

Which one predicts better? comparing different GDP Nowcasting methods with Brazillian Data.

Lucas Gabriel Martins de Oliveira
University of Brasilia

April 2, 2023

List of Figures

1	Neural Network with one hidden layer and three neurons	16
2	Austria model's performance, selected models: DFM + Shrinkage models/Hungary model's performance, selected models: DFM + Shrinkage models + NN.	20
3	Ireland model's performance, selected models: DFM + Shrinkage models/Hungary model's performance, selected models: DFM + Shrinkage models + NN/ Portugal model's performance, selected models: selected models: DFM + Shrinkage models + SVM.	21
4	Jagged Edge Database	25
5	MIDAS estimation-out of sample comparison	31
6	RMSE out of sample forecasts comparison	35
7	Selected Variables - LASSO	38
8	Out of sample forecast comparison - adaLASSO	39
9	out of sample comparison - ridge regression	40
10	Out of sample forecast comparison MIDAS/fully informed market expectations .	49
11	Information Criteria following BAI & NG (2002), source: dfms	49
12	Extracted Factor series, source: dfms	50
13	DFM (3,3) out of sample comparison	50
14	DFM (2,2) Factors extraction.	51
15	DFM (2,2) out of sample comparison.	51
16	Out of sample comparison dfm agg(3,3)	52
17	Out of sample comparison dfm agg(2,2)	52
18	RMSE/MAE comparison out of sample forecasts (Combination of classical models vs market)	53
19	Out of sample forecast comparison - LASSO 2	53
20	λ definition for ridge regression	54
21	Elastic Net regression plot	54
22	Elastic Net regression Out of sample comparison	55
23	Out of sample comparison combination of machine learning models and fully informed market expectations	55

24 RMSE/MAE comparison out of sample forecasts (combination of machine learning models vs market) 56

25 Out of sample forecast comparison 57

List of Tables

1	MIDAS Model Regression Results	30
2	Diebold-Mariano Test Results (MIDAS vs Naive)	31
3	Diebold-Mariano Test Results (MIDAS vs Fully Informed Market)	32
4	DFM (3,3) Bridge Equation Summary	33
5	DFM (2,2) Bridge Equation Summary	33
6	Significance DFM(3,3) with Mariano Murasawa Aggregation on it's factors . . .	34
7	Diebold-Mariano Test Results (DFM(2,2) vs Naive)	50
8	Diebold-Mariano Test Results (DFM(3,3) vs Naive)	50
9	Significance DFM(2,2) with Mariano Murasawa Aggregation on it's factors . . .	51

Contents

1	Introduction	9
1.1	Timeline - Early literature development (Before 2000's)	9
1.2	Timeline - Literature Development - Post 2000's	9
2	Nowcasting Literature Review	10
2.1	Current Understanding of Nowcasting Economic Activity	10
2.2	Regime Switching Models	11
2.3	Bridge Equations	12
2.3.1	Bridge Equation - Standard Formulation	12
2.4	MIDAS - Mixed Data Sampling Regressions	12
2.5	Factor Models	13
2.5.1	Dynamic Factor Model (DFM) Literature	13
2.5.2	Dynamic form	13
2.5.3	DFM Literature	14
2.6	DFM's Estimation	14
2.7	Understanding Dynamic Factor Models	15
2.8	Machine Learning methods	15
2.8.1	Machine Learning Methods - Shrinkage Models	15
2.8.2	Machine Learning Methods - Neural Networks	16
2.8.3	Machine Learning Methods - Bagging or Bootstrap Aggregating	17
2.8.4	Complete subset regression	18
2.8.5	Random Forests	18
2.8.6	Support Vector Machines	19
2.9	Note: Mariano & Murasawa's Aggregation use	20
2.10	A Proposal - IMF work	20
2.11	What we understand from the IMF's Nowcasting work	20
2.12	Target Factor Model and a Master's thesis	21
2.13	Extensions	21

3	Database	22
3.1	Nowcasting main problem: Jagged Edge Database	25
4	Nowcasting Strategy	25
4.1	Unbalanced Panel	26
4.2	Specific Case - MIDAS	26
5	Results Evaluation and COVID Nowcasting Exercise	27
5.1	Error Metrics and Further Accuracy Evaluation Methods	27
5.2	COVID Exercise - Classical vs Machine Learning Methods	28
5.2.1	Classical Models	28
5.2.2	Machine Learning	29
5.2.3	MIDAS	29
5.2.4	Diebold Mariano - MIDAS vs NAIVE Model	30
5.2.5	Dynamic Factor Models	32
5.2.6	Dynamic Factor Models with Factor Aggregation	34
5.3	Classical Models and a Selection Algorithm	34
5.3.1	Selection Algorithm	36
5.4	Machine Learning Methods	37
5.4.1	LASSO	37
5.5	adaLASSO	38
5.6	Ridge Regressions	40
5.7	Elastic Net	40
5.8	Neural Networks	41
5.9	Other models	41
5.10	Comparison ML vs market expectations	41
5.11	Machine Learning Methods using selection algorithm	42
6	Concluding remarks	43
7	Bibliographic references	44

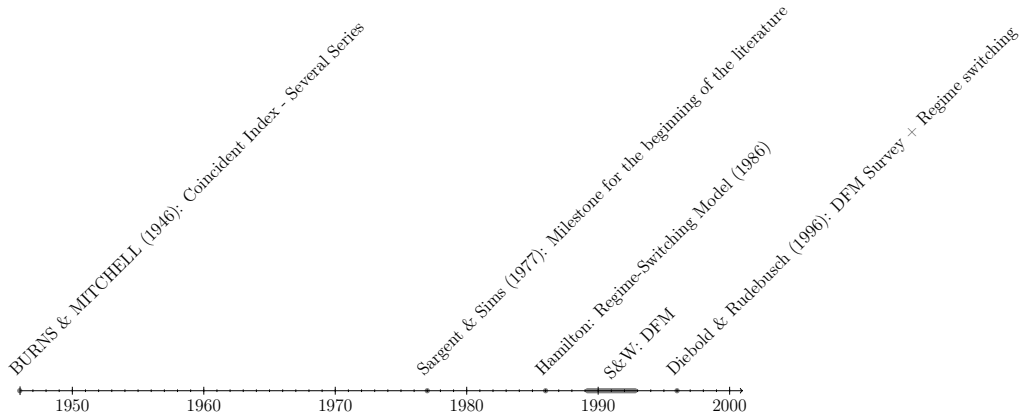
Abstract

This work has the primary objective of raising quantitative tools for assembling a scalable real-time GDP tracking in Brazil possibly later on for southern cone countries. In this work, we survey the literature since the first work on estimating business cycles and document the evolution of this literature until the insertion of machine learning methods used contemporaneously. Additionally, we perform backtesting exercises with an updated Brazilian database, estimate several candidate models for GDP nowcasting, implementing the division of classical models and machine learning models. Finally, we use the Diebold Mariano test to evaluate the forecasts of all models against a naive model and demonstrate that a combination of machine learning models based on the distance of forecasts to the average FOCUS expectations defeats the fully informed market expectations of the FOCUS survey, while the same is not possible only for selected classical nowcasting models, this result is also reached when we use the selection algorithm in Classical and Machine Learning models but the statistical significance of this result is smaller than the first. We also find that Mariano & Murasawa aggregation on factors in dynamic factor models did not improve the quality of the nowcasts.

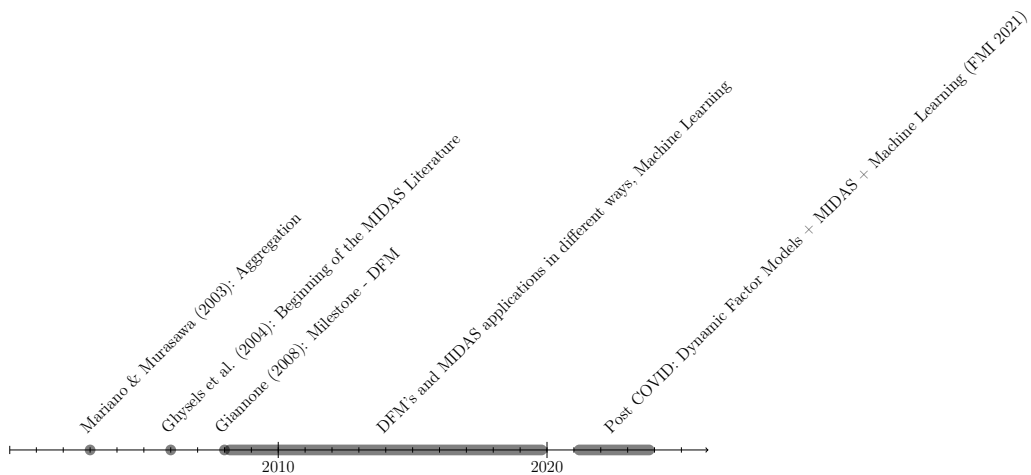
1 Introduction

The objective of this work is to develop a basic framework for the implementation of a GDP nowcasting strategy. Initially, our goal is to identify a scalable strategy that allows us to project the Brazilian GDP, with potential to be applicable in other countries in real time at any point during the current quarter. In this work we detail the survey of classical techniques (of real time GDP projection) and also of techniques usually known by market practitioners and experts as "machine learning methods". We understand that this classification may be confusing for some, since any technique that has as its theoretical basis an optimization problem can be classified as a "machine learning technique". However, we will explain later how the development of the nowcasting literature allows such taxonomy.

1.1 Timeline - Early literature development (Before 2000's)



1.2 Timeline - Literature Development - Post 2000's



2 Nowcasting Literature Review

In their work, Banbura, Giannone, Modugno Reichlin (2013) distinguish between partial models and full models. The full models would be those that include MF-VAR's (Mixed Frequency VAR's) and DFM's. This division is fortuitous because it allows us to classify the different econometric methods into these two large groups. We shall see, there are several methods used to address our problem, any form of classification is welcome.

2.1 Current Understanding of Nowcasting Economic Activity

Monitoring economic activity is no easy task. A country's economy features an immeasurable number of different activities; Obtaining a thermometer for the value produced in each of these activities is a problem that has intrigued economists of different strands since the conception of economic science. Although it is a Homeric task, the well-being of a country is impacted by the quality of its economic policies, which in turn depend on economic monitoring. To corroborate this statement, we can look at recent episodes of relevant economic crises, that is, even if it were not possible to avoid them, a more detailed economic monitoring of economic activity could have guided an early effort to mitigate their effects, saving jobs and directly affecting people's lives. In the absence of a bulwark that configures the objective measure of the active state of the national economy, methodologies for calculating the value added were developed and also, national accounting tautologies were synthesized to develop a variable that performs this task, the GDP. By evaluating the economy using GDP, economic authorities can reach a consensus on the productive state of a country, being able to assess its growth and formulate plans. In addition, it is possible to create metrics of social inequality and look at economic development in a more objective way. In Brazil, the IBGE publishes the GDP result in a quarterly period. Generally, this disclosure takes place on average nine weeks after the reporting period. It is a consensus among economists that some of the macroeconomic variables have "memory"; in fact, often the best prediction of a macroeconomic variable is its lagged value. Even taking this fact into account, it is not possible to make economic policies having as the main measure of economic activity a variable that presents a quarterly interval of disclosure, with the aggravating factor of being a measure lagged by approximately two months when disclosed. Within this context, the effort of public and private economic institutions to build indicators of economic activity and inflation that present less temporality makes sense. We will call this large class coincident indicators. Along these lines, the Central Bank of Brazil (BCB) seeks to produce a monthly indicator of GDP, the IBC-Br. The Institute of Applied Economic Research (IPEA), in turn, develops a coincident indicator, also monthly, for gross fixed capital formation. The indicator is made from accounting identities manifest in the Table of Resources and Uses (TRU). The idea behind developing these indicators is to create theory-driven monthly indicators that, once aggregated quarterly, correspond to GDP and gross fixed capital formation respectively. Despite the commendable effort, these indicators suffer from the same problem as GDP – built "from the bottom up", they leave us lacking data on economic

activity over a relatively long period. To deal with these uncertainties about the future, financial institutions pay economists to develop projections of relevant economic variables, mainly GDP, in different temporal “windows”. In the case of the Brazilian economy, these expectations can be consulted in the BCB’s Focus Report, released every Monday. Due to these problems, it is necessary to build a statistical apparatus, guided by the econometric theory in its “state of the art”, that plays the role of a real-time thermometer of the Brazilian economy in a relevant way. We must bear in mind that the variables that will assemble the set of information for each moment of time will be insufficient. As stated by GIANNONE et al. (2011), the term nowcasting refers to the future, past and present at the same time, having been first used in meteorology and brought to economics for the first time by GIANNONE et al. (2008). Having understood a little about the subject, let’s understand some statistical procedures historically used in nowcasting economic activity.

2.2 Regime Switching Models

Regime-switching models characterize data as falling into different, recurring “regimes” or “states”. Allow the characteristics of time series data, including means, variances, and model parameters to change across regimes. As we know our goal is to nowcast GDP, in order to do this we have to use monthly frequency variables to predict a quarterly one, hence, the regime switching literature has evolved to address mixed frequency variables, mixed frequency regime switching models. There are several contributions in this branch of the literature, some of them we highlight below:

- Hamilton (1989) - Seminal Work
- Diebold & Rudebusch (1996) - Survey + theoretical contribution
- Barsoum Stankiewicz (2013)
- Carstensen et al. (2017) (a work to detect Turning Points).
- Camacho et al. (2012)

Already in these works we have different approaches to deal with the jagged edge problem of the database, an important way of addressing this problem was proposed in Mariano & Murasawa (2003). There are two main models that we want to highlight in this branch, MS-DFM (Markov switching dynamic factor models) and MS-MIDAS (Markov switching mixed data sampling regressions), in the next sections we see in detail each one of these two models. But for now let’s just make a statement about the use of markov switching models in nowcasting. The question is: "Should we use regime switching models to nowcast GDP?". Our investigation suggests that this class of models is ideal to detect turning points in gdp performance, for now, this is nor the case, so the answer is no.

2.3 Bridge Equations

Remember, our goal is to predict a quarterly variable using monthly ones, the bridge equation *per se* would be a partial model, a simple regression to predict GDP, but as we shall see the bridge equations are used as a step in many estimation strategies of more complicated *full models*.

2.3.1 Bridge Equation - Standard Formulation

Take an activity indicator in quarterly frequency, given by its interannual quarter rate YQoQ (%), referred here as $y_t^q = y_t$, if we only use a monthly indicator on the basis of dependent variables (for simplicity), we have that the standard regression that defines a bridge equation is as follows:

$$y_t = \beta_0 + \lambda y_{t-1} + \beta(L)x_t^q + \epsilon_t$$

Where $\epsilon_t \sim WN$, $\beta(L)$ is a polynomial of lags, L is the lag operator. In our case, the database has a higher frequency than that of the dependent variable, an aggregation is performed on this database to then estimate the regression. Suppose we take the simple average of the three months of the quarter on a monthly basis and then do a regression with that data. So, the idea is that you do some kind of aggregation using the monthly data and then you do the following projection:

$$y_{T+1|T+\nu} = \hat{\beta}_0 + \hat{\lambda}y_T + \hat{\beta}(L)x_{T+1|T+\nu}^q + \epsilon_{T+1}$$

In this case, the subscript ν refers to intra-quarter months.

2.4 MIDAS - Mixed Data Sampling Regressions

The problem with Bridge equations models is that they require the estimation of a large number of parameters, to solve this, Ghysels, Santa Clara & Valkanov (2004) propose the MIDAS model. To control the curse of dimensionality, the MIDAS model places a restriction on the parameters of the polynomial lag structure using an aggregation function according to the following specification.

$$y_{t+h} = \beta_0 + \lambda y_t + \beta_1 B(L^{\frac{1}{3}}; \theta)x_{t+\omega}^m + \epsilon_{t+h}$$

Where $\omega = T_y - T_x$, the ends of the samples of y and x and represents the lead of the highest frequency indicator. Additionally, $B(L^{\frac{1}{3}}; \theta) = \sum_{k=0}^K b(k; \theta)L^{\frac{k}{3}}$ is the polynomial of lags and $b(k; \theta) = \frac{\exp(\theta_1 k + \theta_2 k^2)}{\sum_{j=0}^K \exp(\theta_1 j + \theta_2 j^2)}$, known as Almon's exponential polynomial. An important example of application it's Laine & Lindblad (2020) in a study by the Bank of Finland (Helsinki), used MIDAS to nowcast the Finnish GDP using financial variables with different frequencies.

