Forecasting Aggregate Inflation using Disaggregates and Machine Learning

Gilberto Boaretto
Department of Economics, PUC-Rio, Brazil
Email: gilbertoboaretto@hotmail.com

Marcelo C. Medeiros
Department of Economics, PUC-Rio, Brazil
Email: mcm@econ.puc-rio.br

This version: December 20, 2022
Preliminary

Abstract

This paper investigates the performance of forecasting aggregate inflation from the aggregation of disaggregated forecasts considering a large number of predictors and machine learning methods. The benchmarks are survey-based forecasts and traditional time series models forecasting the aggregate directly. Considering the Brazilian inflation, we show the benefits of owning both disaggregated data and machine learning techniques in the forecaster's toolbox. Besides exploring the models in isolation, we propose three model combinations that contribute to obtaining more accurate forecasts: (i) average between model-based forecasts and inflation expectation of the Focus survey available when the econometrician computes your forecast, (ii) distinct methods forecasting different disaggregates to compute the aggregation, and (iii) choice over time between all approached models. The model selections of the two last combinations are based on the minor out-of-sample squared error in a previous period.

Keywords: inflation forecasting; disaggregate inflation; information set; dimensionality reduction; variable selection; model combination.

JEL Codes: C22, C38, C52, C53, C55, E37.

Acknowledgements: Boaretto thanks PUC-Rio, CAPES, CNPq, and FAPERJ for financial support. Medeiros thanks CNPq and CAPES for financial support.

44th Meeting of the Brazilian Econometric Society (SBE)
Fortaleza, CE, Brazil – December 2022
Track: Econometrics

1 Introduction

Economists and econometricians are always interested in owning the best tools to deliver inflation forecasts as well as possible. Considering the behavior of disaggregated prices in different markets or economic classifications can improve the forecasting performance since this could better capture the dynamics in trend, seasonality, and short-term changes (Espasa, Senra, and Albacete, 2002). In other words, using a most detailed individual series would allow the econometric models to capture better the heterogeneity of the data (Bermingham and D'Agostino, 2014). When the data generating process is unknown, whether direct or indirect forecasting via the aggregating of disaggregated forecasts improves or does not improve the forecast accuracy of the aggregate is strictly an empirical question (Lütkepohl, 1984; Hendry and Hubrich, 2011; Faust and Wright, 2013). However, it is essential to impose some dimensionality reduction to control the estimation uncertainty generated by the aggregation of disaggregated forecasts. The potential solution for this consists of using factor- or shrinkage-based models, techniques belonging to the toolbox called "machine learning" (ML). These techniques have been widely used in economic research and proven to be valuable tools (Mullainathan and Spiess, 2017; Athey, 2019; Gogas and Papadimitriou, 2021).

The broad literature on inflation forecasting documents that the predictive performance of the survey-based forecasts is challenging to beat, especially in the short term – current and immediately next months (Thomas, 1999; Ang, Bekaert, and Wei, 2007; Croushore, 2010). In this line, Faust and Wright (2013) argue that "purely subjective forecasts are in effect the frontier of our ability to forecast inflation" because, besides private sectors and central banks having access to econometric models, they add expert judgment to these models. Consequently, "a useful way of assessing models is by their ability to match survey measures of inflation expectations" (Faust and Wright, 2013). A potential explanation is that forecasters are likely to have a richer information set than the econometrician employing a standard set of macroeconomic variables (Del Negro and Eusepi, 2011). Thus, including observed expectations among the predictors is a way to exploit such information set. Following this, Baştürk, Çakmakli, Ceyhan, and Van Dijk (2014), Altug and Çakmaklı (2016), Fulton and Hubrich (2021), and Bańbura, Leiva-León, and Menz (2021) found evidence favorable to the incorporation of survey-based forecasts into forecasting models.

