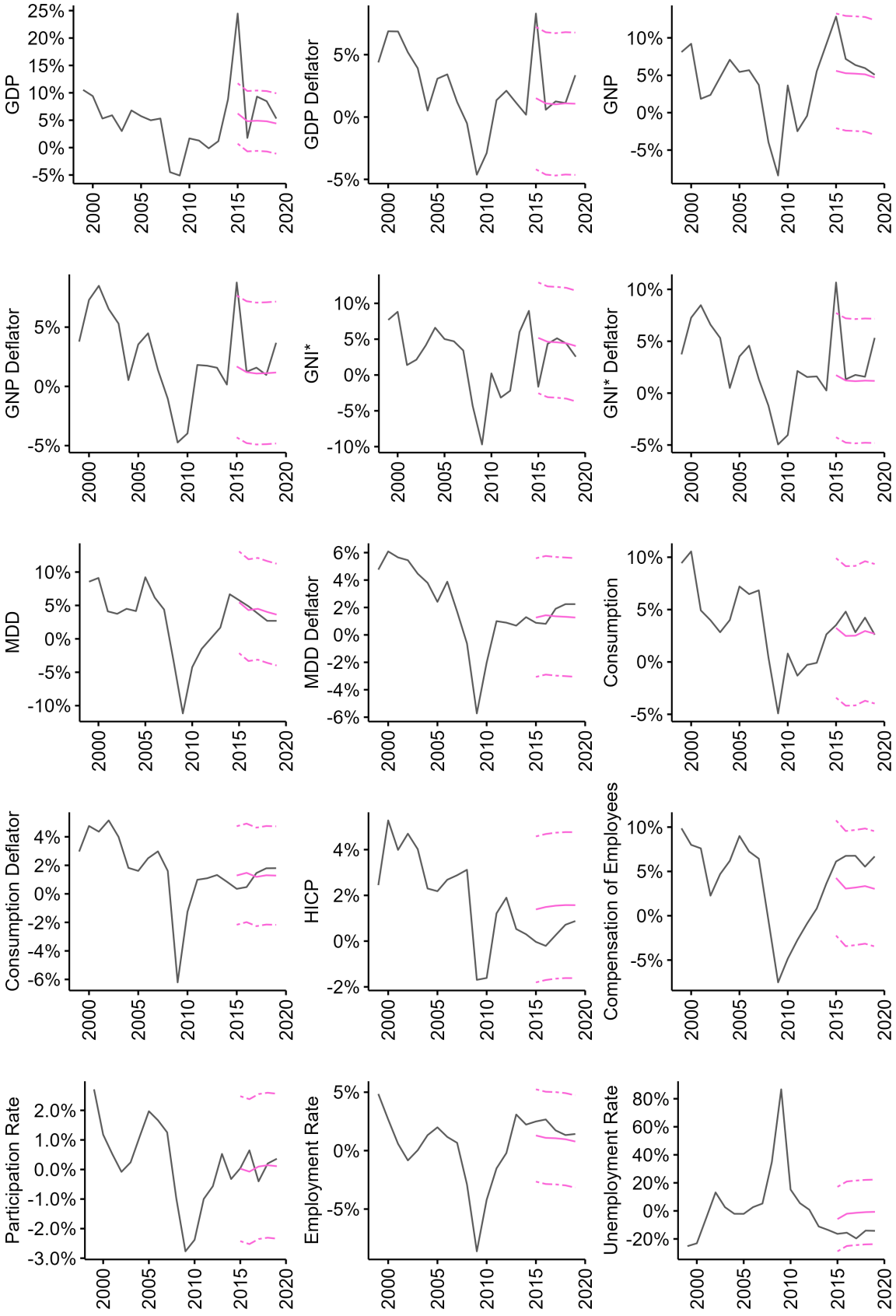

Figure 5: Forecast averages with the 90% prediction intervals for 2015–2019. The values on the y-axes are percentage changes.

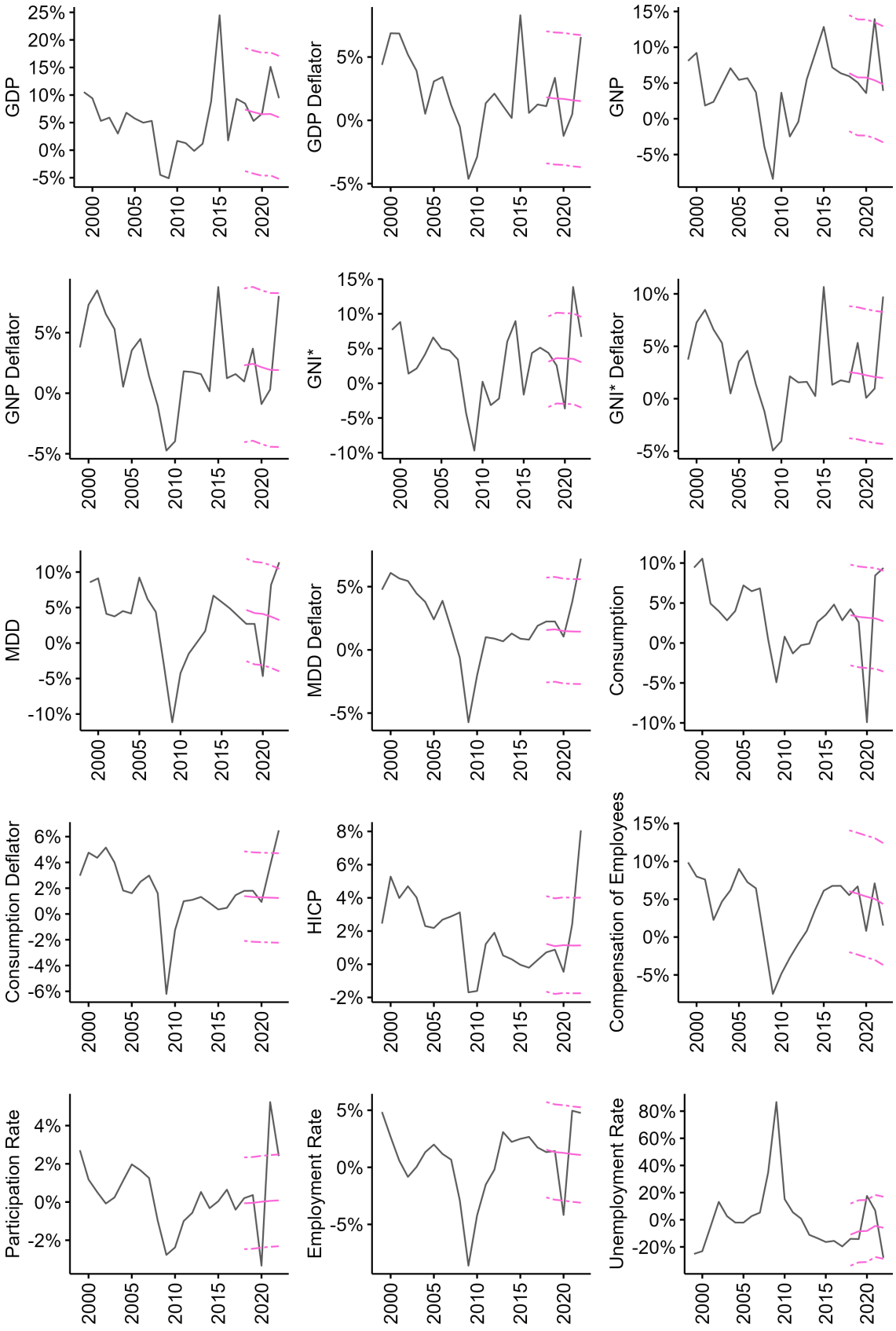

Figure 6: Forecast averages with the 90% prediction intervals for 2018–2022. The values on the y-axes are percentage changes.