2.5 Factor Models

Assume that a reasonable number of indicators are somehow related to economic conditions. According to the works of Stock & Watson. In this case more specifically Stock & Watson (1991), suppose that each variable can be written as the sum of two mutually uncorrelated stochastic components, the first component f_t is assumed to be common to all series in the base, the second component u_{it} refers to the idiosyncratic dynamics of each indicator in the base. Thus, the latent states of monthly GDP growth and monthly base indicators are assumed to follow the following dynamics. In this way, each series in this base can be translated as being composed of an element that measures the *business cycle overall* and the idiosyncratic component of each specific time series indicator.

$$\begin{aligned} y_t^m &= \beta_0 f_t + u_{it} \\ x_{it}^m &= \beta_i f_t + u_{it} \end{aligned}$$

Where $i \in I_N$ and N is the number of monthly indicators used in the analysis. In the first works of Stock & Watson related to factor models we found the specification above, the literature evolved and today we find dynamic factor models, whose specification will be detailed later.

2.5.1 Dynamic Factor Model (DFM) Literature

Again, in a factor model, the relationships between n variables x_1, x_2, \dots, x_n , for which T observations are available, are due to a small number of latent variables called factors, assumed in quantity $r < n$. The link between the observed variables and the factors is assumed to be linear. In this way, each observation x_{it} can be decomposed in the following way:

$$x_{it} = \mu_i + \lambda_i' f_t + e_{it}$$

Where μ_i is the average of x_i , λ_i is a vector $r \cdot 1$ and e_{it} and f_t are two uncorrelated processes. Thus, $\forall t \in I_T$ and $\forall i \in I_N$, x_{it} can be decomposed into a sum of two unobserved orthogonal components: The common component, $\chi_{it} = \lambda_i' f_t$ and the idiosyncratic component, $\epsilon_{it} = \mu_i + e_{it}$.

2.5.2 Dynamic form

In the literature we find different ways of representing the DFM's, their exact form, their dynamic form, their approximate form... Different modes of interpretation lead to different estimation methods according to Stock & Watson (2016). Below its dynamic form.

$$\begin{aligned} X_t &= \lambda(L) f_t + e_t \\ f_t &= \psi(L) f_{t-1} + \eta_t \end{aligned}$$

where X_t is a matrix of observed time series and f_t are the latent factors that will be estimated from this basis, $\lambda(L)$ and $\psi(L)$ are matrices $N \cdot q$ and $q \cdot q$ respectively, η_t is a vector $q \cdot 1$ of serially uncorrelated factor innovations, e_t are idiosyncratic shocks in the observed series. For estimation purposes it is necessary to specify a dynamics for e_t series... In this case: $e_{it} = \delta_i(L)e_{it-1} + \nu_{it}$, with ν_{it} being a white noise.

2.5.3 DFM Literature

The DFM literature is extensive, there are works focused on forecasting different macroeconomic variables, such as American and European GDP (Banbura et al. (2013)), Giannone et al. (2008) and Giannone et al. (2017), GDP of China (Yiu & Chow (2011)), GDP of Norway (Aastvei & Trovik (2010)) and Thorstud (2016), GDP of Turkey Modugno et al. (2016), American annual inflation Giannone et al. (2006) and forecast for global economic growth (Ferrara & Marsilli (2014)). Liu et al. (2011) develop five GDP nowcasting models using the DFM literature for ten Latin American countries. Dahlhaus et al. (2015) made projections of the GDP of the countries Brazil, Russia, India, China and Mexico (BRIC + M).

2.6 DFM's Estimation

Stock & Watson (2010) document the evolution of factor models according to their estimation methods, divided into three generations:

- i) First generation: time domain maximum likelihood via the kalman filter - In this procedure the EM method is used to estimate the parameters via Maximum Likelihood and then, given the parameters, the Kalman Filter can be used to compute the likelihood and estimate filtered values for F_t and hence f_t . Advantage: This method can handle data irregularities in a better way, Drawback: the number of parameters estimated is proportional to the number of series in the database, in this way, direct MLE is historically prohibitive.
- ii.1) Second generation: non-parametric averaging methods - the reason to consider cross-section averaging methods of X_t is that the idiosyncratic disturbances will converge to zero by the weak law of large numbers, in this case, only the linear combination of the factors will remain.
- ii.2) A specific method of Weighted average estimators is component analysis. Strength: the second generation methods can handle large datasets very well. Drawback: The method cannot deal very well with missing data and jagged edge datasets.
- iii) Third generation: Hybrid principal components and state space methods - The third generation of methods for estimating the factors merges the statistical efficiency of the state space approach with the robustness and convenience of the principal components

approach. This merged procedure occurs in two steps, which are described in more detail in works such as Giannone, Reichlin & Small (2008) and Doz, Giannone & Reichlin (2006).

- First step: The factors are estimated by principal components
- Second step: In the second step the estimated factors \widehat{F}_t are used to estimate the unknown parameters of the state space representation.

2.7 Understanding Dynamic Factor Models

The Hybrid Method of the third generation is largely used in the literature after the seminal work of Giannone et al. (2008), the consistency properties of this method is studied by Doz et al (2011). To better understand this class of models, one may want to take a look at the following subject, called state space models and it's formulations as we show below:

$$x_{nt} = \begin{bmatrix} \lambda_n & 0_{n \times r} & \dots & 0_{n \times r} & 0_{n \times r} \end{bmatrix} \begin{pmatrix} f_t \\ \dots \\ f_{t-p+2} \\ f_{t-p+1} \end{pmatrix} + \xi_{nt}$$

which is just the specification pointed out by Stock & Watson, a time series indicator can be decomposed into a factor component plus an idyosincratic therm. In space state form, this is called a measurement equation. Aditionally, the factors are supposed to follow a simple VAR(p) process:

$$F_t = AF_{t-1} + G\mu_t$$

2.8 Machine Learning methods

As we mentioned before, there are not many works that used Machine Learning for GDP projection before the 2000s, however, with all the discussion about the entry of machine learning methods in all sciences, machine learning methods are now being **largelly** used in nowcasting. Specially in post covid times, where the uncertainty about economic activity raised and economic activity forecasting was requested. In this area we highlight two works, the first work is from the IMF and the second is a master's dissertation from PUC-RJ.

2.8.1 Machine Learning Methods - Shrinkage Models

Shrinkage Models are a well-established alternative to factor models, the basic idea behind this modelling is to reduce the parameters that correspond to irrelevant variables towards zero. To quickly understand these models we can follow Maynard's specification (Maynard (2021)): Shrinkage models are based in the simple optimization problem:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{t=1}^T (y_t - x_t' \beta)^2 + \lambda \left[\alpha \sum_{j=1}^N \frac{|\beta_j|}{\omega_j} + (1 - \alpha) \sum_{j=1}^N \frac{|\beta_j|^2}{\omega_j} \right]$$

Setting the following vector of settings: $\psi = (\alpha, \lambda, \omega_j)$

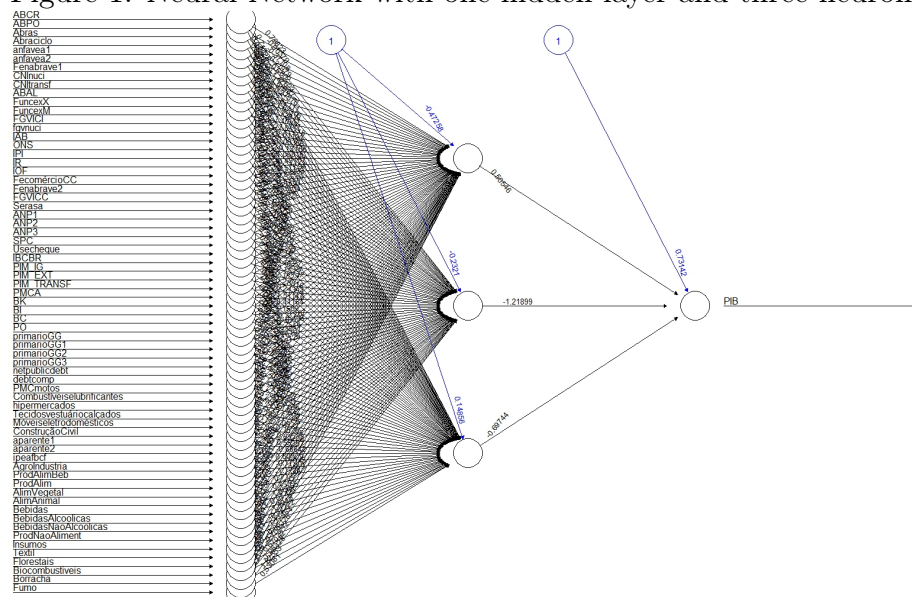
- Ridge Regression: $\alpha = 0$ and $\omega_j = 1 \forall j \in I_N$ where N is the number of parameters in the regression model.
- LASSO: $\alpha = 1$ and $\omega_j = 1 \forall j \in I_N$
- Elastic Net: $\alpha \in (0, 1)$ and $\omega_j = 1 \forall j \in I_N$

In this branch of the machine learning literature we also use the adaptive LASSO, or adaLASSO, this model is an evolution of LASSO techniques and it has the oracle properties for the choice of λ . The literature also stem away from shrinkage models, there's a variety of machine learning models that can be used for our task, the recent literature focuses in models like Random forests, complete subset regression, support vector machines, etc... For our purposes i'm gonna just step out of the detailed explanation of these models.

2.8.2 Machine Learning Methods - Neural Networks

A neural net is a method of artificial intelligence that allows computers to process data in a way that is somewhat linked to the way the human brain learns. It is a type of machine learning process more closely related to deep learning, the method uses neurons, interconnected in hidden layers, hence the analogy to the human brain. A neuron in a neural network takes

Figure 1: Neural Network with one hidden layer and three neurons



one or more inputs, applies a mathematical function to the inputs, and produces an output. The mathematical operation performed by each neuron is called the activation function. Let

x be the input vector of size n , w be the weight vector of size n , and b be the bias term. The output h of the neuron is given by the activation function f applied to the weighted sum of the inputs plus the bias term:

$$h = f(W^T x + b)$$

In a neural network with multiple layers, the output of one layer becomes the input to the next layer. Let $x^{(l)}$ denote the input to layer l , and $h^{(l)}$ denote the output of layer l for $l = 1, 2, \dots, L$. Then the output of the entire network is given by:

$$h^{(L)} = f^{(L)}(W^{(L)}h^{(L-1)} + b^{(L)})$$

where $W^{(L)}$ is the weight matrix for layer L , $b^{(L)}$ is the bias vector for layer L , and $f^{(L)}$ is the activation function for layer L . During training, the network's parameters are updated using the backpropagation algorithm. The loss function $L(y, \hat{y})$ measures the difference between the predicted output \hat{y} and the true output y , and the network's parameters are updated to minimize this loss function.

2.8.3 Machine Learning Methods - Bagging or Bootstrap Aggregating

Bagging is a type of ensemble method in machine learning algorithms that uses bootstrap and aggregating to form an ensembled model, where it's given a sample of data. Bootstrap samples are pulled, then a decision tree is formed on each of the bootstrapped subsamples. After each subsample decision tree has been formed; an algorithm is used to aggregate on the decision trees to form the most efficient predictors. The procedure used here is better described in Inoue, A., Kilian, L. (2008). This algorithm is largely used to reduce variance and prevent from overfitting. The basic idea behind bagging is to train multiple models on different subsets of the training data and then combine their predictions. The subsets are created by randomly sampling the training data with replacement, a process known as bootstrapping. This means that some instances may be sampled multiple times, while others may not be sampled at all.

Once the models have been trained, their predictions are combined using a simple averaging or voting scheme, depending on the problem being solved. This combination of models is known as an ensemble.

Mathematically, let $(x_1, y_1), \dots, (x_n, y_n)$ be the training dataset, where each x_i is a feature vector and y_i is the corresponding target value. We want to train T models $f_1(x), \dots, f_T(x)$ on different subsets of the training data.

To create each subset, we randomly sample n instances from the training dataset with replacement. This means that some instances may be sampled multiple times, while others may not be sampled at all. Let D_1, \dots, D_T be the T subsets of the training data.

Next, we train T models $f_1(x), \dots, f_T(x)$ on the T subsets of the training data, using the same learning algorithm for each model.

Finally, to make predictions for a new instance x , we combine the predictions of the T models using a simple averaging or voting scheme. For regression problems, we can average the predictions:

$$y(\hat{x}) = \frac{1}{T} \sum_{t=1}^T f_t(x)$$

Bagging can improve the performance of high-variance algorithms, such as decision trees, by reducing overfitting and improving generalization. By training multiple models on different subsets of the training data, bagging reduces the variance of the ensemble and improves its stability.

2.8.4 Complete subset regression

Complete subset regression is a technique presented in Elliott, Gargano and Timmermann (2013), the rationale behind this method is that all the subsets of a given sample data are selected and the forecasts computed on these subsamples are then compared.¹ Suppose we have a response variable Y and p predictor variables denoted by X_1, X_2, \dots, X_p . The complete subset regression problem involves fitting all possible linear regression models using different subsets of the predictor variables. There are 2^p possible subsets of predictor variables, which means there are 2^p possible models. The algorithm then fits a linear model on the selected subsample of regressors, using OLS. The best subset regression is the regression that minimizes information criterion AIC or BIC.

2.8.5 Random Forests

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression. As we mentioned before, the random forests algorithm of machine learning is closely related with that proposed by Inoue, A., Kilian, L. (2008) (bagging). The Random Forests methodology per se was first described in Breiman (2001), as it's explained perfectly in Maynard (2021). Suppose we have a training dataset with n observations and p predictor variables. To construct a random forest, we first generate B bootstrap samples of the training data, each containing n observations randomly sampled with replacement from the original data. For each bootstrap sample², we construct a decision tree using a random subset of m predictor variables selected at each split.

¹for more information in using these methods to nowcast economic variables see: <https://github.com/gabrielrvsc/HDeconometrics/blob/master/R/bagging.R>

²A bootstrap sample is a sample of data that is drawn from the original dataset by sampling with replacement. In other words, a bootstrap sample is obtained by randomly selecting observations from the original dataset, with replacement, until a sample of the desired size is obtained.

To grow a decision tree, we recursively split the data into smaller and smaller subsets based on the values of the predictor variables. At each split, we select the best variable and split point that maximizes the reduction in variance or the information gain.

After constructing all B decision trees, we obtain a prediction for a new observation by averaging the predictions from all the trees. For a regression problem, the prediction is simply the mean of the predicted values from each tree. For a classification problem, we use majority voting to determine the predicted class.

The random forest algorithm also provides estimates of variable importance based on the decrease in node impurity or the reduction in mean squared error resulting from splits involving a given variable. These estimates can be used to identify the most important predictor variables for the response variable. Random forests are known to be robust to overfitting and it's also great at dealing with large number of predictor variables.

2.8.6 Support Vector Machines

Support vector machines or simply SVM's is a supervised machine learning model that uses classification algorithms. In high dimensional data, what the SVM algorithm does is simply try to separate or classify a pre selected class of data using a hyperplane. For time series forecasting, the use of SVM is promising, as it's shown in Kyoung-Jae Kim (2003). One approach to using SVMs for time series prediction is to convert the time series into a supervised learning problem by framing it as a regression task. This involves selecting a window of past observations as input features, and using the next observation in the time series as the target variable.

For example, suppose we want to predict the next value in a time series based on the previous three values. We can frame this as a supervised learning problem by using the first three observations as input features and the fourth observation as the target variable. We can then use an SVM regression model to learn the relationship between the input features and the target variable.