This paper investigates the effectiveness of forecasting aggregate inflation from the aggregation of disaggregated forecasts computed using several methods. Analyzing the Brazilian case, we compared the predictive performance of the aggregation approach with more traditional approaches found in the literature – survey-based forecasts and forecasting directly from the aggregate. Considering different disaggregations, we run forecast exercises considering both aggregate and aggregating disaggregate forecast approaches. The increase in informational granularity is a potential advantage of considering disaggregates. Besides considering specific effects of the traditional macro-variables related to money, economic activity, government, and external sector, we can consider lagged, crossed effects between the disaggregates in the disaggregated analysis. We also consider as a predictor a survey-based expectation *available* when the econometrician computes her/his forecasts. As emphasized earlier, this variable can potentially add information not contained in other variables for the econometrician. Finally, we employ traditional time series techniques as well as linear and nonlinear ML techniques to deal with a larger number of predictors. More specifically, we consider these modeling possibilities:

1. Traditional time series methods: random walk (RW), historical mean, and autoregressive (AR) models;
2. Shrinkage models with or without sparsity, namely, Ridge and adaptive LASSO (adaLASSO);
3. Factor and target factor augmented models;
4. FarmPredict, a model that bridges both common factor and sparsity structures. With this method, we can explore the possibility of remaining sparse idiosyncratic effects after controlling for common factors, and autoregressive, expectation and deterministic components;
5. Complete Subset Regression (CSR), an ensemble method that combines estimates from all possible linear regression models keeping the number of predictors fixed;
6. Random Forest (RF), a "bagged" tree-based model.

The Brazilian case is interesting for several reasons. First, the official Brazilian price index, the Broad Consumer Price Index (*Índice de Preços ao Consumidor Amplo* – IPCA), is monthly computed and has a rich structure to be explored. There are several disaggregations and all their weights in the consumption basket are available. Second, there is an extensive survey of inflation expectations in daily frequency that covers multiple forecast horizons and is managed by the Central Bank of Brazil: the Focus survey. This survey reflects the opinion of market agents (experts) and may contain inside information that is not available to the econometrician. Besides being used as a predictor to generate model-based forecasts, Focus inflation expectations can be used as a benchmark to verify whether it is possible to improve the survey-based forecast at some horizon. Third, due to Brazil's inflationary history, in addition to the official price index, the country has several price indices that can be used as potential predictors of the inflation. Thus, another question is to investigate whether this price set is useful to forecast Brazilian inflation for some horizons ahead.

Findings. Among the main results of this paper are:

(i) it is not easy to outperform the survey-based forecast (Focus) in the short forecast horizons, but it is possible for greater ones;

(ii) taking disaggregated prices into account contributes to improving monthly inflation forecasts from next month onwards (considering all disaggregations analyzed);

(iii) in general, machine learning techniques performed better than traditional time series techniques, especially adaLASSO and FarmPredict – potentially, this derives from the possibility of dealing well with a large number of predictors extracting relevant information;

(iv) the variety of predictors selected by adaLASSO and FarmPredict was notable and, in general, the available survey-based inflation expectations, price variables, and a long-term interest rate were the most relevant variables in forecast aggregate Brazilian inflation as well as its disaggregates, mainly nontradables;

(v) combining (via average) model-based forecasts with available Focus survey expectations improved inflation forecasts for shorter horizons;

(vi) combining different models forecasting different disaggregations tended to improve the accuracy of forecasts for aggregate inflation, especially when we considered a broader disaggregation (i.e., groups and subgroups) that, in general, had not performed as well when employing the same technique to forecast all disaggregates;

(vii) combining all approached models to obtain a forecast for aggregate performs very well for intermediate and longer horizons;

(viii) finally, also for all combinations cases, computing the average between combinations and survey-based forecasts led to improvements to forecast short-term aggregate inflation.

Contributions for the literature. We can summarize the contributions of this paper in two fields. First, this paper contributes to the literature on inflation forecasting via aggregation of disaggregated forecasts by considering many predictors for each disaggregation, as well as various statistical and econometric methods little or not explored in that literature. Many papers have employed traditional time series models and a limited number of predictors. In this context, some papers have found evidence in favor of aggregating disaggregated forecasts for the Euro Area (Espasa, Senra, and Albacete, 2002; Espasa and Albacete, 2007) and various countries (Bruneau, De Bandt, Flageollet, and Michaux, 2007; Duarte and Rua, 2007; Moser, Rumler, and Scharler, 2007; Capistrán, Constandse, and Ramos-Francia, 2010; Aron and Muellbauer, 2012; Carlo and Marçal, 2016; Fulton and Hubrich, 2021). On the other hand, some papers have found that aggregating forecasts by components does not necessarily improve aggregate inflation forecasting (Benalal, Diaz del Hoyo, Landau, Roma, and Skudelny, 2004; Hubrich, 2005; Hendry and Hubrich, 2011). In its turn, Ibarra (2012) and Bermingham and D'Agostino (2014) highlight the benefits of aggregating a large number of disaggregates. Florido (2021, Ch. 1) point out the benefits of considering disaggregations for nowcasting Brazilian aggregate inflation. Our work extends the literature by showing that forecasts that consider aggregation of different disaggregations and use *available* survey-based expectation as a predictor can be more accurate for several horizons than *current* survey-based forecasts and forecasts based on traditional methods.