7 Conclusion

This paper has employed a VAR setup to develop self-contained multivariate models for the medium-term forecasting of (often volatile) Irish macroeconomic variables. It has applied the VAR setup to hierarchical Bayesian VAR, Random Forest, XGBoost, and LightGBM. The paper has also used random walk and auto-ARIMA as the baseline models. Furthermore, it has utilized a forecast averaging approach, taking the averages of the forecasts from these individual models.

The results have showed that different models perform better or worse than others, depending on contexts. It has also been found that the forecast averages from all individual models used perform in a stable and consistent manner and seem to be a credible choice for actual forecasting. The paper has proposed a simulation method to generate the prediction interval of a forecast average.

By these findings, the paper contributes to the literature on macroeconomic forecasting, particularly to the one for small open economies (e.g., Botha et al. 2023; Carroll 2020; Conroy and Casey 2017; Gupta and Kabundi 2010; Hou, Nguyen, and Zhang 2023; Kenny, Meyler, and Quinn 1998; Marcellino and Sivec 2021). Future research might try the same approach to different countries than Ireland. In addition, from a policy perspective, it will be useful to compare the predictive performance of each model used in this paper in real time, for further fine-tuning and model development.

8 Appendix A: Data Sources

- Real GDP, real GNP, real GNI*, real MDD, real consumption, HICP, participation rate, employment rate, unemployment rate: Central Statistics Office (2023i, 2023g, 2023k, 2023j, 2023a, 2023l), downloaded on 12th December 2023.
- GDP deflator, GNP deflator, GNI* deflator, MDD deflator, consumption deflator: the author's calculations, based on the data on the current and constant values from the Central Statistics Office (2023i, 2023f, 2023g, 2023k, 2023j), downloaded

on 12th December 2023.

- HICP-adjusted compensation of employees: the author's calculations, based on the data on the current value of the compensation of employees and the HICP from the Central Statistics Office (2023c, 2023b, 2023a), downloaded on 12th December 2023.

9 Appendix B: Predicted Values per Model

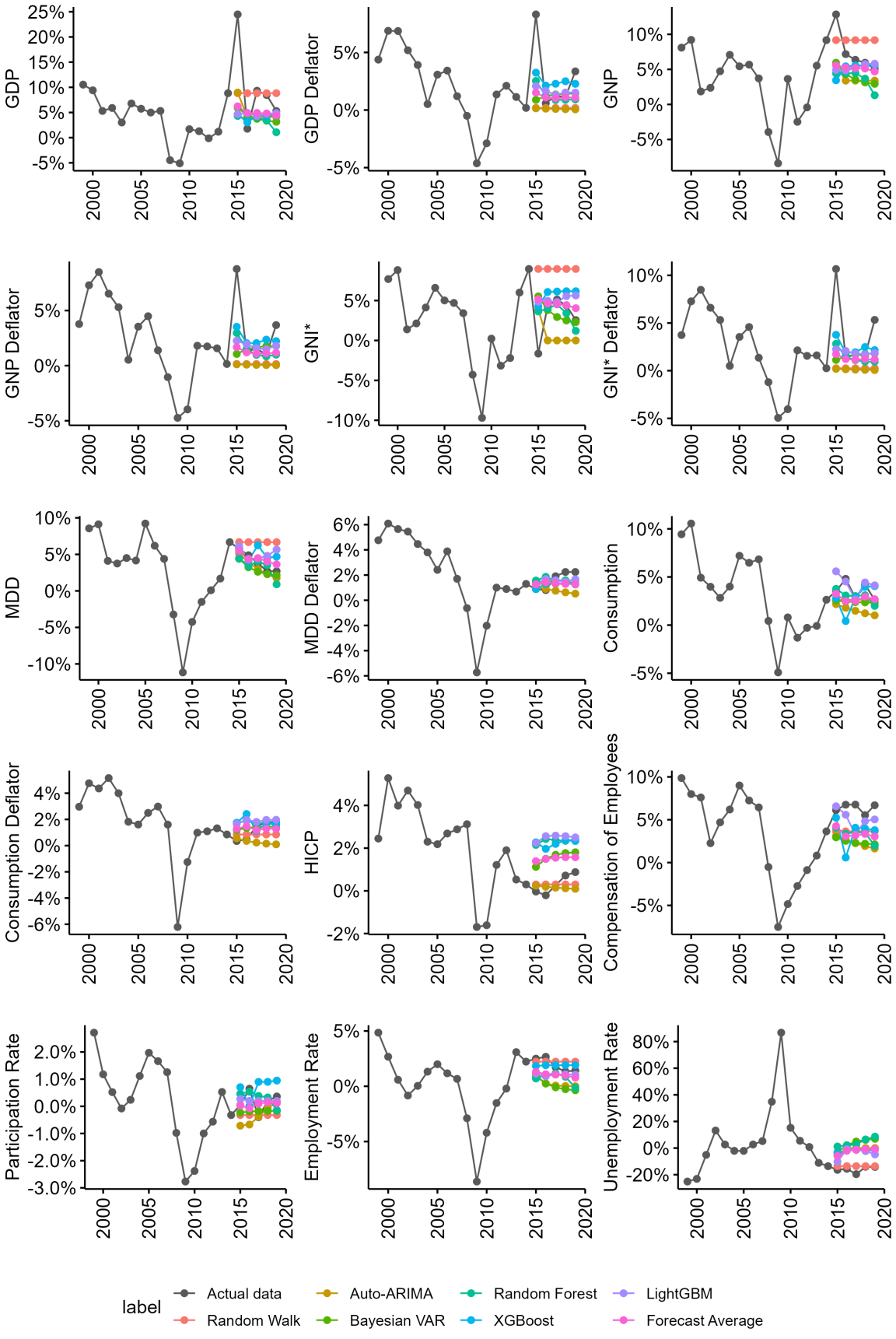

Figure 7: Predicted values over the 5-year horizon of 2015–2019. The values on the y-axes are percentage changes.