Another approach to using SVMs for time series prediction is to treat the time series as a sequence of data points and use an SVM-based sequence prediction model. This involves training an SVM to predict the next value in the sequence based on the previous values, similar to recurrent neural networks (RNNs).

Overall, SVMs can be a useful tool for time series prediction tasks, particularly for data with short-term dependencies and where interpretability is important. However, selecting appropriate input features and tuning the model parameters can be challenging, and more complex models such as RNNs may be necessary for certain types of time series data.

2.9 Note: Mariano & Murasawa’s Aggregation use

MIDAS regression models and more explicitly the MF-VAR are based on the aggregation proposed by Mariano & Murasawa (2004), what they demonstrated in this work is that if the sample mean of the three months of a quarter can be well approximated by the geometric mean then the quarterly growth rate can be decomposed into moving averages of the monthly rates, in particular, the quarterly growth rate can be expressed:

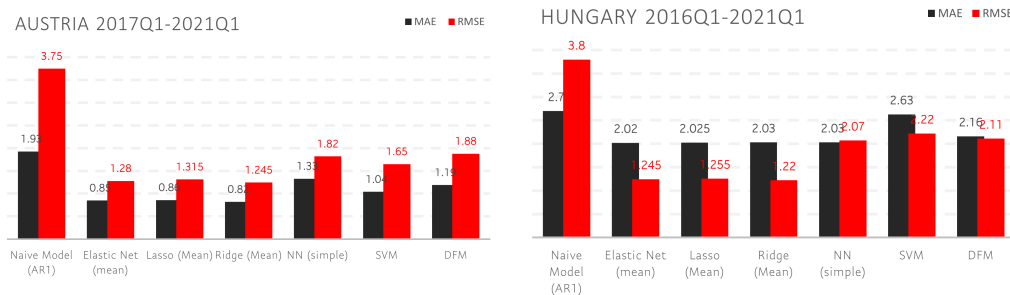
$$y_t^q = \frac{2}{3}y_t^m + \frac{1}{3}y_{t-1}^m + y_{t-2}^m + \frac{1}{3}y_{t-3}^m + \frac{2}{3}y_{t-4}^m$$

As we shall see, the dynamic factor model, with aggregation on it’s factors is one notable example of the aplicability of this relation.

2.10 A Proposal - IMF work

In recent work (Nowcasting GDP: a scalable approach using DFM, Machine Learning & Novel Data, Applied to European Economies), the IMF compared machine learning models with factor-calculated models using out of sample-sized error measurements. In another work (Nowcasting GDP with Machine Learning Models: Evidence from the US), a master’s thesis at PUC-RJ, Lucas Maynard also compared Machine Learning models with simulated factor models. Neither of these works used the dynamic factor model with the aggregation of Mariano & Murasawa in the factors, Valk et al. (2017) and Martins (2020) showed that this aggregation improves the GDP nowcast of dynamic factor models. Below we show graphs of the exercises made by the IMF in this important work, measuring out of sample works of different models.

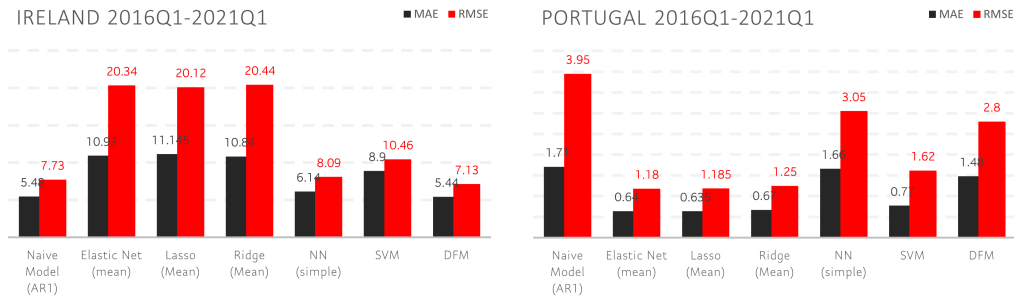
Figure 2: Austria model’s performance, selected models: DFM + Shrinkage models/Hungary model’s performance, selected models: DFM + Shrinkage models + NN.



2.11 What we understand from the IMF’s Nowcasting work

In some countries, for some reason, the data are not of sufficient quality for a satisfactory level of marginal monitoring. We also notice that there is some primacy in the projections made

Figure 3: Ireland model’s performance, selected models: DFM + Shrinkage models/Hungary model’s performance, selected models: DFM + Shrinkage models + NN/ Portugal model’s performance, selected models: selected models: DFM + Shrinkage models + SVM.



by models from the machine learning literature over dynamic factor models, with the exception of a few countries. We do not know the mechanism by which this happens, but neural networks, LASSO, ridge regressions, and elastic net seem to have greater primacy.

2.12 Target Factor Model and a Master’s thesis

According to Maynard (2021):

- Ridge Regressions together with DFM’s are a powerful tool for nowcasting.
- Target Factor model: The steps to make a target factor model, according to Maynard (2021) are:
 - step 1) Fit the following regression:

$$GDP_t = \alpha + \alpha_1 GDP_{t-1} + \alpha_2 S_{it} + \alpha_3 S_{it-1} + \alpha_4 S_{it-2} + \alpha_5 S_{it-4} + \epsilon_{it}$$

where S_{it} is each time series of the database.

- step 2) Discard every time series in the database that has a $p - value > 5\%$ in the regression aforementioned.
- step 3) Extract the factors using PCA and select the number of factors following information criteria (Bai & Ng (2002)).
- step 4) Fit a linear regression of GDP_t on the selected factors.

According to maynard (2021), this Target Factor Model outperforms the dynamic factor model in some vintages of his work.

2.13 Extensions

There are several other works associated with this literature, we left out some important points and branches of the real-time monitoring discussion, they are: BVAR’s, forecasting the

GDP GAP, early selection of regressors in the base of nowcasting, sparsity discussion (Giannone (2021)), etc.

3 Database

The main idea of our work is to create a scalable approach to implement nowcasting of economic activity using real time data, in this first moment we used a database of brazilian economic activity indicators of several sorts and frequencies, below we detail this feature of our work. We should note that the "Abrás" series was discontinued by the font, also, we have a small number of series to follow services dynamics, also, labor market indicators in Brazil are deprecated, the formal jobs series was discontinued and it has a recent break. The same problem happens with unemployment data and general income data because the oficial numbers are published in PNAD (IBGE) that only starts in 2012 (Services numbers are from 2014(IBGE)), in this exercise we prefer to let this series behind to preserve degrees of freedom in our estimations. One way of overcoming this issue would be by retropolating PNAD using old Monthly Employment Resarch (IBGE) and Services (PMS) by using old series as well. Another alternative would be to back interpolate these series using some clever empirical strategy. Below we detail the database used in our exercise, the nowcasting main algorithm may or may not have all the series here described, as we're constantly trying new configurations of regressors in order to better nowcast GDP.

Data ID	Series Explained	Frequency/type
PIM_IG	Industrial Production General Index	Monthly/Industry
PIM_EXT	Industrial Production Extractive Index	Monthly/Industry
PIM_TRANS	Industrial Production Transformation Index	Monthly/Industry
PIM_BK	Industrial Production Capital Goods	Monthly/Industry
PIM_BC	Industrial Production Consu Goods	Monthly/Industry
PIM_BCD	Industrial Production Durable Cons Goods	Monthly/Industry
PIM_BI	Industrial Production Intermediary Goods	Monthly/Industry
PIM_ITCC	Industrial Production Typical Build.	Monthly/Industry
PMC_IG_VA	Retail Sales - General enlarged volume	Monthly/Retail
PMC_IG_VR	Retail Sales - General enlarged volume	Monthly/Retail
CS_FGV	Services Confidence Indicator (FGV)	Monthly/Sentiment
IBC_BR	Economic Activity Monthly Indicator	Monthly/General
ABCR	Tolled Vehicle Flow Index	Monthly/leading
ABPO	Corrugated Paper Index	Monthly/leading
Abraciclo	Motorcycle License Plate Index	Monthly/leading
anfavea2	motor vehicle financed index 2	Monthly/leading
anfavea3	motor vehicle financed index 3	Monthly/leading
Fenabrave...19	License Plate Vehicles Index	Monthly/leading
CNI...20	Industry Indicator Index (hours worked)	Monthly/leading
ABAL	Produced Aluminum Index	Monthly/leading
funcex1	External sector data (Exports)	Monthly/leading
funcex2	External sector data (Imports)	Monthly/leading
funcex3	External sector data (Therms of trade)	Monthly/leading
funcex4	External sector data (Therms of trade)	Monthly/leading
fgv1	Confidence Consumer index 1	Monthly/Sentiment
fgv2	Idleness index	Monthly/Sentiment
IAB	Steel Production Index	Monthly/leading
ONS	Electric Charging Index	Qtly/General
rf1	Families Income 1	Qtly/General
rf2	Families Income 2	Qtly/General
rf3	Families Income 2	Qtly/General
Abras	Supermarket Sales Index	Qtly/General
CNI	Industrial Production Volume produced cni	Monthly/leading

Data ID	Series Explained	Frequency/type
Fecomércio	Sentiment Index business indicator	Monthly/leading
Fenabreve	Vehicles Distributed Index	Monthly/leading
Serasa	Credit Default Indicator Index 1	Monthly/leading
ANP1	Oil Produced Index 1	Monthly/leading
ANP2	Oil Produced Index 2	Monthly/leading
ANP3	Oil Produced Index 3	Monthly/leading
SPC	Credit Default Indicator Index 2	Monthly/leading
Usecheque	Payment Indicator (Checks)	Monthly/leading
IPEA_FBCF1	IPEA gross fixed capital formation 1	Monthly/Coincident
IPEA_FBCF2	IPEA gross fixed capital formation 2	Monthly/Coincident
IPEA_FBCF3	IPEA gross fixed capital formation 3	Monthly/Coincident
IPEA_FBCF4	IPEA gross fixed capital formation 4	Monthly/Coincident
LSPA1	Agricultural production 1	Monthly/Agric.
LSPA2	Agricultural production 2	Monthly/Agric
ICI - FGV	Confidence Industry indicator 1	Monthly/Sentiment
ISA - FGV	Actual Conjuncture Confidence Indicator	Monthly/Sentiment
est_termo	Therm Structure of interest rate	Monthly/Financial
IBOVESPA	number of operations in BR Stock Ex.	Monthly/Financial
IE - FGV	Expectations Indicator	Monthly/Sentiment
NUCI - FGV	Idleness Indicator	Monthly/coincident
PIB(1)	First Difference of GDP	Qtly/General
PIB(2)	Second Difference of GDP	Qtly/General
RNB	Gross Domestic Income	Qtly/General
CONS_CNT	Families Consumption	Qtly/General
FBKF_CNT	Gross Fixed Capital Formation	Monthly/Industry
cambio_real	Dolar/BRL Currency	Monthly/International
usa_atividade	USA Economic Activity	Qtly/International
gtrends	google searches - unemployment	Monthly/General
cdi_eua	American inter-bank deposit rate	Monthly/Financial
jobs_survey	American employment survey	Monthly/Financial

3.1 Nowcasting main problem: Jagged Edge Database

One of the main problems in Nowcasting Economic Activity is the fact that in general, along the dates defining the *vintages*, the database becomes unbalanced, that is, as if it were not enough for the variable of interest to have a frequency different from that in which the data are observed, the nature of the intra-quarter releases causes the database to present the characteristic that the literature calls *jagged-edge*. Due to the large number of models estimated,



Figure 4: Jagged Edge Database

we will not present a unique solution to this problem, for the simple reason that each model requires a different solution. For example, in the case of the MIDAS model the balancing of the base must follow a "frequency alignment", whereas in the case of machine learning algorithms and dynamic factor models the balancing can be done in the same way. In general, to set the database in the real-time exercises using the entire surveyed database, when necessary, we will estimate simple autoregressive models to fill in the *missing values*, thus dealing with the Jagged-Edge problem. nowcasting

4 Nowcasting Strategy

A GDP Nowcasting Strategy is a method used to predict the current level of a country's Gross Domestic Product (GDP) in real-time or near-real-time using the most recent available data. In this work the nowcasting strategy is simple and must be adapted to the fact that we estimate a final projection that is a combination of several models, namely, classical and machine learning models as we have already stated.

4.1 Unbalanced Panel

As we put it earlier, updating the time series matrix in real time results in a razor tooth or "jagged edge" database, the strategy used to deal with this problem requires a panel balancing method, some works use as a balancing method the simple naive forecast (from a well calibrated autoregressive model) applied on each column of the observations matrix. In our case, we use the following procedure to deal with the unbalanced basis problem.

- **Step 1:** Select the column of the updated series (or series) on each day of nowcasting.
- **Step 2:** Using the date of the update of these series as the projection horizon for the most lagged series select the machine learning algorithm that performed best (according to the metrics used (RMSE, MAE)) in the past vintages.
- **Step 3:** Make out-of-sample projections up to the selected horizon for each of the series with missing values.
- **Step 4:** Run models (training set - balanced panel) and compute final one step ahead projection using balanced panel, computing the nowcast for the current quarter.

Following these steps you should have a balanced time series panel, the only missing values should be the values of the GDP to nowcast at the quarter of concern. The mechanism from which we derive the final nowcast using all the individual nowcasts of ML and Classical models will be detailed later on.

4.2 Specific Case - MIDAS

The only model used in our strategy that does not fit in this procedure is the MIDAS model, because the structure of the model requires a separate frequency alignment. Therefore for the MIDAS model the Nowcasting strategy is done separately, in this case we emphasize that the number of rows of the matrix must be equal to three times the number of observations of the lower frequency series (in this case GDP), in some vintages this rule will not be respected, to estimate MIDAS we will then need to use some of the machine learning models to make the frequency alignment satisfied. In this topic we follow Ghysels et al. (2016), addressing the frequency alignment issue by aggregating the high-frequency data to match the low-frequency data. This is done by using weighted averages of the high-frequency data, where the weights are estimated as part of the model. Specifically, the authors use polynomial weighting schemes, such as the Almon lag polynomial, the Beta lag polynomial, and the Exponential Almon lag polynomial.

5 Results Evaluation and COVID Nowcasting Exercise

To evaluate the accuracy of the models used in our nowcasting algorithm, we formulated the following Backtesting Strategy:

- 1) Separate the database into training set and test set, as we noticed, a general interest of the scientific community was to interpret the behavior of these models after the COVID-19 event, knowing this, we used a pertinent first time window for the backtesting, for the first nowcast then, the training set is the entire pre covid period, i.e., 1996 (year of the first YQoQ(%) release of Brazilian GDP) until March 2020, thus we will have the out-of-sample estimations after the COVID-19 period. That is, we put ourselves in a period in time that from which all data for Q2 2020 GDP was available except the own GDP (Hence the name fully informed expectations). After that, we repeated this procedure by re-estimating the models up to August 2022 and computing the pertinent metrics.
- 2) After separating the base into training set and test set, we estimate all models as if we were on the eve of the covid 19 events, i.e., the vintage used in the exercise done here is that of February 28, 2020. This procedure is done only to evaluate the accuracy of the models, the main code of work will allow the selection of any vintage anywhere in time in order to get backtesting out-of-sample results.
- 3) Calculation of selected error metrics (RMSE, MAE) to evaluate the predictive potential of these predictions compared to a naive model.
- 4) Diebold-Mariano against naive models and FOCUS projections (in that selected vintage).

5.1 Error Metrics and Further Accuracy Evaluation Methods

In this section we explain how we use the diebold mariano test in order to compare out of sample projections at the selected vintage. Frequently in practical applications, multiple time series models exist that can be utilized for predicting a specific variable of concern, we generally have: Let I_N be a set of integers until the number N indexed. Here, N is the length of a forecast vector.

- Observed Data: $\{y_t; t \in I_N\}$
- Competing Forecasts: $\{\hat{y}_{1t}, \hat{y}_{2t}\}; t \in I_N$

We then Define the forecast errors as:

$$e_{1t} = \hat{y}_{1t} - y_t \text{ and } e_{2t} = \hat{y}_{2t} - y_t$$

Define a Loss function associated with each forecast, generally it's assumed to be a function $h(e_{1,t})$ and $h(e_{2,t})$. The function $h_{it}()$ has the following properties:

- 1) $h = 0$ if the forecast is perfectly accurate.
- 2) the function $h(\cdot)$ can't be negative.
- 3) As errors become greater the function grows in module.