Second, our analysis extends the literature on machine learning (ML) benefits to forecasting inflation by showing a useful application of these methods considering the aggregation of disaggregated forecasts in a data-rich environment – a hitherto unexplored aspect. To forecast aggregate inflation directly, the employ of ML methods started with the application of factor and principal component models (Stock and Watson, 1999, 2002; Forni, Hallin, Lippi, and Reichlin, 2003; Bai and Ng, 2008; Ibarra, 2012), and neural network models (Moshiri and Cameron, 2000; Nakamura, 2005; Choudhary and Haider, 2012). Several other papers expanded the list of methods to shrinkage-based models (e.g., Ridge and LASSO), Bayesian methods, bagging, boosting, random forest (RF), and complete subset regression (CSR), but keeping the focus directly on the aggregate inflation (e.g., Inoue and Kilian, 2008; Medeiros, Vasconcelos, and Freitas, 2016; Garcia, Medeiros, and Vasconcelos, 2017; Zeng, 2017; Baybuza, 2018; Medeiros, Vasconcelos, Veiga, and Zilberman, 2021; Araujo and Gaglianone, 2022). Recently some works have combined ML estimation and aggregation of disaggregated forecasts. Florido (2021, Ch. 1) considers several disaggregations to predict short-term aggregate inflation and found good results. Araujo and Gaglianone (2022) forecast inflation by considering a disaggregation into administered prices, services, industrial goods, and food at home at several horizons employing ARMA, adaLASSO, and RF. However, their results are not favorable to the disaggregated analysis. Our paper extends this piece of the literature by indicating that the aggregation of disaggregated forecasts obtained employing different machine learning methods in a data-rich environment generates an attractive expansion of possibilities. The simple investigated model combinations obtained more accurate forecasts than survey-based inflation expectations for at least half of the forecast horizons.

Structure. This paper has five more sections in addition to this Introduction. Section 2 presents the forecasting methodology. Models, estimation, and metrics and tests used to compute and assess the results are described in section 3. The data and setup are displayed in section 4. Section 5 reports the results and brings an economic discussion about them. Finally, section 6 concludes.

2 Forecasting Methodology

2.1 "Traditional" Inflation Forecasting

Let π_t be the (aggregate) inflation at period t. In practice, inflation is calculated from the percentage change in a price index constructed from a typical consumption basket. For forecasting purposes, assume that there are J predictors for inflation. Let $\boldsymbol{x}_t$ be a J-dimensional vector of these explanatory variables *observed* at t, that is, the set of information available to the econometrician to perform the forecasting. Notice that $\boldsymbol{x}_t$ can contain both the last available realizations of the predictor variables as well as lags of these variables. Lastly, let

$\mathcal{M}_{t,h}$ be a time-varying mapping between explanatory variables (predictors) and inflation h periods ahead. This mapping is obtained via estimation based on past data. As the mapping is usually obtained from moving windows, the mapping is conditional on time, which is indicated by the subscript t.

There are several possibilities to estimate the mapping $\mathcal{M}_{t,h}$. Initially, we choose between linear or non-linear specifications. Furthermore, in a rich-data environment, we can consider dimensionality reduction or shrinkage with or without selecting explanatory variables. Whatever the choices, we must be careful to avoid overfitting. Finally, an h-periods-ahead forecast is given by

$$\widehat{\pi}_{t+h\,|\,t} = \widehat{\mathcal{M}}_{t,h}(\boldsymbol{x}_t)$$

where hats indicate estimation.

2.2 Aggregation of Disaggregated Forecasts

Besides the general price index and its percentage change, the aggregate inflation π_t, now consider the availability of D disaggregated price indexes indexed by $i = 1, \ldots, D$. Let π_{it} be the percentage change of the disaggregation i at period t. Let ω_{it} be the weight of the disaggregation i in the general price index at period t. Note that these weights are time-varying since the composition of the representative consumption basket may change over time. The relationship between inflation and price changes in disaggregated indices is given by

$$\pi_t = \sum_{i=1}^{D} \omega_{it}\, \pi_{it}, \tag{2.1}$$

that is, the aggregate inflation is a weighted average of the "disaggregate inflations" (price changes) of each disaggregation.