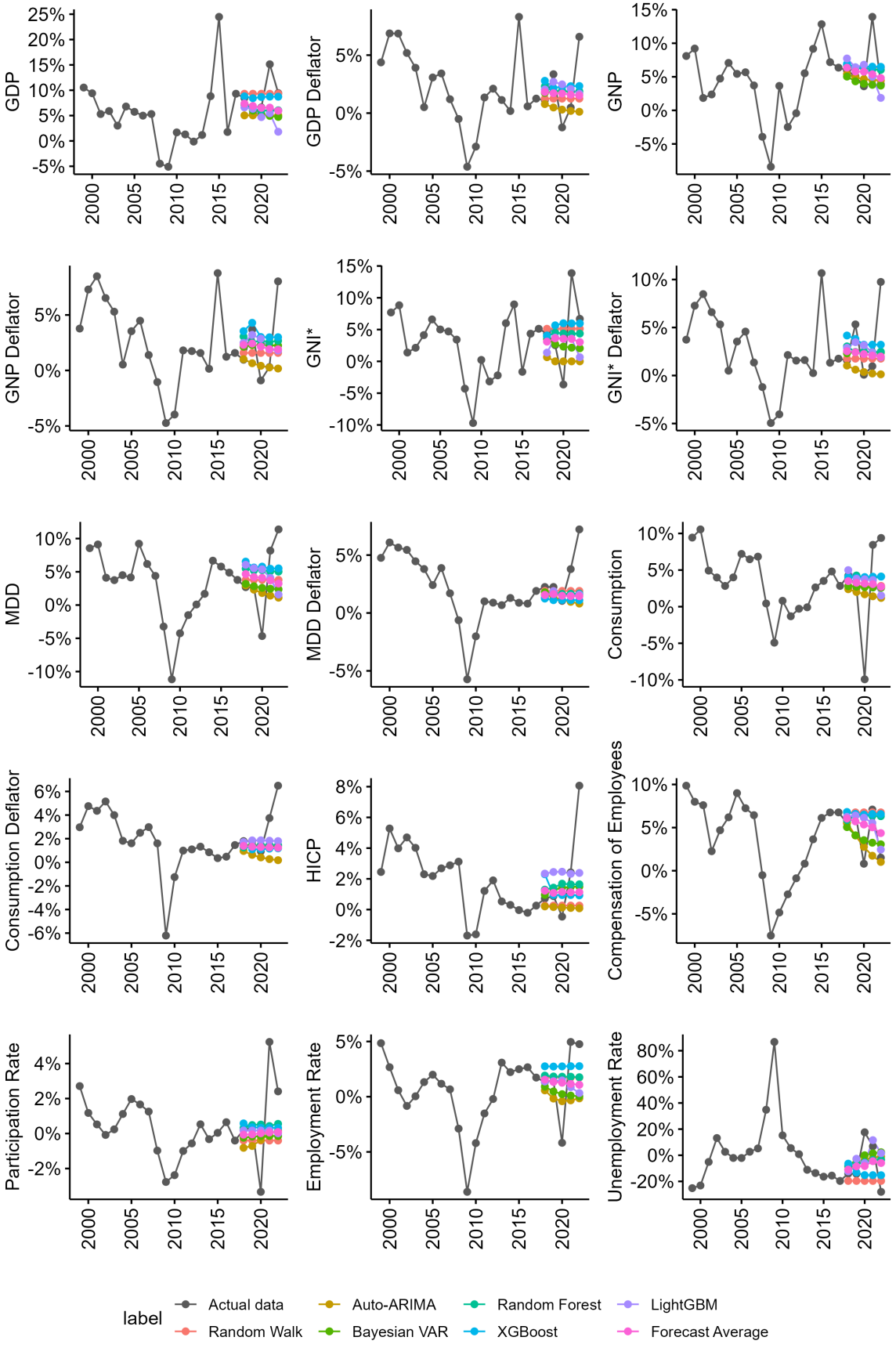

Figure 8: Predicted values over the 5-year horizon of 2018–2022. The values on the y-axes are percentage changes.

10 Appendix C: Additional Results of 2016–2020 and 2017–2021

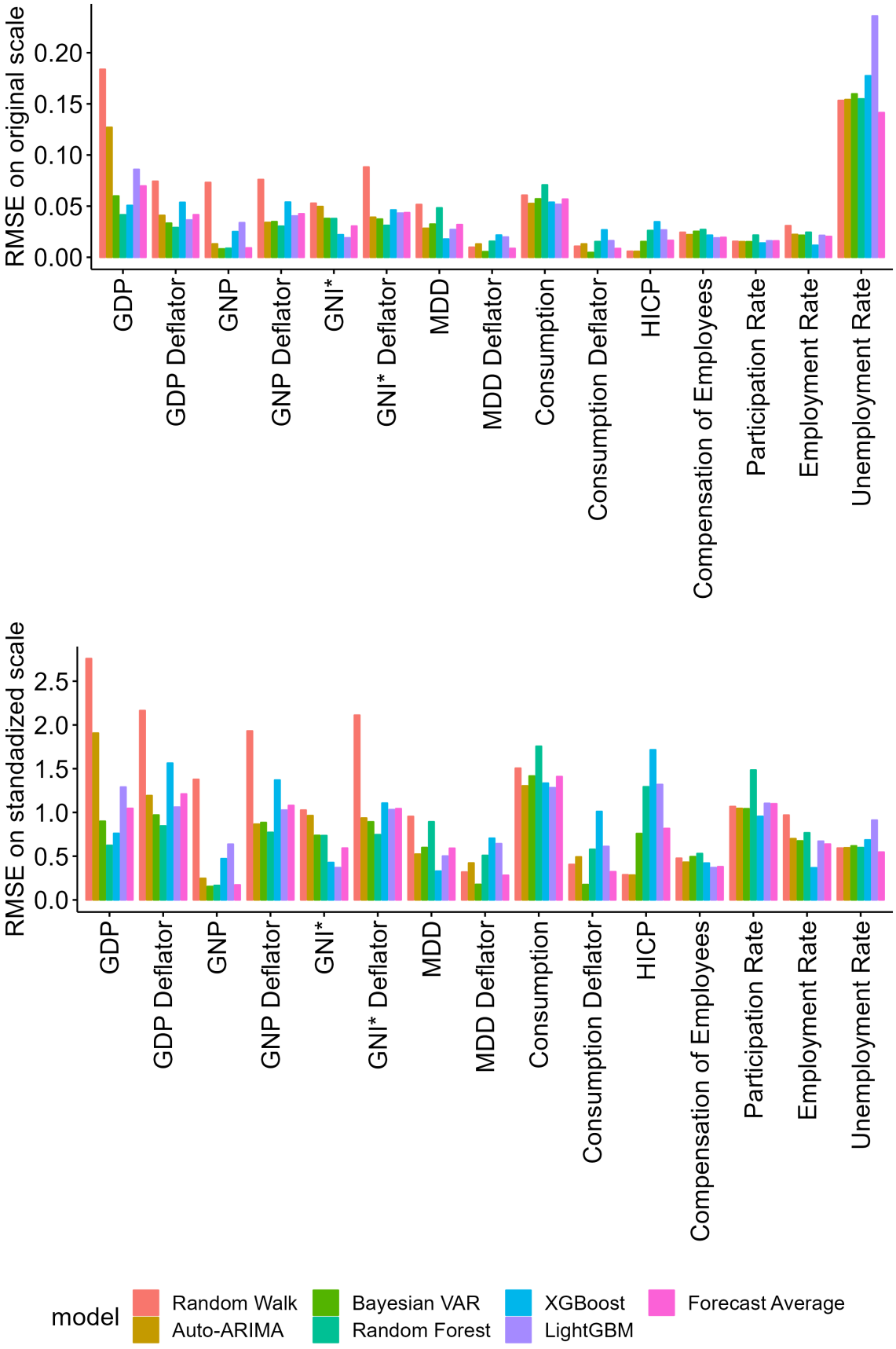

Figure 9: RMSEs over the 5-year horizon of 2016–2020.