The function $h(\cdot)$ generally it's a modular function or a squared function, when it's appropriate for this function to be symmetric. We then define the loss differential:

$$d_t = h(e_{1t}) - h(e_{2t})$$

We then say that the two forecasts have equal accuracy \iff the differential above mentioned satisfies:

$$E[d_t] = 0 \quad \forall t \in I_N$$

The Diebold-Mariano test's Null Hypothesis is then defined:

$$H_0 : E[d_t] = 0 \quad \forall t \in I_N$$

Asymptotic properties and more details about the test can be seen in Diebold (2015).

5.2 COVID Exercise - Classical vs Machine Learning Methods

This Exercise was made following the next steps:

- 1) Make pseudo real-time databases in final quarters of each year, in this way we set different training samples and one step ahead test samples.
- 2) Estimation of the models in training sample, from 2020 and so on until second quarter of 2022, the exercise was made in final quarter vintages of each year.
- 3) Evaluations of out of sample projections in test sample (Data Available After Second Quarter of 2020 and so on until second quarter of 2022).

Using this procedure we can now divide the models falling into two categories, Classical and ML models.

5.2.1 Classical Models

The exercises done are in the sense of estimating models traditionally classified as "machine learning" methods and methods that, in some classifications, may also fall under this nickname but are estimated in the classical nowcasting literature, that is, we will classify these methods as "classical models". In our classification these models are classical in the sense that they have been part of the nowcasting literature, both for inflation and GDP, for much longer than the techniques under the "Machine Learning" umbrella. In the classification of classical models, we estimate:

- 1) MIDAS (Mixed Data Sampling Regression)
- 2) Dynamic Factor Model (two step)
- 3) Dynamic Factor Model (two step - Mariano & Murasawa factor aggregation)

5.2.2 Machine Learning

Under this classification we estimate:

- LASSO
- adaLASSO
- Ridge
- Elastic Net
- Neural Networks
- Bagging & Bootstrapping methods
- Complete Subset Regression
- Random Forests
- Support Vector Machines

5.2.3 MIDAS

To estimate the MIDAS model we follow Ghysels et al. (2016), this work highlights the estimation of the mixed frequency regressions perfectly, as we put before, the only challenge here is to adjust the database according to the required "frequency alignment". It is interesting to note that the real-time data update procedure is much simpler for this model than all the other estimated models. As we have pointed out, we used the base described above and the result for the projections was not satisfactory, for the MIDAS model we then decided to test a more sparse database in the follow-up and updating of the Brazilian GDP, in this case we opted for the following regression:

$$PIB_t^Q = \alpha + \beta_1 PIB_{t-1}^Q + \beta_2 PIM_t^M + \beta_3 PMC_t^M + \beta_4 IBCBR_t^M + \epsilon_t$$

it is clear that this would be a regression with simple time series in the way it was put, however it should be understood that in all variables within the model it was used the aggregation with the Almon polynomial explained earlier, this aggregation allows to map variables of different frequencies in the same simple regression, the estimation of the MIDAS model is done by non-linear least squares. The superscripts Q and M mean that the series have quarterly and

Table 1: MIDAS Model Regression Results

Variable	Estimate	Standard Error	t-statistic
Intercept	0.5485***	0.15882	3.468
lag(GDP)	0.1211***	0.0324	3.725
Industrial Prod. (IBGE)	-0.174643***	0.02559	-6.824
Retail Sales (IBGE)	0.0686**	0.0231	2.992
IBC Br (Central Bank)	59.1823***	2.9312	20.190

Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' ' ,

monthly frequencies respectively. Table 1 shows the summary of our MIDAS regression results.

Regarding the choice of regressors, the industrial production measured by the monthly industrial survey is remarkably one of the variables that most explains the Brazilian GDP, being on a par with the monthly trade survey. We did not include the SMP (Services) because its sample size is too small, we did not find regressors that could explain with any quality the behavior of the agricultural and external sectors, in further tests we want to include sentiment indicators for this sector. To supply this need we have included the IBC-BR, which in its composition presents in a way a little of these two indicators. Future research may show how to select the regressors in a more satisfactory way for the composition of the MIDAS regression. Below is the out-of-sample comparison of the GDP projection made in the MIDAS model placed above and the actual GDP observed. The MIDAS model is very different from the other models estimated here, it is obtained with a separate database, the variance of its nowcasting will also be smaller, because we can only perform nowcasting with the MIDAS model after the update of one of these few series from its base. Nevertheless, we consider the MIDAS model a good candidate in our final nowcasting, even if to combine with the nowcast of other models in some vintages we have to take only an average while the hard data in its base is not updated.

Notice that the data observed right after the hecatomb generated by the covid crisis are of poor quality, in all the estimated models we observe a difficulty in capturing the real movement of GDP in this period, some perform better than others.

5.2.4 Diebold Mariano - MIDAS vs NAIVE Model

We then evaluate projections $h = 10$ steps ahead using the Diebold Mariano test using the standard error computation method to compute the loss function. For the framework of the test, we will use the null hypothesis that the forecast of method 1 (MIDAS) is less accurate than that of method 2 (Well specified ARIMA). Thus, if the p-value is smaller than a certain predetermined value, we can vouch for the quality of the method 1 projection. In this case, we attest with a significance of 1% that the MIDAS model projections beat a naive model. We can say that an educated guess regarding the behavior of the GDP can very easily beat the naive model, the real test to decide whether the use of a model adds or does not add value in

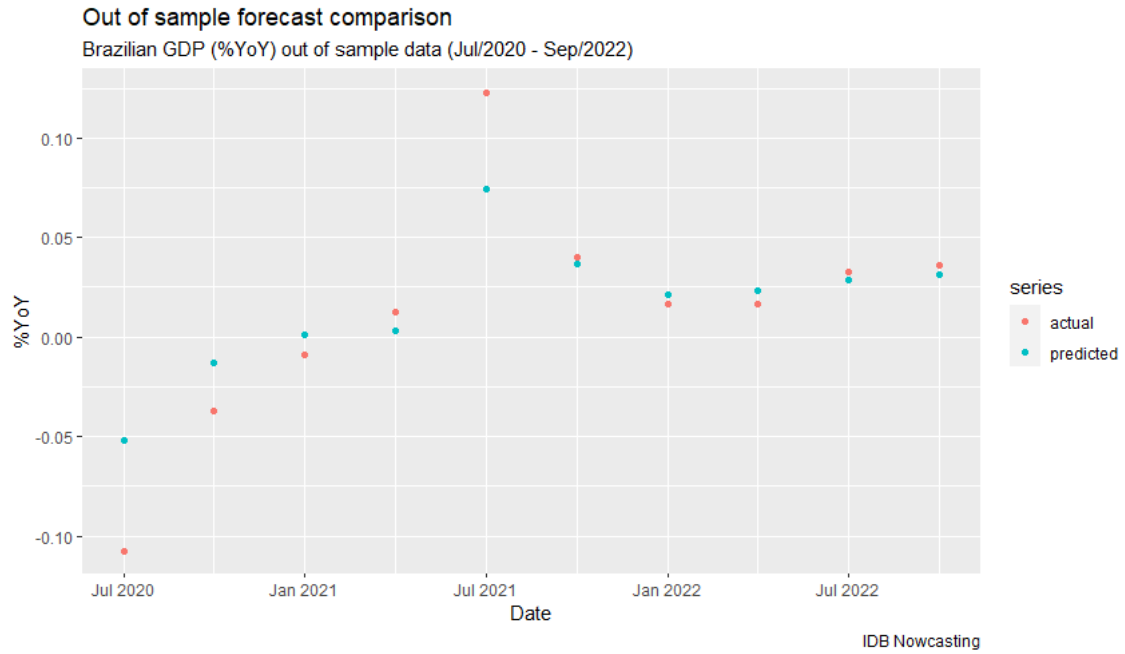


Figure 5: MIDAS estimation-out of sample comparison

Table 2: Diebold-Mariano Test Results (MIDAS vs Naive)

Forecast Horizon	Test Statistics	
	DM	p-value
10	-3.5478	0.003119

the projection of a macroeconomic variable is its comparison with the market projections, in Brazil these projections are carried out by the FOCUS survey. However, I would like to make a parenthesis. In many works, there is a mistaken notion of what market expectations are, that is, the time at which they are collected makes a lot of difference. In the case of GDP, the relevant data for the following quarter all come out around the second month after the reference quarter, so using expectations collected before that underestimates the power of market expectations. To emphasize that we deal with this problem by putting expectations on an equal base with our projections, we will call these expectations collected two months after the reference period fully informed expectations. Below is the result of the DM test comparing the MIDAS model and the fully informed expectations. As we can see in **Figure 6**, the MIDAS model with only

Table 3: Diebold-Mariano Test Results (MIDAS vs Fully Informed Market)

Forecast Horizon	Test Statistics	
	DM	p-value
10	3.8818	0.9981

the default configuration does not present a greater predictive power than that presented by the fully informed market. Although the distance between the two forecasts is much smaller in the post-covid crisis periods, it is notable that the market performs better than our model because of the poor performance in the covid crisis where the indicators stopped making sense at some level.

5.2.5 Dynamic Factor Models

In order to estimate Dynamic Factor Models, first, using the full base described above, we selected the number of factors via information criteria. As we can see, the information criteria point to three, two, and seven factors to extract, to preserve degrees of freedom in our estimation we selected two factors only. Remember that estimating dynamic factor models involves specifying a VAR, whose degrees of freedom are lost depending on the number of independent variables included and the number of lags required in the model structure. For example, take a VAR(3) on a database with 3 variables, then each VAR equation will have $3 \cdot 3 + 1 = 10$ parameters including the constant, this number is acceptable with a small series like GDP, however if we consider 7 extracted factors, as pointed out by the third criterion, we will have in a VAR(7) $7 \cdot 7 + 1 = 22$ parameters, if we understand that the GDP series within the sample considered in this estimation presents a little more than 85 observations we realize the size of the problem³. First we estimate Dynamic factor models using two factors and then three factors extracted, using the two steps method, and then, we estimate the same models using two steps aggregating the factors extracted by using Mariano Murasawa aggregation. Figure 6 shows information criteria for dynamic factor model with three dynamic and static factors extracted. Using these criteria, we estimated factor models with 3 lags and 3 dynamic

³More at: <https://kevinkotze.github.io/ts-7-var/>

factors extracted and also with 2 lags and 2 dynamic factors only. As we can notice at a first glance, the dynamic factor model with 3 extracted factors and 3 lags did not perform so well in our vintages, the same behavior observed in the previous model occurs, that is, right after the pandemic the model performs poorly and then becomes better, however it is notable that this estimation does not beats the fully informed expectations. On the statistical significance side, our bridge equation presented stastically significant results for this estimation. However, the significance remains at optimal level as we can see in the table below.

Table 4: DFM (3,3) Bridge Equation Summary

Variable	Estimate	Standard Error	t-statistic
Intercept	0.02227***	0.00084	26.6
Factor 1	0.00510***	0.00016	30.5
Factor 2	0.00185***	0.00038	4.8
Factor 3	-0.00342**	0.00044	-7.7

Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '

With this, we set out to estimate the DFM(2,2) model as the information criteria pointed out. At a first glance there doesn't seem to be a very satisfactory performance either. However, we need statistical evaluation to determine whether the models perform well or not.

Table 5: DFM (2,2) Bridge Equation Summary

Variable	Estimate	Standard Error	t-statistic
Intercept	0.02227***	0.00091	24.3
Factor 1	0.00510***	0.00016	30.5
Factor 2	0.00508***	0.00018	28.1

Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '

After estimating these models we can compare then with the Naive model using Mariano Murasawa test, below we present these results for both. The test rejects the Null hypothesis that DFM's are less accurate than the naive model in forecasting the Year - Quarterly GDP Growth at 1% of significance for both cases.

5.2.6 Dynamic Factor Models with Factor Aggregation

The classic literature of dynamic factor models has used Mariano and Murasawa’s aggregation on factors to improve the quality of nowcasting models ⁴, here we use the same method in order to evaluate these nowcasts compared. We used the Mariano & Murasawa aggregation as seen below:

$$y_t^q = \frac{2}{3}y_t^m + \frac{1}{3}y_{t-1}^m + y_{t-2}^m + \frac{1}{3}y_{t-3}^m + \frac{2}{3}y_{t-4}^m$$

Where y is the variable of interest in quarters and the superscript m tells us that we are approximating it monthly.

Table 6: Significance DFM(3,3) with Mariano Murasawa Aggregation on it’s factors

Variable	Estimate	Standard Error	t-statistic
Intercept	0.0219***	0.000816	25.2
Factor 1 agg	0.00062***	0.00002	27.8
Factor 2 agg	0.00125**	0.00004	2.7
Factor 3 agg	0.00008*	0.00008	-2.3

Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' ' ,

5.3 Classical Models and a Selection Algorithm

In this section of our work we present a forecast selection algorithm based on FOCUS expectations, applying this procedure in classic models we show that their forecast ability is enhanced. We computed the out of sample RMSE for classical models in our previous backtesting exercise, below we show a comparison of these models with **fully informed market expectations**, as we can see, no model alone can beat the market at the eve of release of the GDP, nevertheless, it’s important to note that this result might not hold under the test of different vintages (i.e. if we get market expectations when the data available for GDP it’s incomplete it might be the case that the forecast power of these expectations it’s not the case). Diebold-Mariano tests for the dynamic factor models with Mariano & Murasawa aggregation revealed that Classic models are more accurate than naive ones both at 1% significance, however there is no evidence that the projections from these models are more accurate than those made by the perfectly informed market. A slightly concerning result for estimations of classical models jointly, i.e., MIDAS, DFM(2,2), DFM(3,3), DFM(2,2) agg, DFM(3,3) agg is that there is no linear combination among the models’ projections that can outperform the perfectly informed market’s projections. One feature we note in these estimations is the fact that their variance and the ability of these models to capture nonlinearities is not that large, going contrary to these results, we’ll demonstrate that machine learning literature presents an answer to this problem. A qualification here is necessary: **We have not** estimated all the models in the classical literature, an extension of this work should be one that goes towards encompassing the

⁴Stock & Watson, 2016

Bayesian literature as put forth earlier, in addition to these, it would be interesting to estimate change of state (Markov Switching) models trying to capture nonlinearities that appear to have been the problem of the classical models in this exercise. Of course there are vintages in which some of the classical models presented a better projection than all the others, however when we take the overall average this does not seem to be the case.