Let $\widetilde{\boldsymbol{x}}_t$ be a $\widetilde{J}$-dimensional vector of explanatory variables for price changes *observed* at t by the econometrician. Note that it is expected that $\widetilde{J} > J$ since in the disaggregated case we potentially have more information: in addition to all the other explanatory variables available in the aggregated case, we can use the lagged price changes of the other disaggregations as explanatory variables for a specific disaggregation. Thus, the question arises whether it is now possible to capture and explore the crossed dependence between disaggregate prices to improve the inflation forecasting. Let $\mathcal{M}_{i,t,h}$ be a time-varying mapping between explanatory variables and h-periods-ahead price variation of the disaggregation i. Following (2.1), an h-periods-ahead forecast for the aggregate inflation is given by

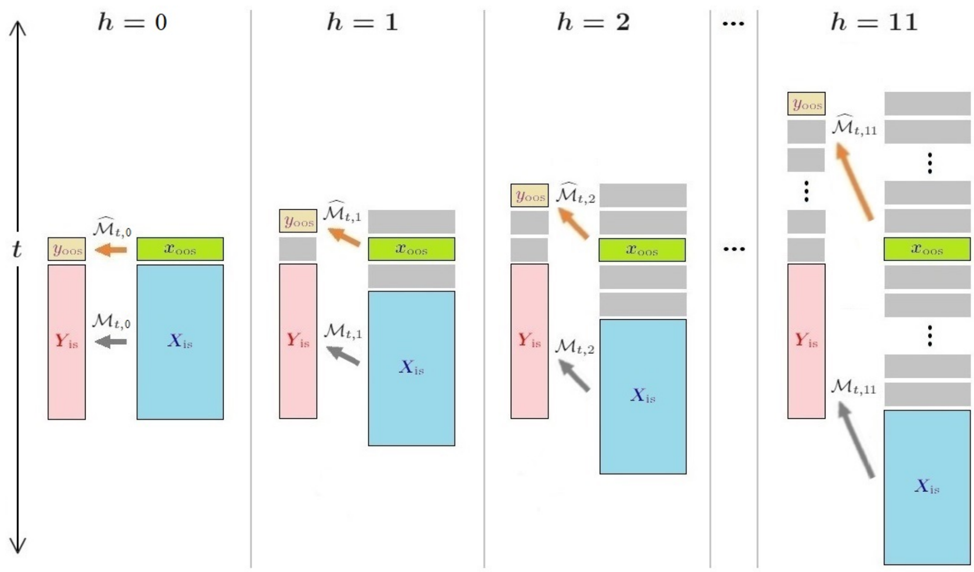
$$\widehat{\pi}_{t+h\,|\,t} = \sum_{i=1}^{D} \widetilde{\omega}_{it}\, \widehat{\pi}_{i,t+h} = \sum_{i=1}^{D} \widetilde{\omega}_{it}\, \widehat{\mathcal{M}}_{i,t,h}(\boldsymbol{x}_t)$$

where $\widetilde{\omega}_{it}$ is the weight of disaggregation i in the aggregate index *observed* at t by the econometrician, that is, the *last available* weight at period t and not the weight *evaluated* for the period t – which we previously indicated simply by ω_{it}.

2.3 Direct Forecasting Approach and Rolling Window Scheme

In this work, we employ a direct approach considering (mainly) 10-year rolling windows for each horizon $h \in \{0, 1, \ldots, 11\}$. We take the time-adjusted explanatory variables to fit the mapping between them and inflation in this approach. For example, suppose we want to generate a forecast for the current period ($h = 0$). In that case, we consider the most recently *available* information to estimate the desired mapping. On the other hand, if we want to compute an one-period-ahead forecast ($h = 1$), we consider the information available up to the previous period in which the forecast is estimated. We followed this reasoning until we calculated the eleven-period-ahead forecast in which we assumed the information *available* ten periods earlier. Figure 1 illustrates the exercise. After computing forecasts based on a given period, we roll the time window forward one period and repeat the estimation procedure followed by calculating forecasts for all horizons of interest.