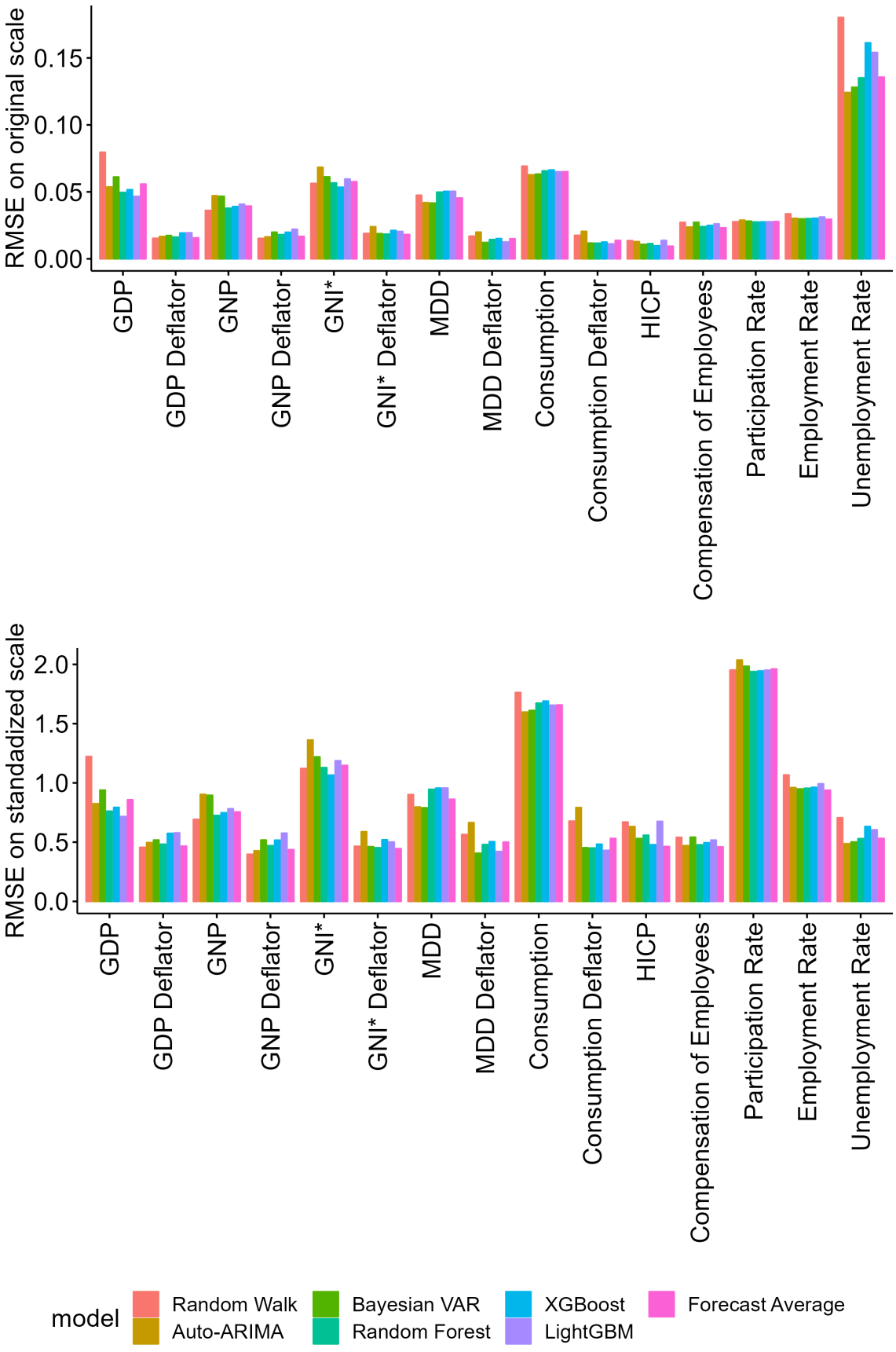

Figure 10: RMSEs over the 5-year horizon of 2017–2021.

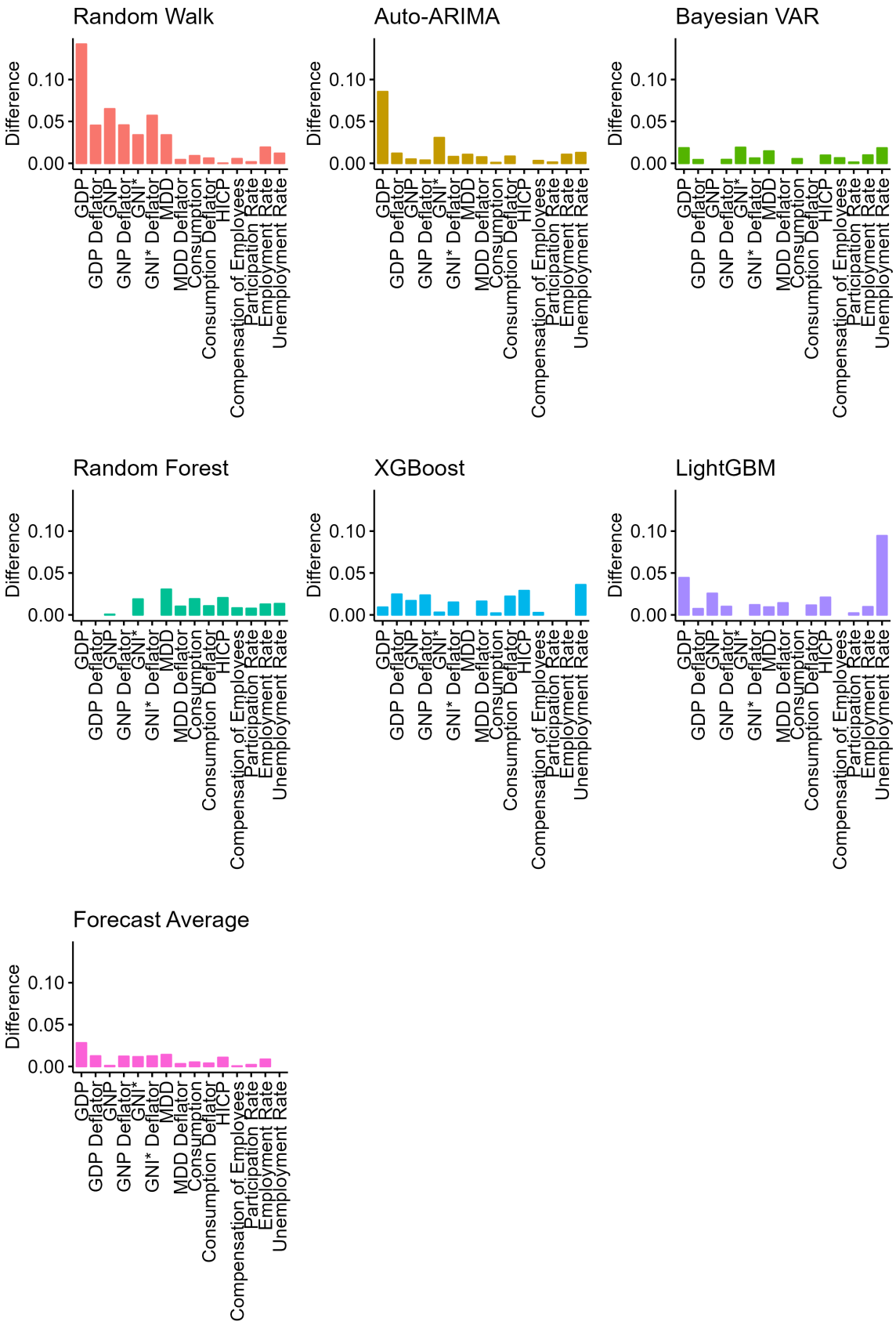

Figure 11: Difference between each of the models and the best performer in RMSEs over the 5-year horizon of 2016–2020. A zero difference means that the model of concern is the best performer.