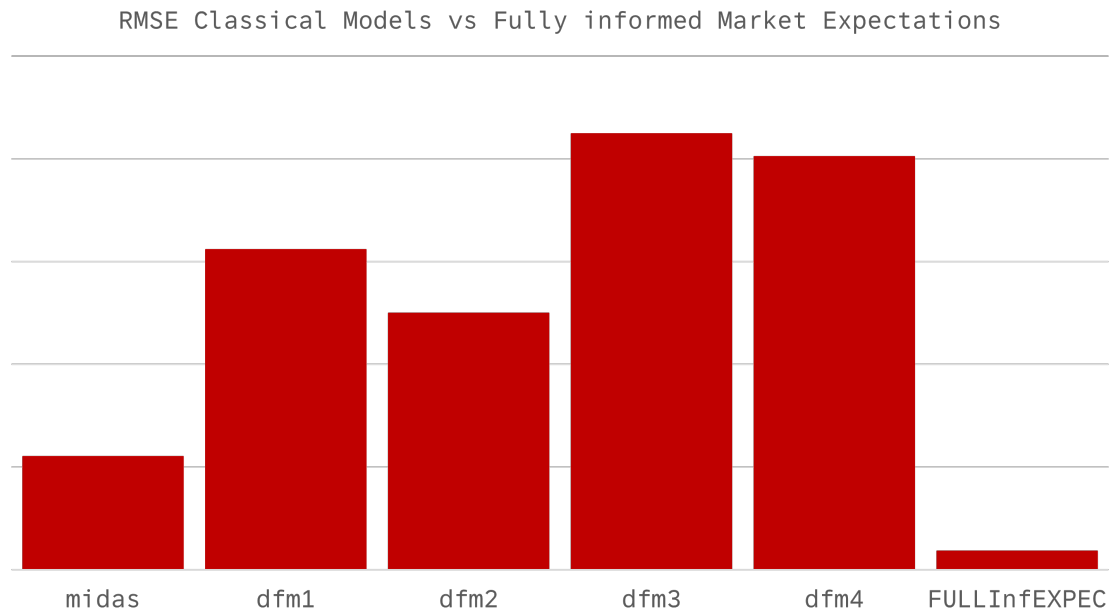


Figure 6: RMSE out of sample forecasts comparison

5.3.1 Selection Algorithm

Since we have the expectations beforehand, we can encompass them in the projection of the models in an algorithm, that is, we can select the models that are closest (above or below) the perfectly informed market projection, so we can somehow improve the projection in an automatic way, below we perform this experiment with the classical models and compare them with the perfectly informed projection. The selection algorithm takes into account a metric $d(\mathcal{M}, \mathcal{E})$ defined as follows: Let \mathcal{M}, \mathcal{E} be respectively the set of model projections taken into account and perfectly informed market expectations respectively. Additionally define \mathcal{FF} as the set of models selected at the final forecast. We detail the selection algorithm below:

- 1) Define a real value for α
- 2) Calculate $d(\mathcal{M}_i, \mathcal{E}) = |\mathcal{M}_i - \mathcal{E}|, \forall i \in I_j$, where I_j is the set of natural numbers that index all models encompassed in this algorithm.
- 3) If the signal of the forecast is different from the signal of the expectation discard that model from \mathcal{FF} in this step.
- 4) If $d(\mathcal{M}_i, \mathcal{E}) > \alpha$, discard \mathcal{M}_i from \mathcal{FF} , else, include \mathcal{M}_i in \mathcal{FF} .
- 5) If no model succeed in the step 3 or 4, select only the expectation.
- 6) The final forecast is: $\hat{\mathcal{F}} = \frac{\sum_{i=1}^k \mathcal{M}_i}{k}$, where k is the number of models selected.

Performing this procedure with $\alpha = 2\%$ for the classical models we get the following result:

- Q1) MIDAS
- Q2) MIDAS, DFM4
- Q3) MIDAS, DFM4
- Q4) MIDAS, DFM2, DFM4
- Q5) MIDAS
- Q6) MIDAS, DFM1, DFM2
- Q7) MIDAS, DFM1, DFM2, DFM3, DFM4
- Q8) MIDAS, DFM1, DFM2, DFM3, DFM4
- Q9) MIDAS, DFM1, DFM2, DFM3, DFM4
- Q10) MIDAS, DFM2, DFM4

This procedure gives us the following results:

Measure Model	FULLInfEXPEC	Final Forecast
RMSE	0.95	1.19
MAE	0.88	1.01

As we can see, the selection algorithm greatly improves the final projection, given that the out-of-sample errors were much larger than these, however this is not enough to beat the fully informed market expectations. Perhaps if we had a few more models in the "classical models" wing the results would be a little better, however, what is worrying about these estimations is the low variance between the configurations.

5.4 Machine Learning Methods

For machine learning methods we will skip some details regarding estimation, the packages used in R for estimating these models are in general very standard.

5.4.1 LASSO

The LASSO (Least Absolute Shrinkage and Selection Operator) is a popular shrinkage method for selecting relevant economic variables in forecasting and nowcasting models. It combines regularization with variable selection, resulting in a more interpretable and parsimonious model.

LASSO works by adding a penalty term to the ordinary least squares (OLS) objective function, which forces the sum of the absolute values of the coefficients to be less than a predetermined value. The penalty term shrinks the coefficients of less important variables towards zero and sets the coefficients of irrelevant variables to exactly zero. This process effectively removes them from the model, leading to automatic variable selection.

In the context of nowcasting economic variables, LASSO can be particularly useful due to the following reasons:

- High-dimensional data: Economic datasets often contain a large number of potential predictors. LASSO helps to navigate this high-dimensionality by identifying a subset of relevant variables, leading to more accurate and interpretable models.
- Multicollinearity: Economic variables often exhibit multicollinearity, with high correlations between predictors. LASSO helps to mitigate multicollinearity issues by generating stable and unique coefficient estimates.
- Overfitting: By shrinking the coefficients of less important variables, LASSO prevents overfitting, leading to better out-of-sample forecasting performance. here we perform a

LASSO estimation and select, from the nowcasting balanced panel the following series: The variable that presents the greatest weight in the GDP projection according to the

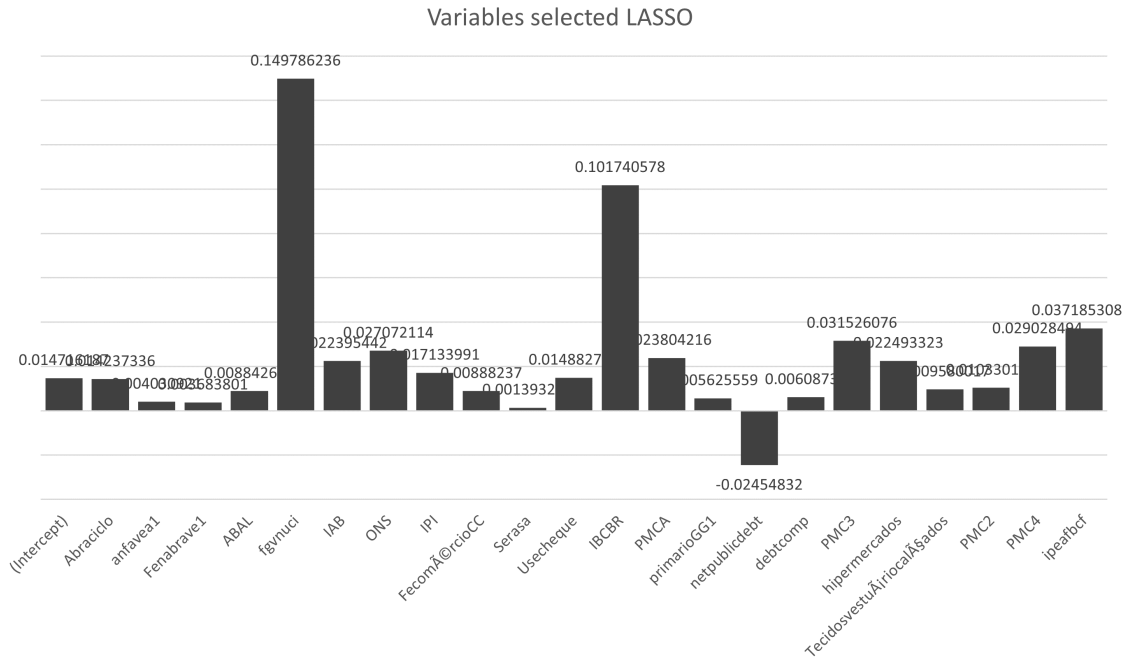


Figure 7: Selected Variables - LASSO

LASSO selection was the FGV’s idle capacity indicator, the NUCI. Also noteworthy is the negative impact on Brazilian GDP growth of the general government net debt. We also estimate a second LASSO model, the difference of this second estimation is only the pre-treatment of the data, in the former we used only the original database without lags, in the second estimation we used lags of the own nowcasting database. We did not have much change in the regressors selected for the final estimation between the first and the second estimations of LASSO. When we evaluate both LASSO 1 and LASSO 2 we repeat the result as seen before, these models are superior to the naive one at 0.2% significance. On the other hand they are not superior, by themselves, to the perfectly informed market estimates.

5.5 adaLASSO

LASSO (Least Absolute Shrinkage and Selection Operator) and adaLASSO (Adaptive LASSO) are both regularization methods for linear regression that aim to improve model accuracy and variable selection. The main difference between them lies in the penalty term applied to the coefficients. For a linear regression model, let \mathbf{y} be the response variable, \mathbf{X} be the predictor matrix, and β be the coefficient vector. The objective functions for LASSO and adaLASSO are as follows:

– LASSO:

$$\hat{\beta}_{\text{LASSO}} = \arg \min_{\beta} \left\{ \frac{1}{2n} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_1 \right\},$$

– adaLASSO:

$$\hat{\beta}_{\text{adaLASSO}} = \arg \min_{\beta} \left\{ \frac{1}{2n} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \sum_{j=1}^p w_j |\beta_j| \right\},$$

In the adaLASSO formula, w_j represents the adaptive weights for each coefficient, which are typically chosen based on some preliminary estimates of the coefficients, such as the OLS estimates. By introducing these adaptive weights, adaLASSO improves upon LASSO's variable selection consistency and estimation accuracy. As we can see, adaLASSO

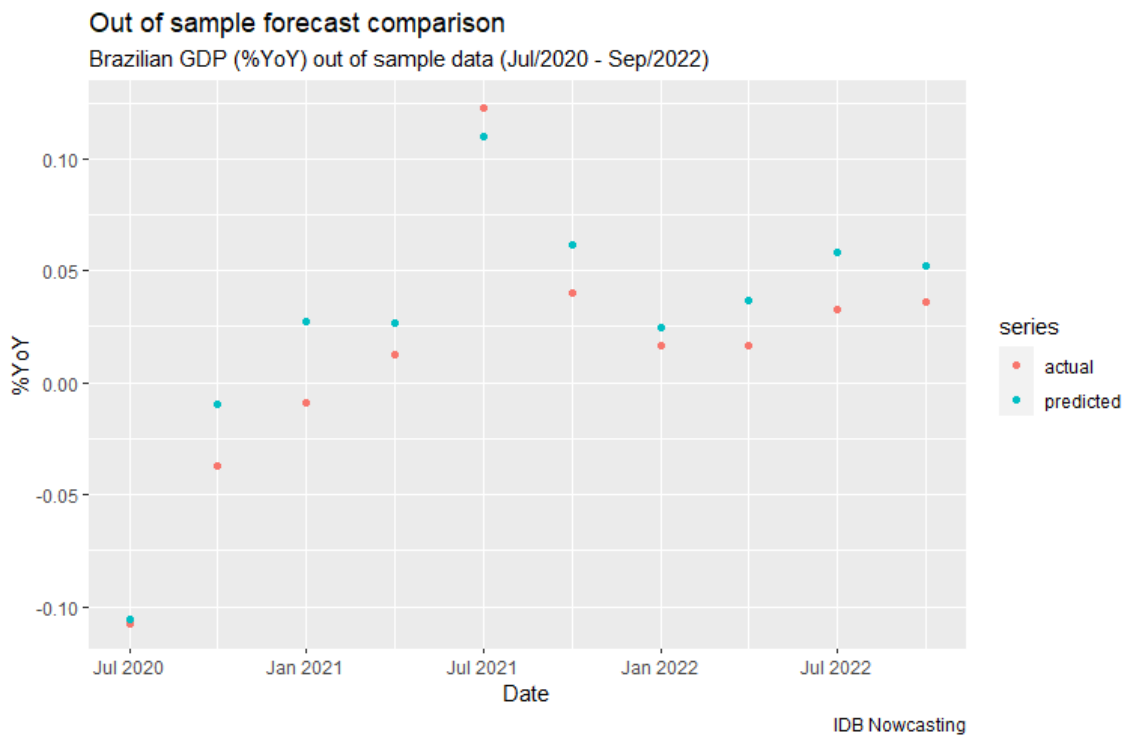


Figure 8: Out of sample forecast comparison - adaLASSO

performs very similarly to the previous ones in the out-of-sample projection. Concerning the evaluation of forecasts for this model, The Diebold-Mariano test for adaLASSO gives us a p-value of 99% for the comparison against the market and a p-value of 0.151% for the comparison against the naive model, again, the model by itself is better than a naive model but does not beats the market.

5.6 Ridge Regressions

Now we estimate a ridge regression, we use the cross validation method to obtain the optimal values for λ^5 , taking the package default, we use 10 cross-validations to obtain this parameter. After the estimation we can look at the realized values against the estimated values in an out-of-sample exercise. Evaluating these exercises we get the

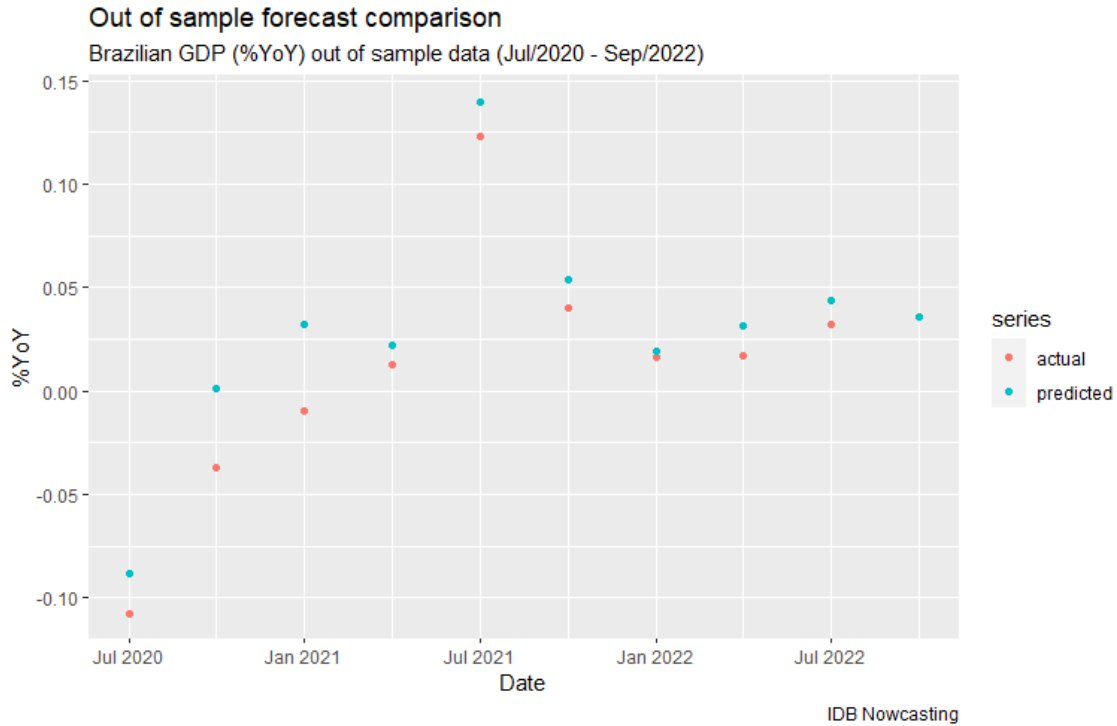


Figure 9: out of sample comparison - ridge regression

following results, the Ridge model is superior to the naive model, using the Diebold-Mariano test we obtained a p-value of 0.9% using 10 projection horizons. The alternative hypothesis is that the Ridge model is better (more accurate) than the AR model. The comparison result with perfectly informed expectations is repeated and the ridge model alone is not superior to the market, the p-value of this test is 99.90% for the same null hypothesis as above.

5.7 Elastic Net

The Diebold - Mariano tests show that the elastic net model performs much better than the naive model, using 10 projection horizons the exercise gives us a p-value of 0.5 percent, in the case of the comparison with the perfectly informed market the result is still unsatisfactory.

⁵The parameter λ , both in LASSO, ridge and elastic net, is the parameter that dictates the degree of penalty that the regression parameters will suffer. In the case of ridge, it is interesting to choose λ with some method, in our case we use cross-validation.

5.8 Neural Networks

In the case of neural networks we made three estimations, the difference between the estimations is the configuration of the neural network, in the first one we used a hidden layer and 2 neurons, in the second one we used a hidden layer and 3 neurons, and in the last estimation we used two hidden layers where the first presents 4 neurons and the second presents 2 neurons. We would like to point out that the determination of these parameters is the subject of a separate study, one direction in which we can advance in the quality of our projections is to improve the choice of these networks, additionally, it is interesting to note that there is a long way to go towards the use of *deep learning* techniques involving these networks. In the graph below we show the out of sample comparison of the backtesting exercises using these neural network estimations. The results observed in the latter models are repeated for the neural networks, that is, the models show greater predictive power than the naive model however they do not debunk the fully informed expectations.