Figure 1: Direct Forecasting Approach with Rolling Window Scheme



3 Models and Forecast Evaluation

3.1 Models

3.1.1 Benchmarks

Random Walk (RW). Considering the aggregated case, the forecast of the inflation h periods ahead at period t is given by current inflation, that is, $\widehat{\pi}_{t+h\,|\,t} = \pi_t$.

Historical Mean. Also for the aggregate case, the forecast h periods ahead is given by historical average inflation computed at t, that is,

$$\widehat{\pi}_{t+h\,|\,t} = \overline{\pi}_{t+h\,|\,t} = \frac{1}{S} \sum_{s\,=\,t-S+1}^{t} \pi_s$$

where S is the number of previously observed inflation measures (rolling window length).

Autoregressive model – AR(p). For both aggregated and disaggregated cases, in the direct forecasting approach, for each horizon h, we can be written a p-order AR model as

$$\pi_t = \mu + \sum_{l=1}^{p} \phi_l \, \pi_{t-h-l+1} + \varepsilon_t$$

where ε_t is an error term. The order p can be previously fixed or selected via some information criterion (e.g., BIC). Thus, the inflation forecast h periods ahead is given by

$$\widehat{\pi}_{t+h\,|\,t} = \widehat{\mu} + \sum_{l=1}^{p} \widehat{\phi}_l \, \pi_{t-h-l+1}.$$

where $\widehat{\mu}$ and $\widehat{\phi}$'s are least squares (OLS) estimates.

Augmented Autoregressive model. Including seasonal dummies and inflation expectation, we can write the model

$$\pi_t = \mu + \sum_{l=1}^{p} \phi_l \, \pi_{t-h-l+1} + \eta \; \pi^e_{t\,|\,t-h} + \sum_{m=1}^{11} \delta_m \, d_{mt} + \varepsilon_t \tag{3.1}$$

where $\pi^e_{t\,|\,t-h}$ is the inflation expectation for the period t available at $t-h$, d_{mt} is a seasonal dummy that assumes value 1 for month m, and δ_m is a coefficient associated with seasonal dummy d_{mt}. In this framework, we estimate the coefficients via OLS, and the forecast h periods ahead is given by

$$\pi_{t+h\,|\,t} = \widehat{\mu} + \sum_{l=1}^{p} \widehat{\phi}_l \, \pi_{t-l+1} + \widehat{\eta} \; \pi^e_{t+h\,|\,t} + \sum_{m=1}^{11} \widehat{\delta}_m \, d_{m,t+h}.$$

(Empirical) Phillips Curve. Following and *adapting* price-setting models such as those presented in Galí and Gertler (1999) and Blanchard and Galí (2007), we employ a forecasting model for the aggregate inflation based on a hybrid Phillips curve given by

$$\pi_t = \mu + \sum_{l=1}^{p} \phi_l \, \pi_{t-h-l+1} + \eta \, \pi^e_{t\,|\,t-h} + \theta_1 \, g_{t-h} + \theta_4 \, \Delta s_{t-h} + \varepsilon_t$$

where g_{t-h} is some economic activity measure *observed* at $t-h$, and Δs_{t-h} is an exchange rate measure *observed* at $t-h$. The forecast is computed via

$$\widehat{\pi}_{t+h\,|\,t} = \widehat{\mu} + \sum_{l=1}^{p} \widehat{\phi}_l \, \pi_{t-l+1} + \widehat{\eta} \, \mathbb{E}_t \pi_{t+h} + \widehat{\theta}_1 \, g_t + \widehat{\theta}_2 \, \Delta s_t$$

where $\left(\widehat{\mu}, \widehat{\boldsymbol{\phi}}, \widehat{\eta}, \widehat{\theta}_1, \widehat{\theta}_2\right)$ are OLS estimates.