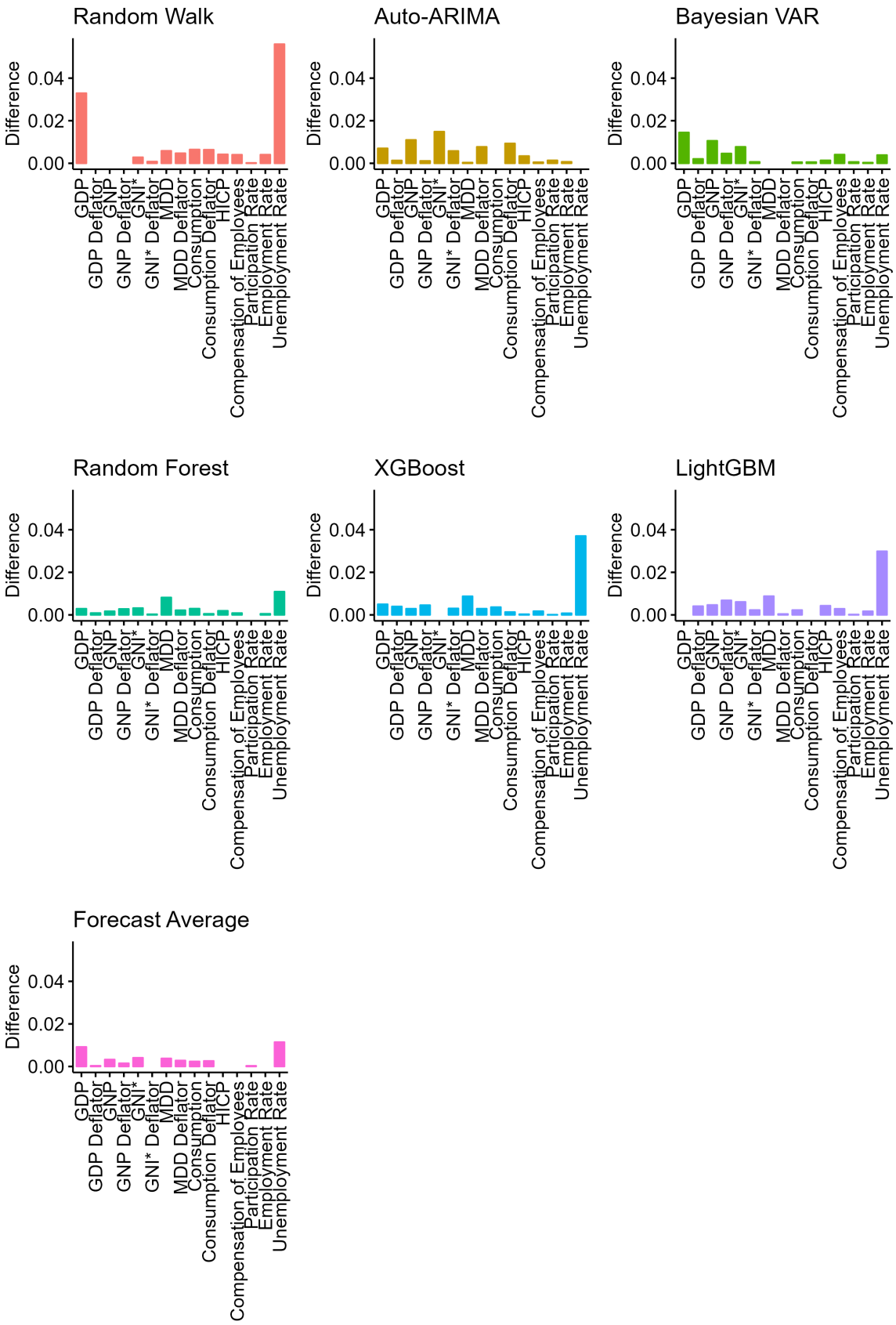

Figure 12: Difference between each of the models and the best performer in RMSEs over the 5-year horizon of 2017–2021. A zero difference means that the model of concern is the best performer.

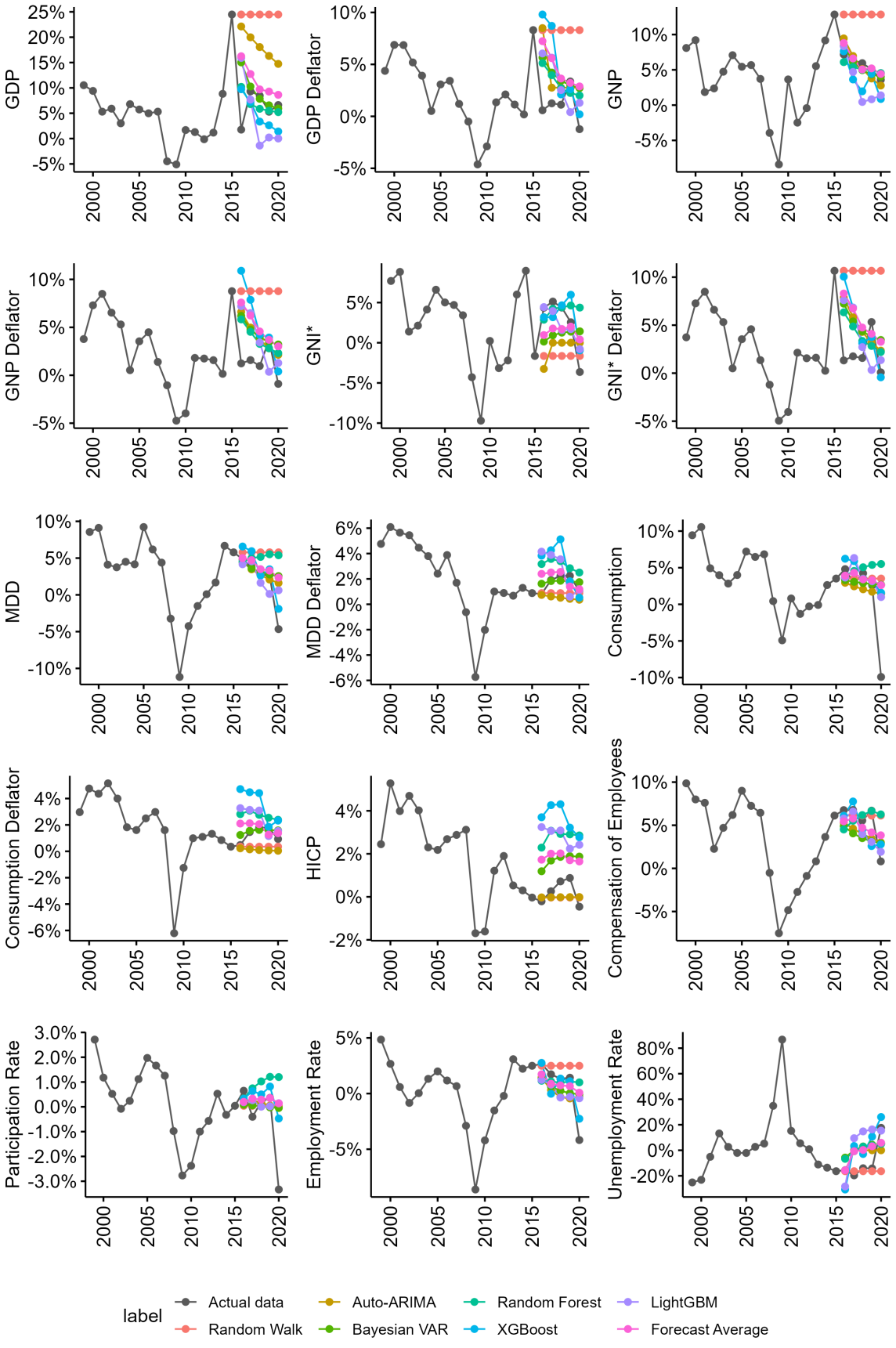

Figure 13: Predicted values over the 5-year horizon of 2016–2020. The values on the y-axes are percentage changes.