5.9 Other models

We also estimated the BAGGING, Complete subset regression and Support Vector Machines models, below a summary table of the evaluation of the projections of these models.

*	*	MAE	*	*
BAGGING	RF	CSR	SVM	NAIVE
0.039	0.025	0.044	0.038	0.047
*	*	RMSE	*	*
BAGGING	RF	CSR	SVM	NAIVE
0.049	0.030	0.059	0.044	0.087

The result is repeated, the models alone fail to beat the market projections, although they do beat the naive model both by the RMSE and MAE values and by the Diebold Mariano test. Below we see the backtesting of these machine learning models compared, notice the good performance of the bagging method after covid-19.

5.10 Comparison ML vs market expectations

Even though, by the Diebold Mariano test, some models from the machine learning literature can in fact outperform the market, it is interesting to note that the RMSE and MAE of these models is on average larger than that of the fully informed market expectations, however it is worth noting that the gap between the expectations and the classical models has been largely diminished when we look at RMSE and MAE values, the next graph put's this argument in perspective.

5.11 Machine Learning Methods using selection algorithm

The application of the model selection algorithm brings good results for the machine learning models, the projection improves a lot, as we said above, some models in this literature already present a better projection than the perfectly informed market, however the use of the algorithm improves this result even more.

- Q1) LASSO2, adaLASSO
- Q2) LASSO2, RNN2
- Q3) Ridge, Elastic Net, RNN1, RNN2, RNN3, BAGGING
- Q4) adaLASSO, Ridge, CSR
- Q5) LASSO2, Elastic Net, RNN1, BAGGING, CSR, RF, SVM
- Q6) LASSO1, LASSO2, adaLASSO, Ridge, Elastic Net, RNN1, RNN2, RNN3, BAGGING, CSR, RF, SVM
- Q7) LASSO1, Ridge, Elastic Net, RNN1, RNN2, RNN3, BAGGING, CSR, RF, SVM
- Q8) LASSO1, Ridge, Elastic Net, RNN1, RNN2, RNN3, BAGGING, CSR, RF, SVM
- Q9) LASSO1, Ridge, Elastic Net, RNN1, RNN2, RNN3, BAGGING, CSR, RF, SVM
- Q10) LASSO1, adaLASSO, ridge, elastic net, RNN1, RNN2, RNN3, BAGGING, CSR, RF, SVM.

The selection algorithm brings us to the next table of error metrics:

Measure Model	FULLinfEXPEC	Final Forecast
RMSE	0.95	0.54
MAE	0.88	0.48

6 Concluding remarks

In this work we show that a combination that machine learning models alone and also a combination of selected classic + machine learning models can beat the fully informed FOCUS expectations at the eve of GDP releases, this performance is reached once we include a selection algorithm with a proposed metric, below we highlight a list of results of our work.

- Machine learning methods perform better in this exercise, however, it's not the case that these models are unconditionally superior, we estimated a very large number of models from this literature, the combination of the projections may have caused the RMSE and MAE to drop.
- The Diebold Mariano test between the combination of machine learning models and perfectly informed expectations reveals that the combination coming from our algorithm is superior to market expectations at 1% significance (p-value of the order of 10^{-5}).
- The model-matching algorithm improves the projection of the classical models but does not generate a projection that is better than the one made by the market on the eve of the GDP release.
- The simple combination of the projection generated by the machine learning models and that generated by the classical models is superior to the market, but the significance is much lower than that derived from the statistical superiority of the machine learning models.
- The aggregation of factors in dynamic models using Mariano & Murasawa's aggregation did not improve the projection of the models, contrary to evidence in the literature.

It's important to note that these results came from an exercise of specific selected vintages, the end of each quarter at the selected testing sample years.

7 Bibliographic references

References

- AASTVEIT, K. A.; TROVIK, T. Estimating the output gap in real time: A factor model approach. *The Quarterly Review of Economics and Finance*, v. 54, n. 2, p. 180–193, 2014.
- BANBURA, M. et al. Nowcasting and the real-time data flow. [s.l.: s.n.]. Disponível em: <<https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1564.pdf>>. Acesso em: 28 set. 2022.
- Bai, J., Ng, S. (2003). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 71(1), 191-221. <https://doi.org/10.1111/1468-0262.00273>
- Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(6), 135-171. <https://doi.org/10.1111/1468-0262.00392>
- BARSOUM, F.; STANKIEWICZ, S. Forecasting GDP Growth Using Mixed-Frequency Models With Switching Regimes. Disponível em: <<https://ideas.repec.org/p/knz/dpteco/1310.html>>. Acesso em: 28 set. 2022.
- BERGMEIR, C., HYNDMAN, R. J., BENÍTEZ, J. M. (2016), Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation. *International Journal of Forecasting*, 32(2), 303–312.
- BREIMAN, L. (2001), Random forests. *Machine learning*, 45(1):5–32.
- BURNS, A. F.; MITCHELL, W. C. Measuring Business Cycles. Disponível em: <<https://www.nber.org/books-and-chapters/measuring-business-cycles>>.
- CAMACHO, M. Mixed-frequency VAR models with Markov-switching dynamics. *Economics Letters*, v. 121, n. 3, p. 369–373, dez. 2013.
- Cascaldi-Garcia, D., Ferreira, T. R. T., Giannone, D., Modugno, M. (2021). Back to the Present: Learning about the Euro Area through a Now-casting Model. *International Finance Discussion Papers*, 1313. Board of Governors of the Federal Reserve System
- DAHLHAUS, T.; GUÉNETTE, J.-D.; VASISHTHA, G. Nowcasting BRIC+M in Real Time. Disponível em: <<https://www.bankofcanada.ca/2015/10/working-paper-2015-38/>>. Acesso em: 28 set. 2022.
- DIEBOLD, F. X.; RUDEBUSCH, G. D. Measuring Business Cycles: A Modern Perspective. *The Review of Economics and Statistics*, v. 78, n. 1, p. 67–77, 1996.
- Elliot, G., Gargano, A., Timmermann, A. (2013). Complete subset regressions. *Journal of Econometrics*, 177(2), 357-373.

- FERRARA, L.; MARSILLI, C.; ORTEGA, J.-P. Forecasting growth during the Great Recession: is financial volatility the missing ingredient? *Economic Modelling*, v. 36, n. C, p. 44–50, 2014.
- GIANNONE, D.; REICHLIN, L.; DE MOL, C. Forecasting using a large number of predictors: Is Bayesian regression a valid alternative to principal components? Disponível em: <<https://econpapers.repec.org/paper/ecbecwps/2006700.htm>>. Acesso em: 28 set. 2022.
- GIANNONE, D.; REICHLIN, L.; SMALL, D. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, v. 55, n. 4, p. 665–676, 2008.
- DOZ, C.; GIANNONE, D.; REICHLIN, L. A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, v. 164, n. 1, p. 188–205, 2011.
- Bok, B., Caratelli, D., Giannone, D., Sbordone, A. M., Tambalotti, A. (2017). *Macroeconomic Nowcasting and Forecasting with Big Data*. Staff report, 830. Federal Reserve Bank of New York.
- Giannone, D., Lenza, M., Primiceri, G. E. (2021). *Economic Predictions With Big Data: The Illusion of Sparsity*. *Econometrica*, 89(5), 1685-1729.
- GHYSELS, E.; SANTA-CLARA, P.; VALKANOV, R. The MIDAS Touch: Mixed Data Sampling Regression Models. Disponível em: <<https://econpapers.repec.org/paper/circirwor/2004s-20.htm>>.
- HAMILTON, J. D. Rational-expectations econometric analysis of changes in regime. *Journal of Economic Dynamics and Control*, v. 12, n. 2-3, p. 385–423, jun. 1988.
- HEINRICH, M. et al. Predicting Ordinary and Severe Recessions with a Three-State Markov-Switching Dynamic Factor Model. An Application to the German Business Cycle. Disponível em: <<https://econpapers.repec.org/paper/zbwvfsc17/168206.htm>>. Acesso em: 28 set. 2022.
- IMF (2022) *Nowcasting GDP - A Scalable Approach Using DFM, Machine Learning and Novel Data, Applied to European Economies*. Disponível em: <<https://www.imf.org/en/Publications/WP/Issues/2022/03/11/Nowcasting-GDP-A-Scalable-Approach-Using-DFM-Machine-Learning-and-Novel-Data-Applied-to-513703>>.

- INOUE, A., KILIAN, L., How Useful is Bagging in Forecasting Economic Time Series? A Case Study of U.S. Consumer Price Inflation. *Journal of the American Statistical Association*, 103(482), 511–522.
- KYOUNG-JAE KIM (2003), Financial time series forecasting using support vector machines.
- LAINE, O.-M.; LINDBLAD, A. Nowcasting Finnish GDP growth using financial variables: a MIDAS approach. Disponível em: <<https://econpapers.repec.org/paper/zbwbofecr/42020.htm>>. Acesso em: 28 set. 2022.
- MAYNARD, L. S. (2020), Nowcasting GDP with Machine Learning Models: Evidence from the US.
- MARIANO, R. S.; MURASAWA, Y. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, v. 18, n. 4, p. 427–443, 2003.
- Global Inflation Forecasting: Benefits from Machine Learning Methods. Available at SSRN: <https://ssrn.com/abstract=4145665> or <http://dx.doi.org/10.2139/ssrn.4145665>
- Modugno, M. (2013). Now-casting inflation using high frequency data. *International Journal of Forecasting*, 29(4), 664-675.
- OLIVEIRA, L.,M.,NAKANE, M. I. - Nowcasting do PIB brasileiro: atualização do exercício e recomendações qualitativas. 2020. Trabalho de Conclusão de Curso (Graduação em Economia) - Faculdade de Economia, Administração e Contabilidade - USP.
- THOMAS, S.; CHRISTOPHER, S. Business cycle modeling without pretending to have too much a priori economic theory. 1977.
- SIMS, C. A. Macroeconomics and Reality. *Econometrica*, v. 48, n. 1, p. 1, jan. 1980.
- STOCK, J.; WATSON, M. *The Leading Economic Indicators: New Approaches and Forecasting Records*. [s.l.] Cambridge University Press, 1991. Disponível em: <<https://scholar.harvard.edu/stock/publications/probability-model-coincident-economic-indicators>>. Acesso em: 28 set. 2022.
- Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics.
- STOCK, J. H.; WATSON, M. W. Core Inflation and Trend Inflation. *The Review of Economics and Statistics*, v. 98, n. 4, p. 770–784, 1 out. 2016.

- THORSRUD, L. A. Nowcasting using news topics. Big Data versus big bank. Disponível em: <https://ideas.repec.org/p/bno/worpaper/2016_20.html>. Acesso em: 28 set. 2022.
- VALK, S. DE; MATTOS, D. DE; FERREIRA, P. Nowcasting: An R Package for Predicting Economic Variables Using Dynamic Factor Models. *The R Journal*, v. 11, n. 1, p. 230–244, 2019.
- YIU, K., CHOW, M., RESEARCH, H. K. I. FOR M. AND F. Nowcasting Chinese GDP: Information Content of Economic and Financial Data. Disponível em: <<https://papers.ssrn.com/sol3/papers.cfm?abstractid=1769611>>. Acesso em: 28 set. 2022.
- gabrielrvsc. (2013). HDeconometrics: High-dimensional econometrics. R package version 0.1.0. <https://CRAN.R-project.org/package=HDeconometrics>. Accessed January 13, 2023.
- Koopman, S.J., and van Dijk, D. (2018). midasr: MIDAS Regressions in R. R package version 0.1. <https://CRAN.R-project.org/package=midasr>. Accessed January 13, 2023
- Friedman, J., Hastie, T. and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1), 1-22. <https://CRAN.R-project.org/package=glmnet>. Accessed January 13, 2023.
- Diebold, F. X., Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business Economic Statistics*, 13(3), 253-263.
- Diebold, F. X., Mariano, R. S. (2002). Comparing Predictive Accuracy. In J. J. Heckman E. E. Leamer (Eds.), *Handbook of Econometrics* (pp. 5897-5967). Elsevier.
- Diebold, F. X., Mariano, R. S. (2003). Comparing Density Forecasts Using Various Information Criteria. *Journal of Business Economic Statistics*, 21(4), 600-608.
- Drachal, K. (2022). multDM: Multivariate Version of the Diebold-Mariano Test. R package version 0.1.0. <https://CRAN.R-project.org/package=multDM>.
- Mariano, R.S. and Preve, D. (2012). "A multivariate version of the Diebold-Mariano test for equal predictive ability of multiple forecast comparison", *Journal of Econometrics*, vol. 170(2), pp.453-467, Elsevier. doi:10.1016/j.jeconom.2012.01.014
- Krantz, S., Bagdziunas, R. (2023). Dynamic Factor Models (dfms) R package version 0.1.4. <https://CRAN.R-project.org/package=dfms>.
- Schumacher, C. "MIDAS and bridge equations." Discussion Paper, Deutsche Bundesbank, no. 26/2014.
- Götz, T.B. Knetsch, T.A. "Google data in bridge equation models for German GDP" *International Journal of Forecasting*, vol. 35, no. 1, pp. 45-66, Jan-Mar 2019.

- de Valk, S., de Mattos, D., Ferreira, P. "Nowcasting: An R Package for Predicting Economic Variables Using Dynamic Factor Models." *The R Journal*, 11(1), 230-244, 2019.
- Gomes, G. B., Issler, J. V. (2018). Nowcasting Brazilian GDP: A Performance Assessment of Dynamic Factor Models. Master's thesis, Graduate School of Economics, Rio de Janeiro, Brazil.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.
- Banbura, M., Giannone, D., Reichlin, L. (2010). Large Bayesian VARs. *Journal of Applied Econometrics*, 25(1), 71-92.
- Bernanke, B. S., Mihov, I. (1998). Measuring Monetary Policy. *Journal of Economic Perspectives*, 12(4), 75-92.
- Buell, B., Chen, C., Cherif, R., Seo, H.J., Tang, J., Wendt, N. (2021). Impact of COVID-19: Nowcasting and Big Data to Track Economic Activity in Sub-Saharan Africa. IMF Working Paper, 21-124.