3.1.2 Shrinkage

Ridge (with incomplete information). For disaggregated cases, we consider the Augmented AR model (3.1) with the addition of other lagged disaggregations:

$$\pi_{it} = \mu + \sum_{i'=1}^{D} \sum_{l=1}^{p} \phi_{i'l} \, \pi_{i',t-h-l+1} + \eta_i \, \pi^e_{t\,|\,t-h} + \sum_{m=1}^{11} \delta_{im} \, d_{mt} + \varepsilon_{it} \qquad i = 1, \ldots, D.$$

We estimate the coefficients employing Ridge estimator:

$$\left(\widehat{\mu}_i, \widehat{\boldsymbol{\beta}}_{\text{Ridge},i}(\lambda)\right) = \underset{\mu_i, \boldsymbol{\beta}_i}{\text{argmin}} \left\{ \frac{1}{T-h} \sum_{t=1}^{T-h} \left(\pi_{it} - \mu_i - \boldsymbol{\beta}_i \, \boldsymbol{z}_{t-h}\right)^2 + \lambda_i \sum_{j=1}^{J} \beta_{ij}^2 \right\}$$

where λ_i is a regularization parameter, $\boldsymbol{z}_{t-h}$ is a vector with all explained variables, and $\boldsymbol{\beta}_i$ is a vector of associated coefficients. The forecast h periods ahead is given by

$$\widehat{\pi}_{t+h\,|\,t} = \sum_{i=1}^{D} \omega_{it} \, \widehat{\pi}_{i,t+h\,|\,t} \qquad \text{with} \qquad \widehat{\pi}_{i,t+h\,|\,t} = \widehat{\mu}_i + \widehat{\boldsymbol{\beta}}_{\text{Ridge},i} \, \boldsymbol{z}_t.$$

adaLASSO (with full information). For all cases, consider the model with complete information:

$$\pi_{it} = \mu_i + \sum_{i'=1}^{D} \sum_{l=1}^{p} \phi_{i'l} \, \pi_{i',t-h-l+1} + \eta_i \, \pi^e_{t\,|\,t-h} + \sum_{m=1}^{11} \delta_{im} \, d_{mt} + \sum_{j=1}^{J} \sum_{l=1}^{p} \theta_{ijl} \, x_{j,t-h-l+1} + \varepsilon_{it}$$

where $\boldsymbol{x}_t$ is an expanded vector of potential explanatory variables for π_{it}. We estimate this model via adaptive LASSO (adaLASSO), introduced by (Zou, 2006):

$$\left(\widehat{\mu}_i,\, \widehat{\boldsymbol{\beta}}_{\text{adaLASSO}}(\lambda,\boldsymbol{\omega})\right) = \underset{\mu_i,\,\boldsymbol{\beta}_i}{\text{argmin}} \left\{ \frac{1}{T-h} \sum_{t=1}^{T-h} \left(\pi_{it} - \mu_i - \boldsymbol{\beta}_i\, \boldsymbol{x}_{t-h}\right)^2 + \lambda \sum_{j=1}^{J} \omega_{ij}\, |\beta_{ij}| \right\}$$

where λ is a regularization parameter, $\boldsymbol{x}_{t-h}$ is a vector with *all* explanatory variables, and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_J)$ is a vector of weights obtained previously via LASSO – the estimator that assumes $\omega_{ij} = 1$ for all j. More precisely, the weights are computed via

$$\omega_{ij} = \left(\left| \widehat{\beta}_{\text{LASSO},ij} \right| + \frac{1}{\sqrt{T}} \right)^{-1}.$$

As before, the forecast h periods ahead is given by $\widehat{\pi}_{i,t+h\,|\,t} = \widehat{\mu}_i + \widehat{\boldsymbol{\beta}}_{\text{adaLASSO},i}\, \boldsymbol{x}_t$.

3.1.3 Factor models

(Augmented) Factor Model. Considering that all regressors were normalized for both aggregate and disaggregate cases, this model can be described by

$$\boldsymbol{x}_t = \sum_{k=1}^{K} \boldsymbol{\lambda}_k\, f_{kt} + \boldsymbol{u}_t \tag{3.2}$$

$$\pi_{it} = \mu_i + \sum_{i'=1}^{D} \sum_{l=1}^{p} \phi_{i'l}\, \pi_{i',t-h-l+1} + \xi_i\, \pi^e_{t\,|\,t-h} + \sum_{m=1}^{11} \delta_{im}\, d_{mt} + \sum_{k=1}^{K} \beta_{ik}\, \widehat{f}_{k,t-h} + \varepsilon_{it}$$

from which we compute $\widehat{\boldsymbol{f}}_t = \left(\widehat{f}_{1t}, \ldots, \widehat{f}_{Kt}\right)$ and $\boldsymbol{\lambda}_k = (\lambda_{1k}, \ldots, \lambda_{Jk})$ combining Principal Component Analysis (PCA) and OLS. Finally, we compute $\left(\widehat{\mu},\, \widehat{\boldsymbol{\phi}},\, \widehat{\xi},\, \widehat{\boldsymbol{\beta}}\right)$ via adaLASSO. For identification purposes, we assume that