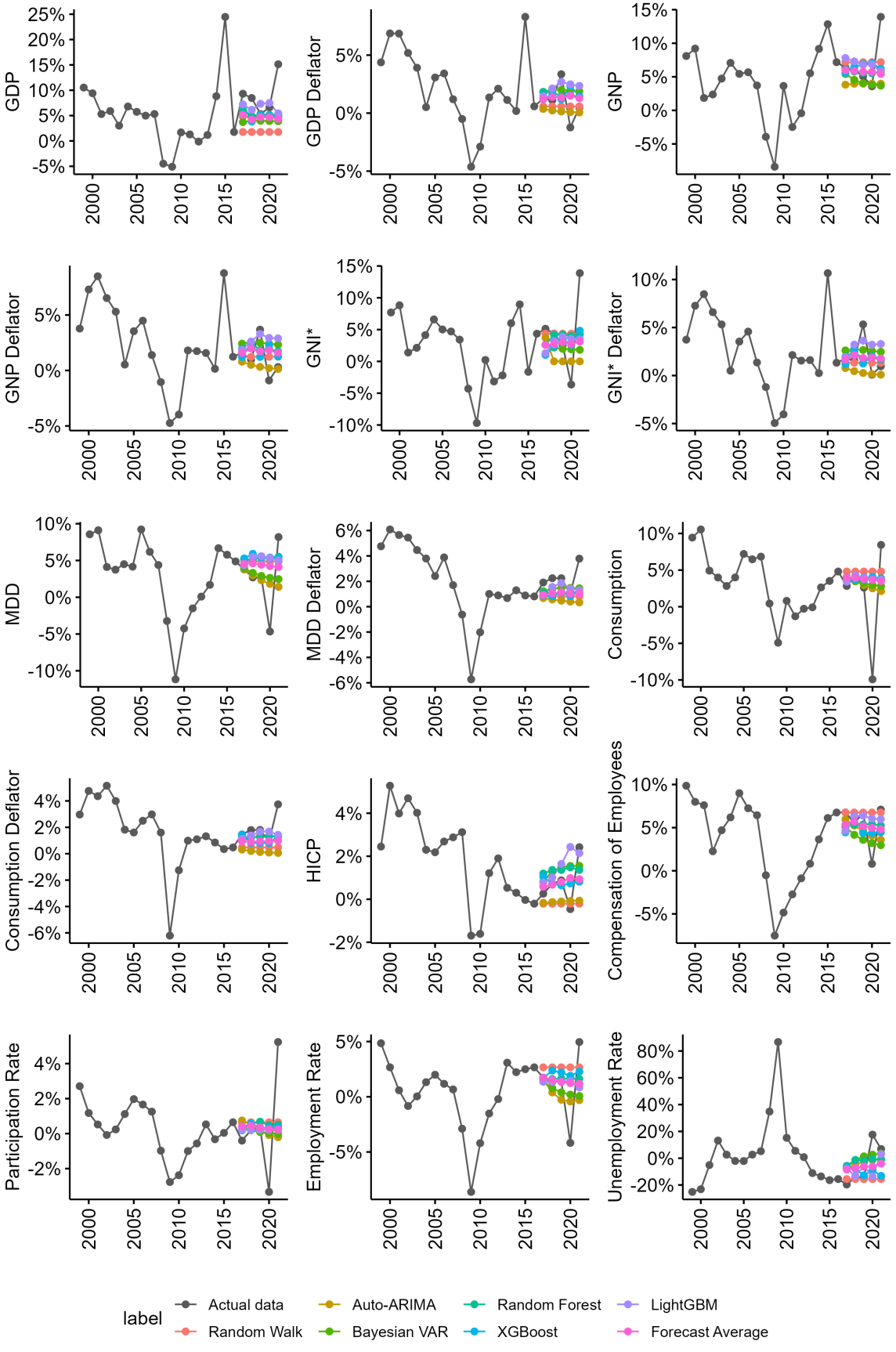

Figure 14: Predicted values over the 5-year horizon of 2017–2021. The values on the y-axes are percentage changes.

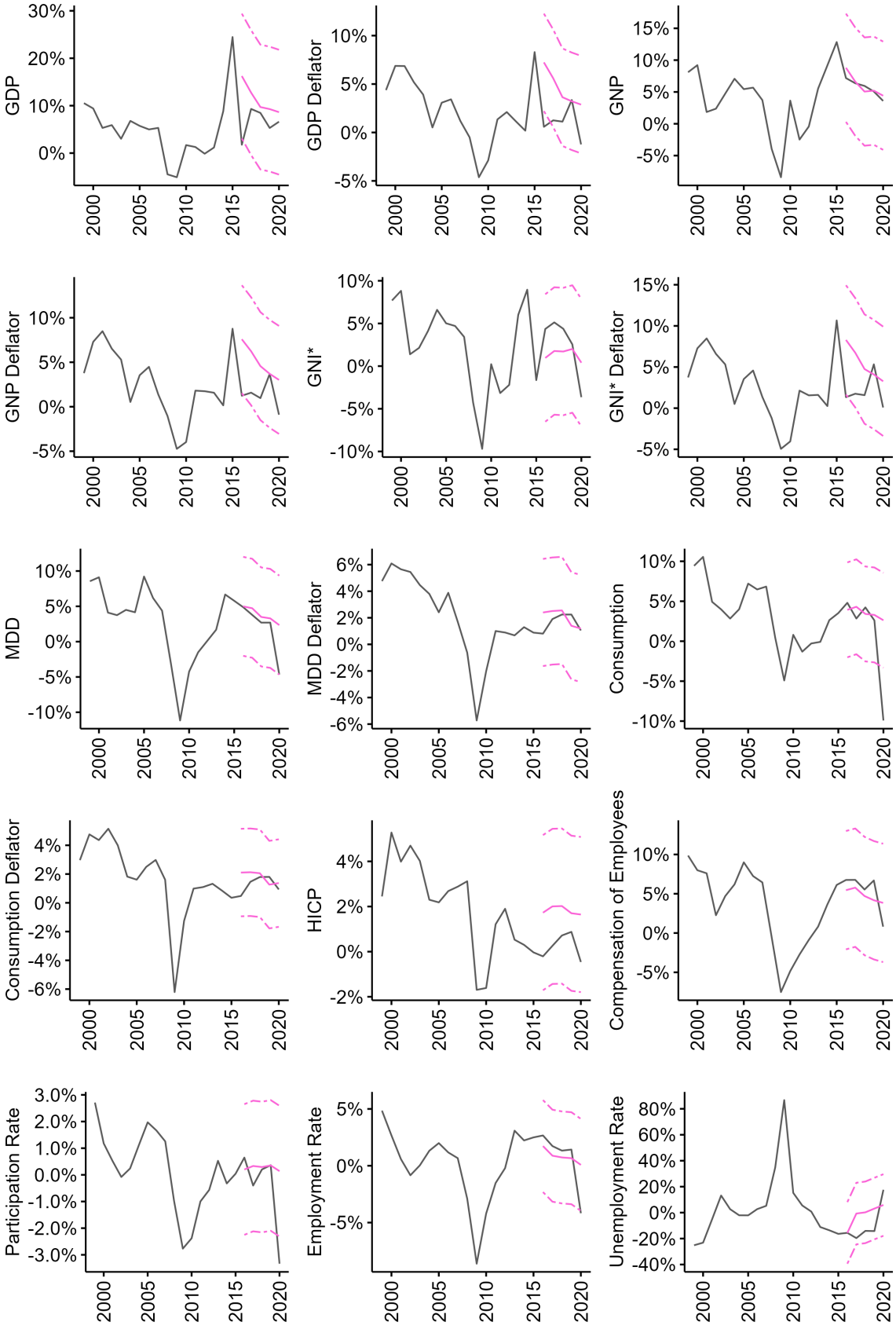

Figure 15: Forecast averages with the 90% prediction intervals for 2016–2020. The values on the y-axes are percentage changes.