8 Anexes

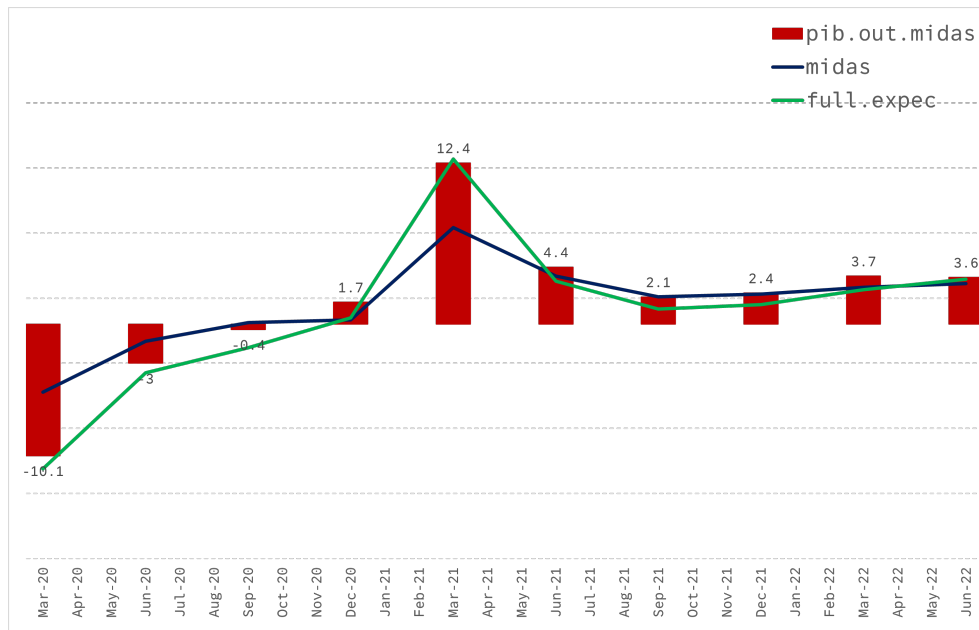


Figure 10: Out of sample forecast comparison MIDAS/fully informed market expectations

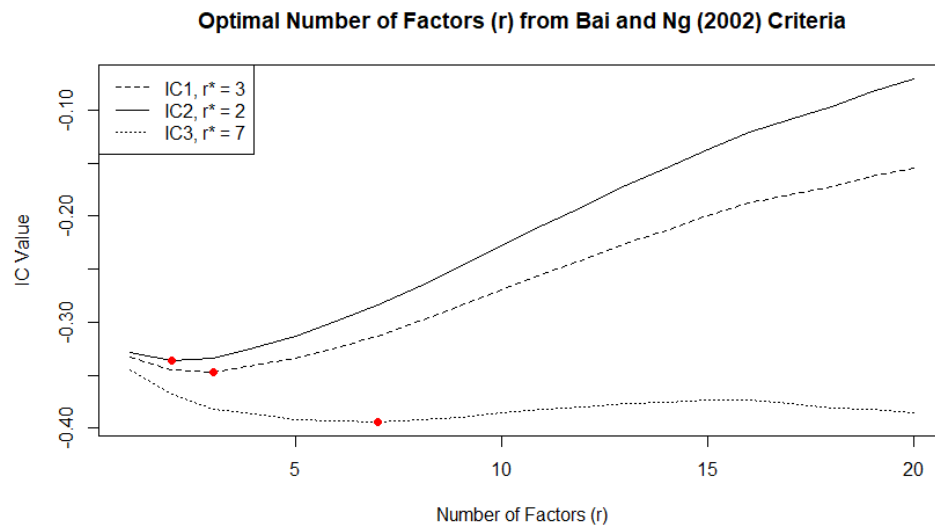


Figure 11: Information Criteria following BAI & NG (2002), source: dfms

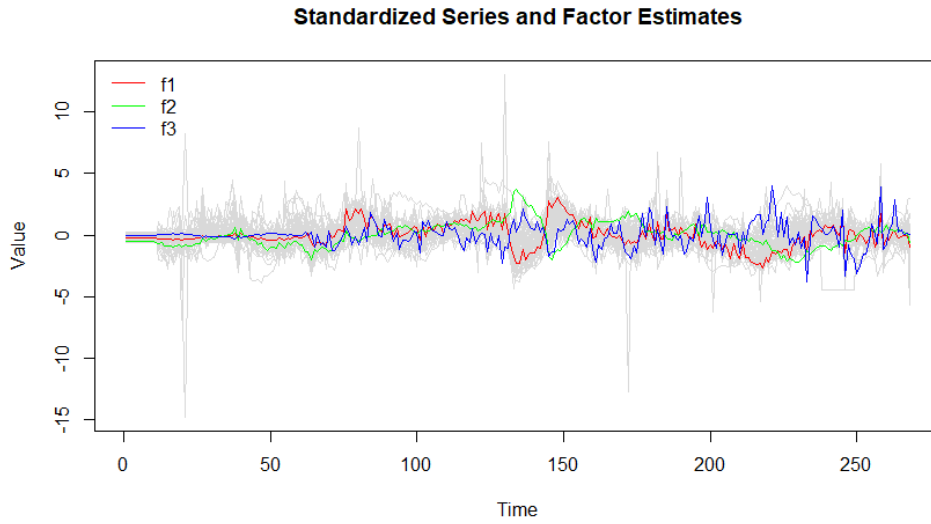


Figure 12: Extracted Factor series, source: dfms

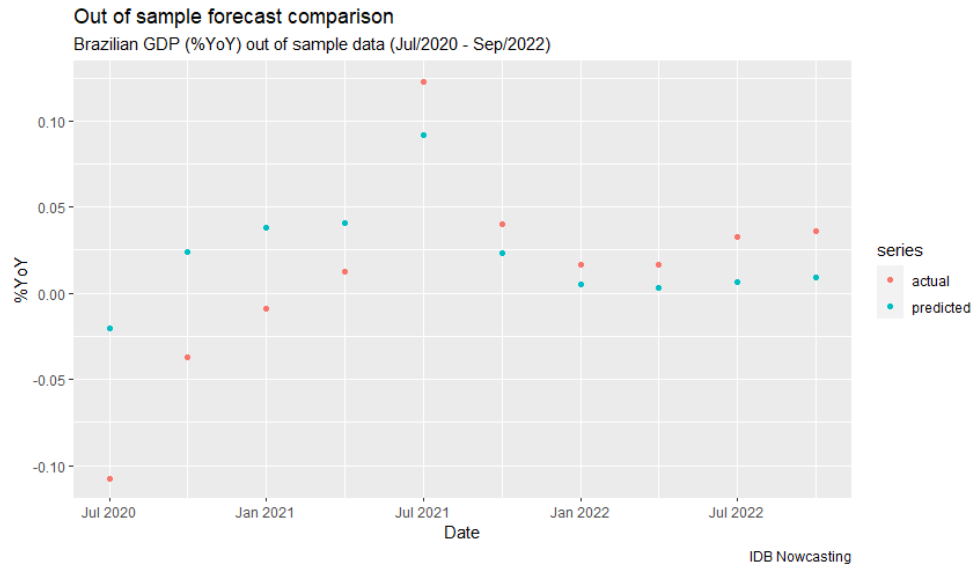


Figure 13: DFM (3,3) out of sample comparison

Table 7: Diebold-Mariano Test Results (DFM(2,2) vs Naive)

Forecast Horizon	Test Statistics	
	DM	p-value
10	-3.54	0.003157

Table 8: Diebold-Mariano Test Results (DFM(3,3) vs Naive)

Forecast Horizon	Test Statistics	
	DM	p-value
10	-3.29	0.004681

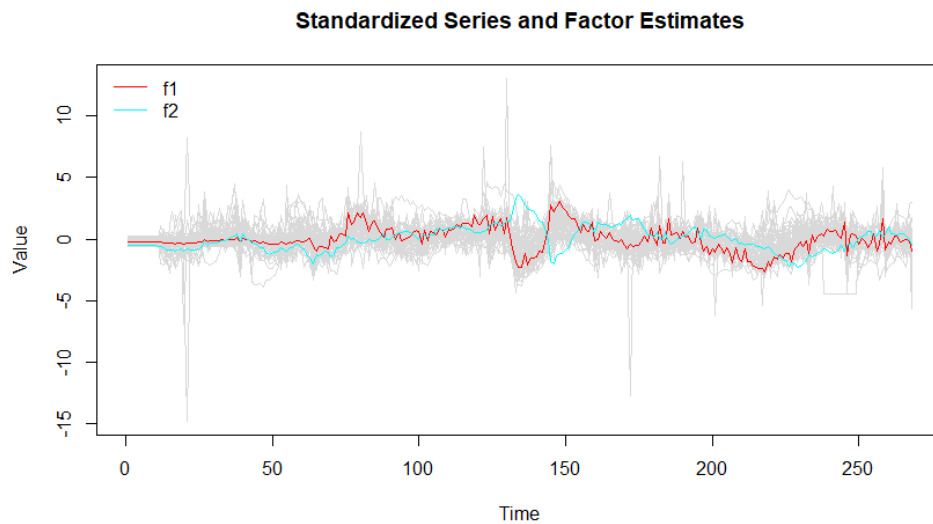


Figure 14: DFM (2,2) Factors extraction.

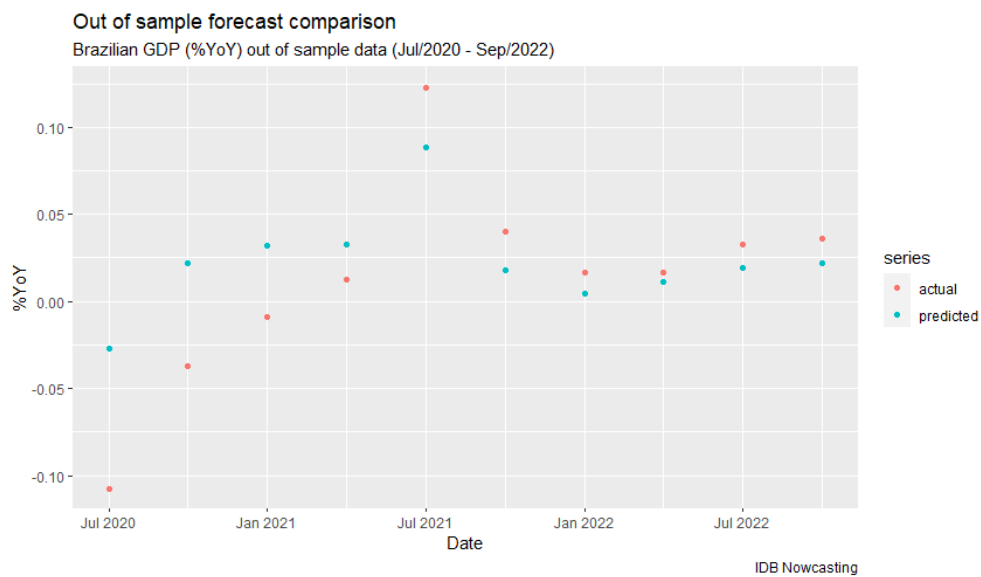


Figure 15: DFM (2,2) out of sample comparison.

Table 9: Significance DFM(2,2) with Mariano Murasawa Aggregation on it's factors

Variable	Estimate	Standard Error	t-statistic
Intercept	0.02182***	0.00008	25.1
Factor 1 agg	0.00643***	0.00026	29.8
Factor 2 agg	0.00012**	0.00046	2.6

Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '

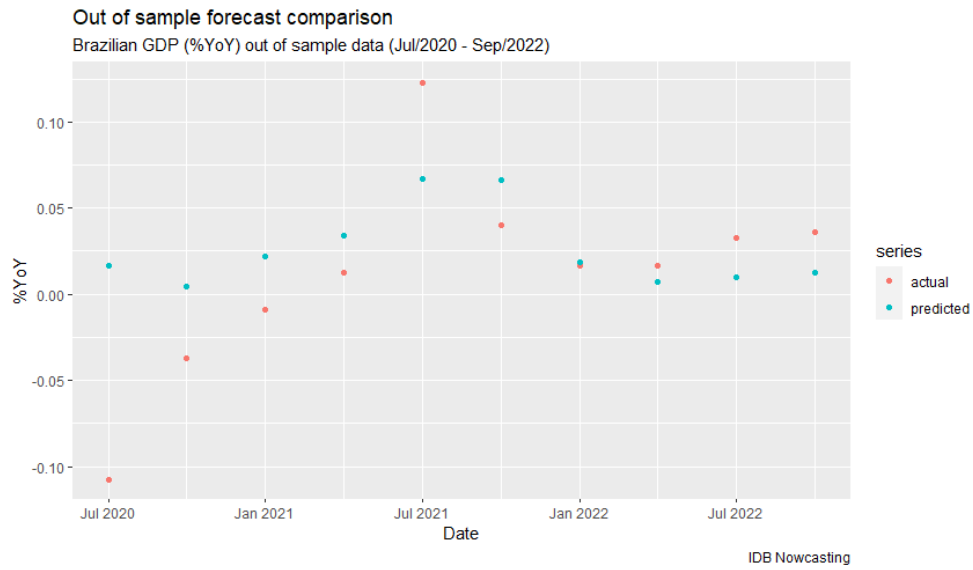


Figure 16: Out of sample comparison dfm agg(3,3)

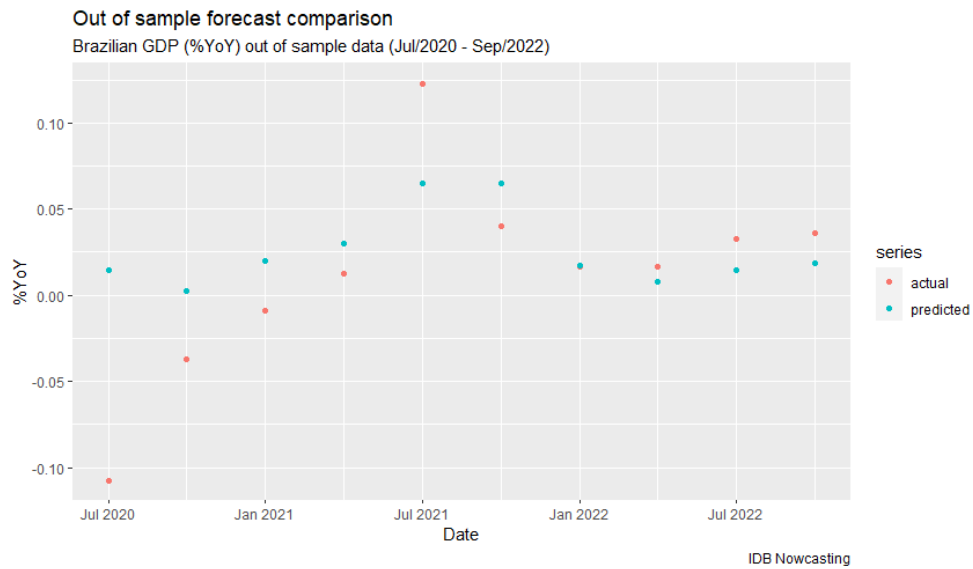


Figure 17: Out of sample comparison dfm agg(2,2)

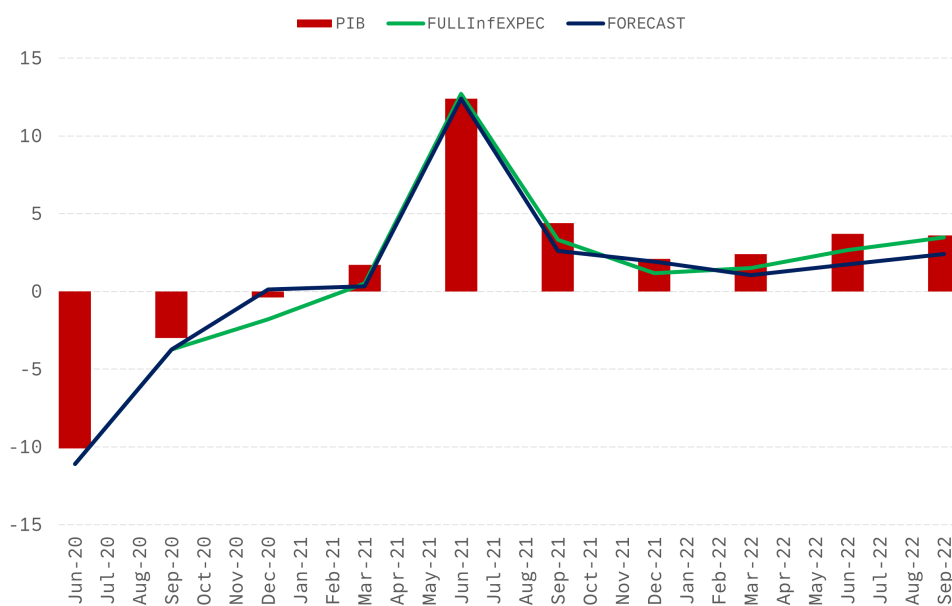


Figure 18: RMSE/MAE comparison out of sample forecasts (Combination of classical models vs market)

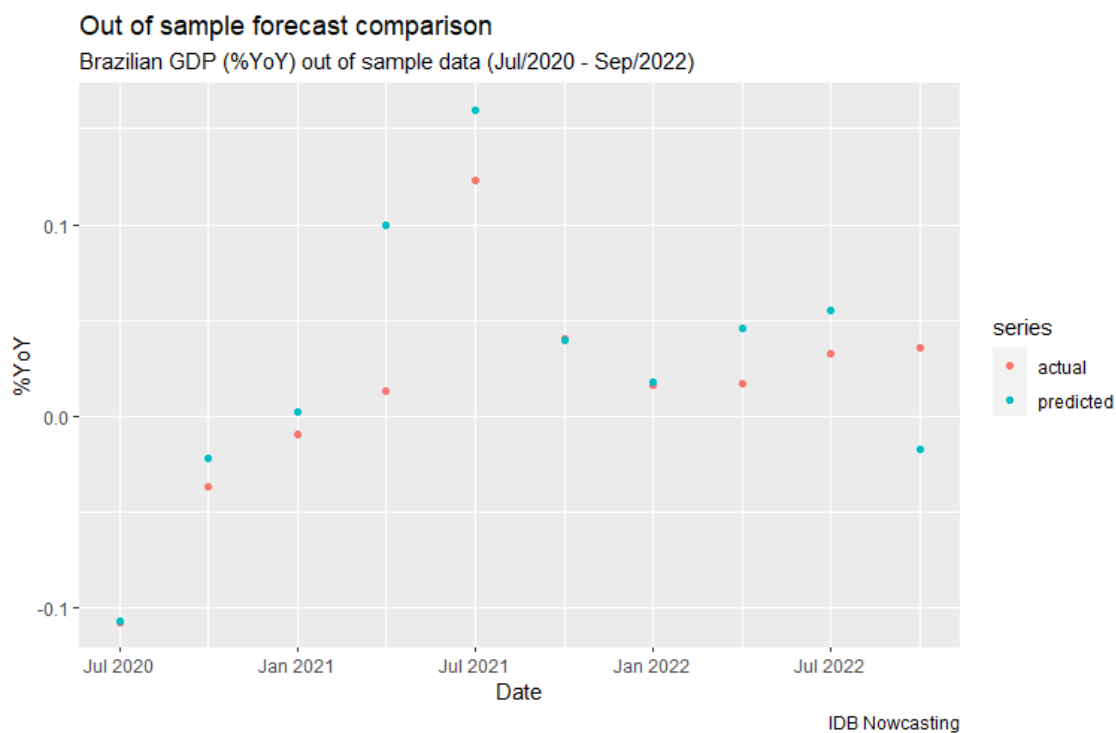


Figure 19: Out of sample forecast comparison - LASSO 2

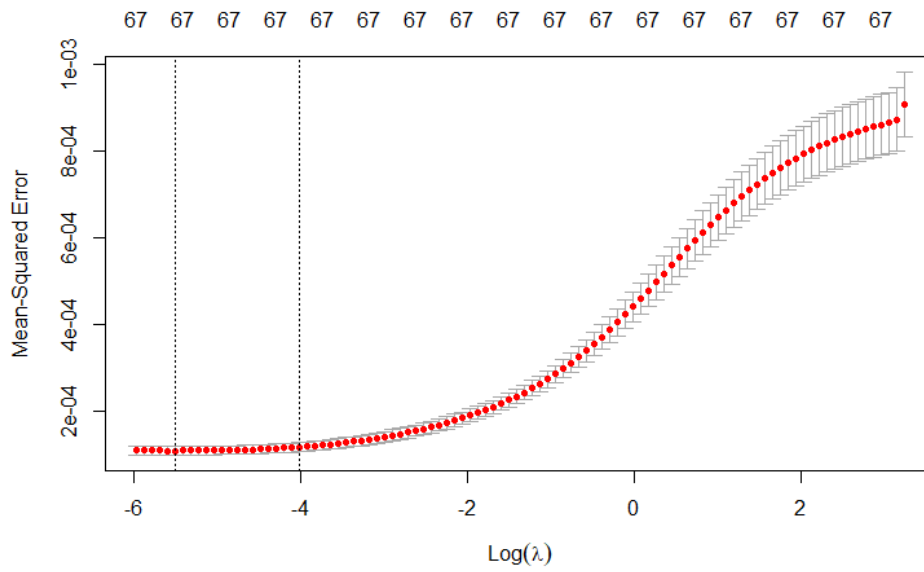


Figure 20: λ definition for ridge regression

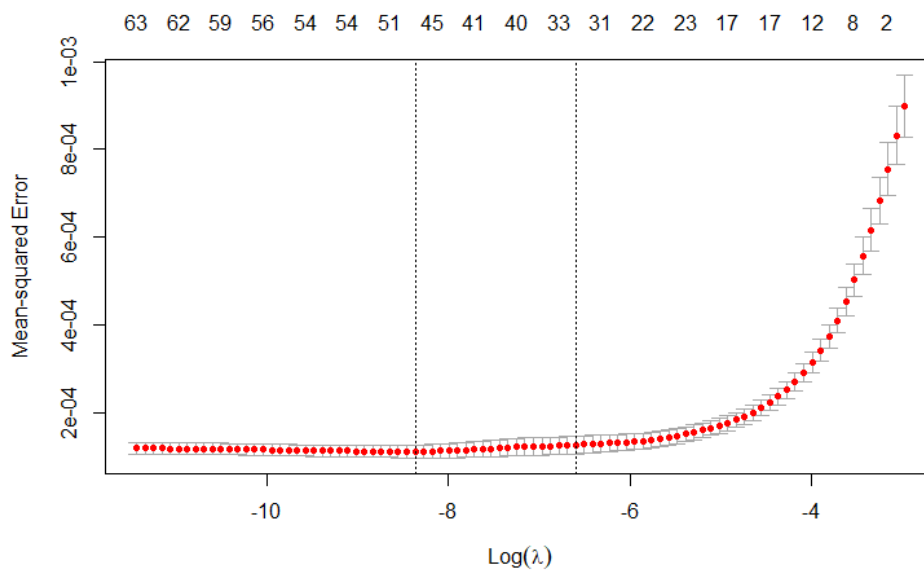


Figure 21: Elastic Net regression plot

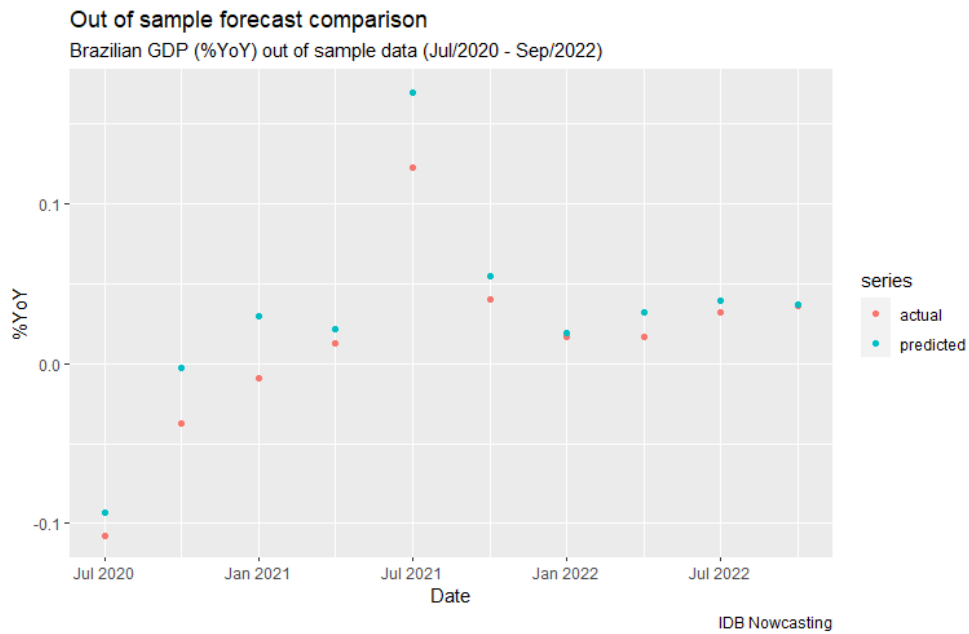


Figure 22: Elastic Net regression Out of sample comparison

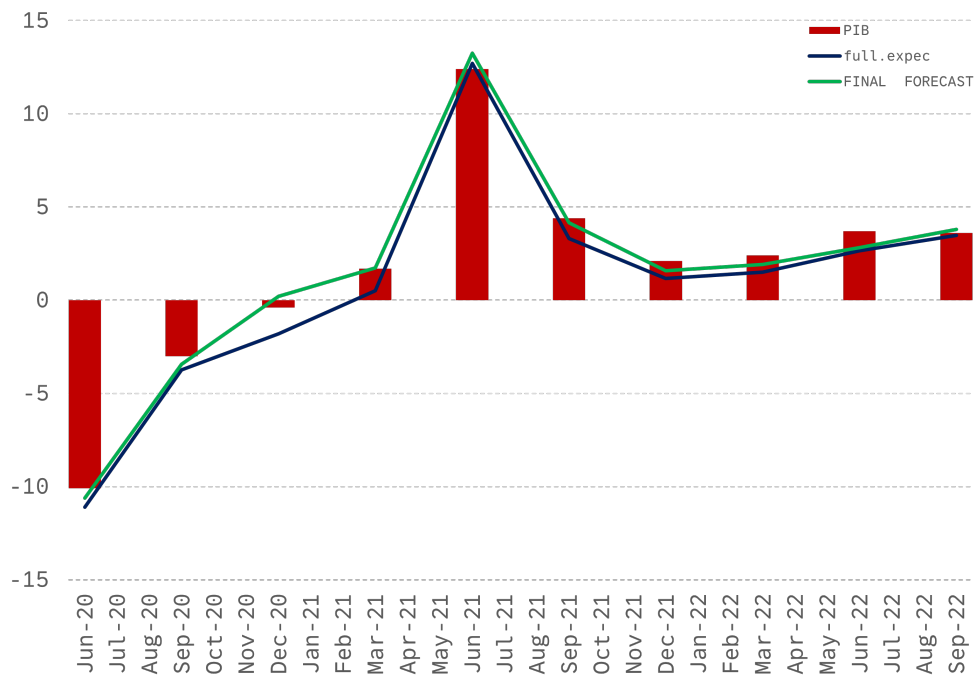
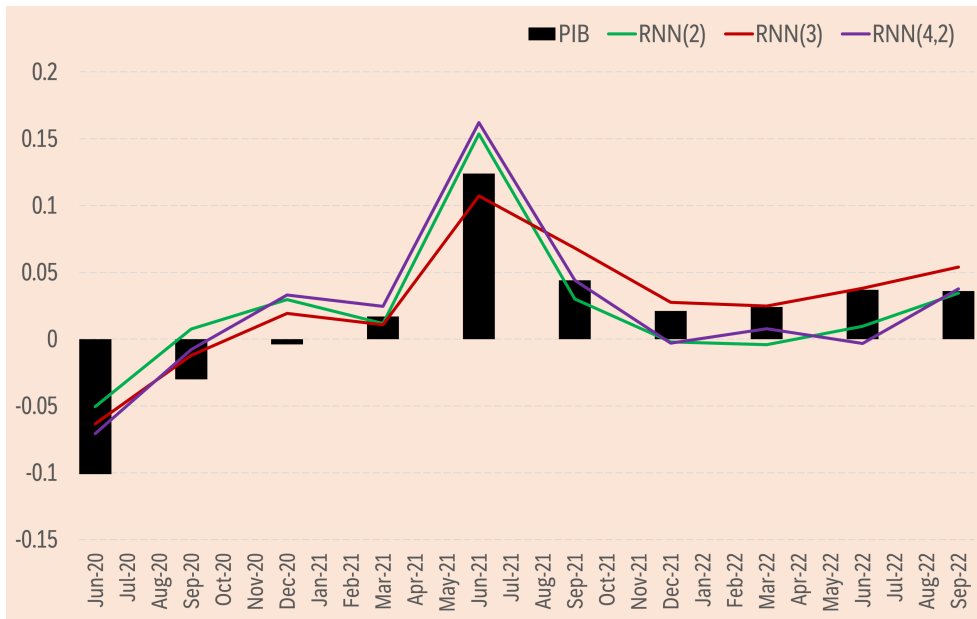


Figure 23: Out of sample comparison combination of machine learning models and fully informed market expectations



• Fig. 33: Out of sample comparison of selected neural networks.

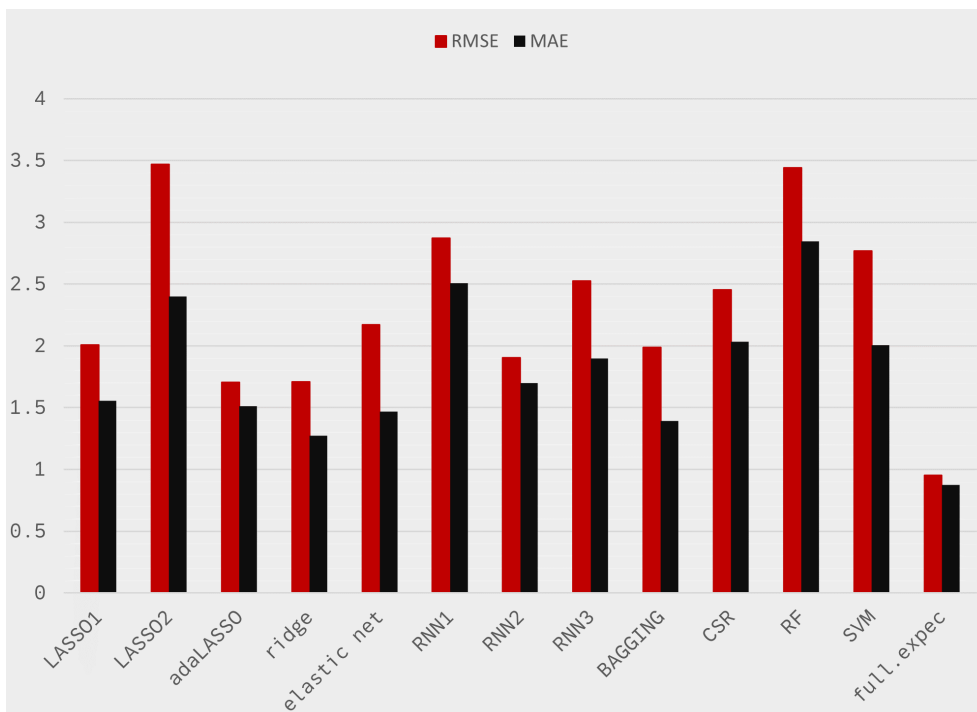


Figure 24: RMSE/MAE comparison out of sample forecasts (combination of machine learning models vs market)

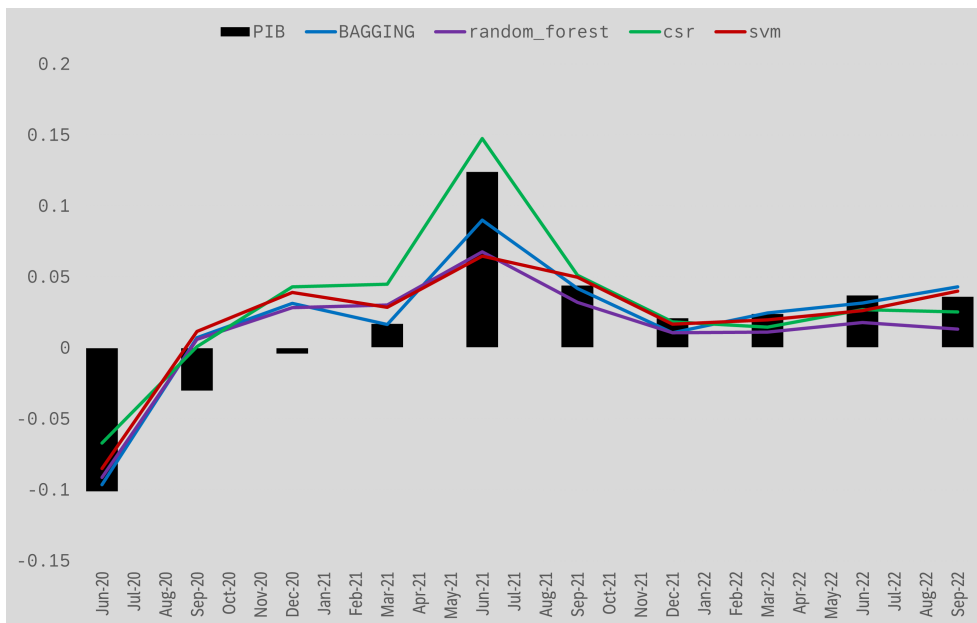


Figure 25: Out of sample forecast comparison