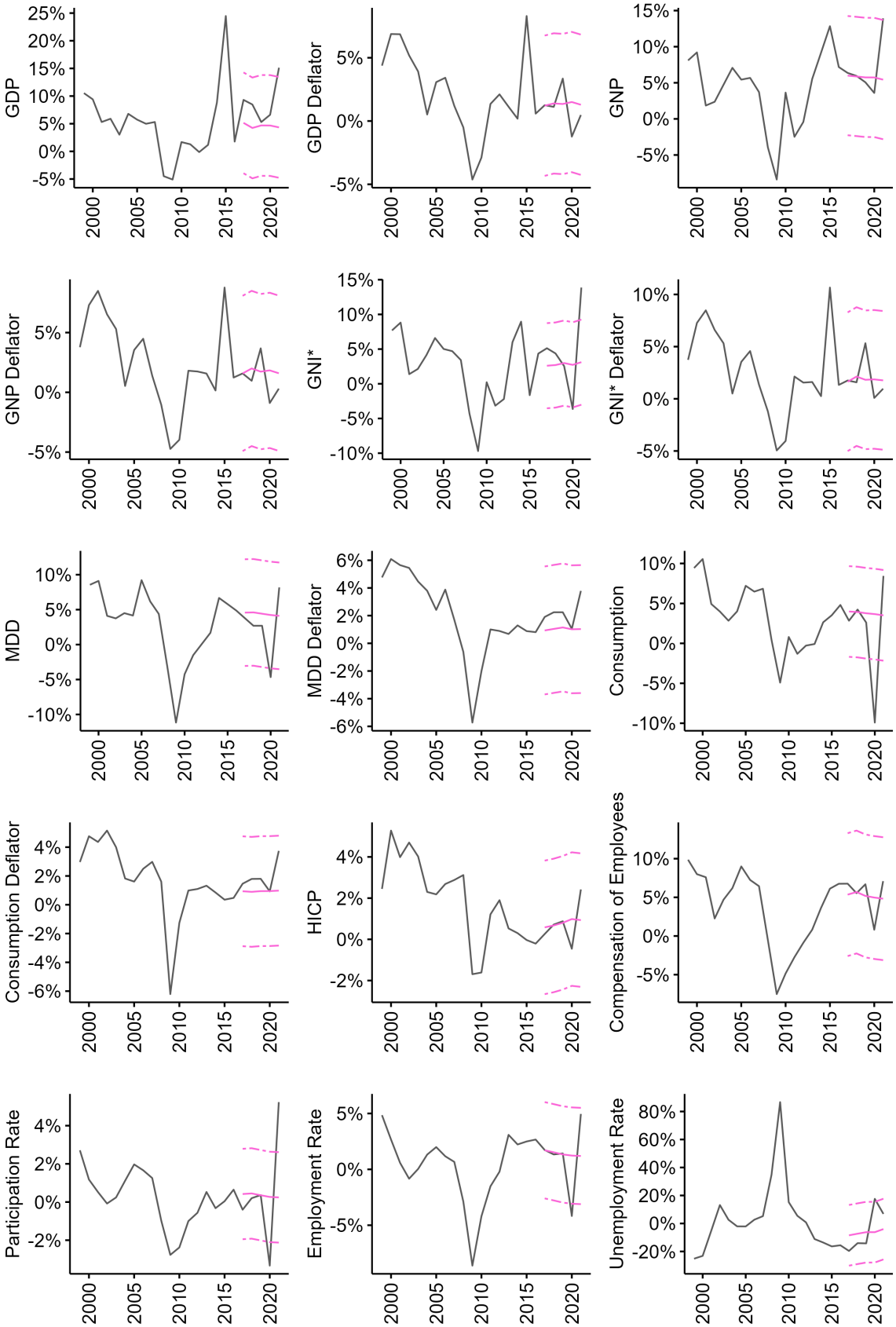

Figure 16: Forecast averages with the 90% prediction intervals for 2017–2021. The values on the y-axes are percentage changes.

References

Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin. 2010. "Large Bayesian Vector Auto Regressions." *Journal of Applied Econometrics* 25 (1): 71–92.

Biau, Olivier, and Angela D'Elia. 2012. *Euro Area GDP Forecasting Using Large Survey Datasets: A Random Forest Approach.* Euroindicators working papers EWP 2011/002. https://ec.europa.eu/eurostat/documents/4187653/5781955/EWP-2011-002-EN.PDF.pdf/56bf3a24-e037-4e01-afdd-d48dd359dd27?t=1414777488000.

Botha, Byron, Rulof Burger, Kevin Kotzé, Neil Rankin, and Daan Steenkamp. 2023. "Big Data Forecasting of South African Inflation." *Empirical Economics* 65 (1): 149–188.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32.

Bureau of Economic Analysis. 2023. "National GDP & Personal Income." U.S. Department of Commerce, October 27, 2023. https://www.bea.gov/itable/national-gdp-and-personal-income.

Carroll, Killian. 2020. *Forecasting Ireland's Macroeconomy: A Large Bayesian VAR Approach.* Working Paper No. 12, Irish Fiscal Advisory Council, September 30, 2020. https://www.fiscalcouncil.ie/wp-content/uploads/2020/09/Forecasting-Irelands-Macroeconomy.pdf.

Casey, Eddie, and Diarmaid Smyth. 2016. "Revisions to Macroeconomic Data: Ireland and the OECD." *Economic and Social Review* 47 (1): 33–68.

Central Statistics Office. 1998. "Quarterly National Household Survey: September - November 1997 First Results," May 26, 1998. https://www.cso.ie/en/media/csoie/releasespublications/documents/labourmarket/1997/qnhs_q31997.pdf.

———. 2023a. "CPM15: EU Harmonised Index of Consumer Prices." Accessed December 12, 2023. https://data.cso.ie/table/CPM15.

Central Statistics Office. 2023b. "ISA03: Annual Current Account." Accessed December 12, 2023. https://data.cso.ie/table/ISA03.

———. 2023c. "ISQ01: Quarterly Current Account." Accessed December 12, 2023. https://data.cso.ie/table/ISQ01.

———. 2023d. "Modified GNI." https://www.cso.ie/en/interactivezone/statisticsexplained/nationalaccountsexplained/modifiedgni/.

———. 2023e. "Monthly Unemployment and COVID-19 Adjusted Estimates August 2020 Technical Note." https://www.cso.ie/en/methods/labourmarket/monthlyunemployment/monthlyunemploymentandcovid-19adjustedestimatesaugust2020technicalnote/.

———. 2023f. "NA001: Modified Gross National Income at Current Market Prices." Accessed December 12, 2023. https://data.cso.ie/table/NA001.

———. 2023g. "NA002: Modified Gross National Income at Constant Market Prices." Accessed December 12, 2023. https://data.cso.ie/table/NA002.

———. 2023h. "NA006: Gross Value Added at Constant Basic Prices by Sector of Origin and Gross and Net National Income at Constant Market Prices," July 14, 2023. https://data.cso.ie/table/NA006.

———. 2023i. "NAQ03: Gross Domestic Product and Gross National Product." Accessed December 12, 2023. https://data.cso.ie/table/NAQ03.

———. 2023j. "NAQ04: Expenditure on Gross National Product." Accessed December 12, 2023. https://data.cso.ie/table/NAQ04.

———. 2023k. "NAQ05: Modified Total Domestic Demand and Components of Modified Gross Domestic Fixed Capital Formation." Accessed December 12, 2023. https://data.cso.ie/table/NAQ05.

———. 2023l. "QLF18: ILO Participation, Employment and Unemployment Characteristics." Accessed December 12, 2023. https://data.cso.ie/table/QLF18.

Central Statistics Office. 2023m. "Total Domestic Demand and Modified Total Domestic Demand." https://www.cso.ie/en/interactivezone/statisticsexplained/nationalaccountsexplained/totaldomesticdemandandmodifiedtotaldomesticdemand/.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 785–794. https://doi.org/10.1145/2939672.2939785.

Chu, Ba, and Shafiullah Qureshi. 2023. "Comparing Out-of-Sample Performance of Machine Learning Methods to Forecast U.S. GDP Growth." *Computational Economics* 62 (4): 1567–1609.

Conroy, Niall, and Eddie Casey. 2017. *Producing Short-Term Forecasts of the Irish Economy: A Suite of Models Approach.* Working Paper No. 4, Irish Fiscal Advisory Council, May. https://www.fiscalcouncil.ie/wp-content/uploads/2015/03/Producing-Short-Term-Forecasts-of-the-Irish-Economy.pdf.

———. 2019. "The Current Account, a Real-Time Signal of Economic Imbalances or 20/20 Hindsight?" *Economic and Social Review* 50 (1): 77–102.

Coulombe, Philippe Goulet, Maxime Leroux, Dalibor Stevanovic, and Stéphane Surprenant. 2021. "Macroeconomic Data Transformations Matter." *International Journal of Forecasting* 37 (4): 1338–1354.

Friedman, Jerome H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *The Annals of Statistics* 29 (5): 1189–1232.

Gamerman, Dani, and Ajax R. B. Moreira. 2002. "Bayesian Analysis of Econometric Time Series Models using Hybrid Integration Rules." *Communications in Statistics - Theory and Methods* 31 (1): 49–72.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian Data Analysis.* 3rd ed. Boca Raton, FL: CRC Press.

Genre, Véronique, Geoff Kenny, Aidan Meyler, and Allan Timmermann. 2013. "Combining Expert Forecasts: Can Anything Beat the Simple Average?" *International Journal of Forecasting* 29 (1): 108–121.

Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri. 2015. "Prior Selection for Vector Autoregressions." *The Review of Economics and Statistics* 97 (2): 436–451.

Gupta, Rangan, and Alain Kabundi. 2010. "Forecasting Macroeconomic Variables in a Small Open Economy: A Comparison between Small- and Large-Scale Models." *Journal of Forecasting* 29 (1–2): 168–185.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2017. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd ed. New York, NY: Springer.

Hewamalage, Hansika, Klaus Ackermann, and Christoph Bergmeir. 2023. "Forecast Evaluation for Data Scientists: Common Pitfalls and Best Practices." *Data Mining and Knowledge Discovery* 37:788–832.

Hou, Chenghan, Bao Nguyen, and Bo Zhang. 2023. "Real-Time Forecasting of the Australian Macroeconomy Using Flexible Bayesian VARs." *Journal of Forecasting* 42 (2): 418–451.

Hyndman, Rob J., and Yeasmin Khandakar. 2008. "Automatic Time Series Forecasting: The forecast Package for R." *Journal of Statistical Software* 27 (3): 1–22. https://doi.org/10.18637/jss.v027.i03.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2021. *An Introduction to Statistical Learning: With Applications in R.* 2nd ed. New York, NY: Springer.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree." In *Advances in Neural Information Processing Systems,* edited by I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

Kenny, Geoff, Aidan Meyler, and Terry Quinn. 1998. *Bayesian VAR Models for Forecasting Irish Inflation.* Central Bank of Ireland Technical Paper 4/RT/98, December. https://www.centralbank.ie/docs/default-source/publications/research-technical-papers/4rt98---bayesian-var-models-for-forecasting-irish-inflation-(kenny-meyler-and-quinn).pdf.

Koop, Gary M. 2013. "Forecasting with Medium and Large Bayesian VARs." *Journal of Applied Econometrics* 28 (2): 177–203.

Kuschnig, Nikolas, and Lukas Vashold. 2021. "BVAR: Bayesian Vector Autoregressions with Hierarchical Prior Selection in R." *Journal of Statistical Software* 100 (14): 1–27. https://doi.org/10.18637/jss.v100.i14.

Liaw, Andy. 2022. *randomForest: Breiman and Cutler's Random Forests for Classification and Regression.* CRAN - Package, May 23, 2022. https://cran.r-project.org/package=randomForest.

Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by randomForest." *R News* 2 (3): 18–22. https://CRAN.R-project.org/doc/Rnews/.

Litterman, Robert B. 1986. "Forecasting with Bayesian Vector Autoregressions: Five Years of Experience." *Journal of Business & Economic Statistics* 4 (1): 25–38.

Makridakis, Spyros, and Robert L. Winkler. 1983. "Averages of Forecasts: Some Empirical Results." *Management Science* 29 (9): 987–996.

Marcellino, Massimiliano, and Vasja Sivec. 2021. "Nowcasting GDP Growth in a Small Open Economy." *National Institute Economic Review* 256 (1): 127–161.

Microsoft Corporation. 2023. "Parameters." LightGBM's Documentation. Revision 2a085655. https://lightgbm.readthedocs.io/en/latest/Parameters.html.

Nowotarski, Jakub, and Rafał Weron. 2015. "Computing Electricity Spot Price Prediction Intervals Using Quantile Regression and Forecast Averaging." *Computational Statistics* 30 (3): 791–803.

OECD. 2016. *Irish GDP Up by 26.3% in 2015?* https://www.oecd.org/sdd/na/Irish-GDP-up-in-2015-OECD.pdf.

Qureshi, Shafiullah, Ba Chu, and Fanny S. Demers. 2021. *Forecasting Canadian GDP Growth with Machine Learning.* Carleton Economics Working Papers CEWP 21-05, Department of Economics, Carleton University, May 17, 2021. https://carleton.ca/economics/wp-content/uploads/cewp21-05.pdf.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Richardson, Adam, Thomas van Florenstein Mulder, and Tuğrul Vehbi. 2021. "Nowcasting GDP Using Machine-Learning Algorithms: A Real-Time Assessment." *International Journal of Forecasting* 37 (2): 941–948.

Saha, Sumit. 2023. "XGBoost vs LightGBM: How Are They Different." MLOps Blog. neptune.ai, August 8, 2023. https://neptune.ai/blog/xgboost-vs-lightgbm.

Stock, James H., and Mark W. Watson. 2004. "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting* 23 (6): 405–430.

"XGBoost." 2023. NVidia. Accessed February 27, 2023. https://www.nvidia.com/en-us/glossary/data-science/xgboost/.

XGBoost Developers. 2022. "XGBoost Parameters." XGBoost Documentation. Revision 0ce4372b. https://xgboost.readthedocs.io/en/latest/parameter.html.

Yoon, Jaehyun. 2021. "Forecasting of Real GDP Growth Using Machine Learning Models: Gradient Boosting and Random Forest Approach." *Computational Economics* 57 (1): 247–265.