$$\mathbb{E}(\boldsymbol{f}_t\,|\,\boldsymbol{u}_t) = 0, \quad \text{Cov}(\boldsymbol{u}_t, \varepsilon_t) = 0, \quad \text{Var}(\boldsymbol{f}_t) = \boldsymbol{I}_K, \quad \text{and} \quad \text{Var}(\boldsymbol{u}_t) = \boldsymbol{\Omega} = \text{diag}\,(\sigma_1^2, \ldots, \sigma_p^2).$$

The number of factors K is selected via information criterion IC_{p2} of Bai and Ng (2002), and the forecast h periods ahead is given by

$$\widehat{\pi}_{i,\,t+h\,|\,t} = \widehat{\mu}_i + \sum_{i'=1}^{D} \sum_{l=1}^{p} \widehat{\phi}_{i'l}\, \pi_{i',t-l+1} + \widehat{\xi}_i\, \pi^e_{t+h\,|\,t} + \sum_{m=1}^{11} \widehat{\delta}_{im}\, d_{m,t+h} + \sum_{k=1}^{K} \widehat{\beta}_{ik}\, \widehat{f}_{kt}$$

where $\widehat{f}_{kt}$ is the k-th factor evaluated at t.

Target Factor Model. Proposed by Bai and Ng (2008), this model controls the participation of normalized explanatory variables in the factor construction. In a previous stage, for each explanatory variables indexed by $j = 1, \ldots, J$, we estimate

$$\pi_{it} = \zeta_i + \sum_{i'=1}^{D} \sum_{l=1}^{p} \gamma_{i'l}\, \pi_{i',t-l} + \eta_i\, \pi^e_{t\,|\,t-h} + \sum_{m=1}^{11} \rho_{im}\, d_{mt} + \theta_{ij}\, x_{j,t-h} + \nu_{it} \qquad i = 1, \ldots, D$$

and run the hypothesis test $\theta_{ij} = 0 \times \theta_{ij} \neq 0$ for some significance level α. If θ_{ij} is statistically different from zero, we employ x_j in the factor estimation. Let $\boldsymbol{x}_t(\alpha, i)$ be the set of selected variables for i-th disaggregation. Finally, we proceed as in the traditional (augmented) factor model: we perform

$$\boldsymbol{x}_t(\alpha, i) = \sum_{k=1}^{K} \boldsymbol{\lambda}_k \, f_{kt} + \boldsymbol{u}_t$$

from which $\widehat{\boldsymbol{f}}_t$ and $\widehat{\boldsymbol{\lambda}}_k$ are computed via PCA and OLS. Then we estimate the augmented (target) factor model via adaLASSO and compute the forecast as before.

3.1.4 FarmPredict: bridging factor and sparse models

Some idiosyncratic errors of the factor model, that is, some $\boldsymbol{u}_t$ entries in equation (3.2), can impact the price variation, which the common factor structure will not capture. Thus, defining $\widehat{\boldsymbol{u}}_t = \boldsymbol{x}_t - \sum_{k=1}^{K} \widehat{\boldsymbol{\lambda}}_k \, \widehat{f}_{k,t}$, a J-dimensional vector, we can introduce lags of $\boldsymbol{u}_t$ on the factor model:

$$\begin{aligned}\pi_{it} = \mu_i + \sum_{i'=1}^{D} \sum_{l=1}^{p} \phi_{i'l} \, \pi_{i',t-h-l+1} + \gamma_i \, \pi^{e}_{t\,|\,t-h} + \sum_{m=1}^{11} \delta_{im} \, d_{mt} \\ + \sum_{k=1}^{K} \sum_{l=1}^{p} \beta_{ikl} \, \widehat{f}_{k,t-h-l+1} + \sum_{j=1}^{J} \sum_{l=1}^{p} \theta_{ijl} \, \widehat{u}_{j,t-h-l+1} + \varepsilon_{it}.\end{aligned} \tag{3.3}$$

This model is a specific form of a general model called FarmPredict that has initially been introduced by Fan, Masini, and Medeiros (2021). Unlike what is proposed in the original paper, we estimate the "final equation" (3.3) with all regressors simultaneously employing adaLASSO. Next, we compute the forecast.

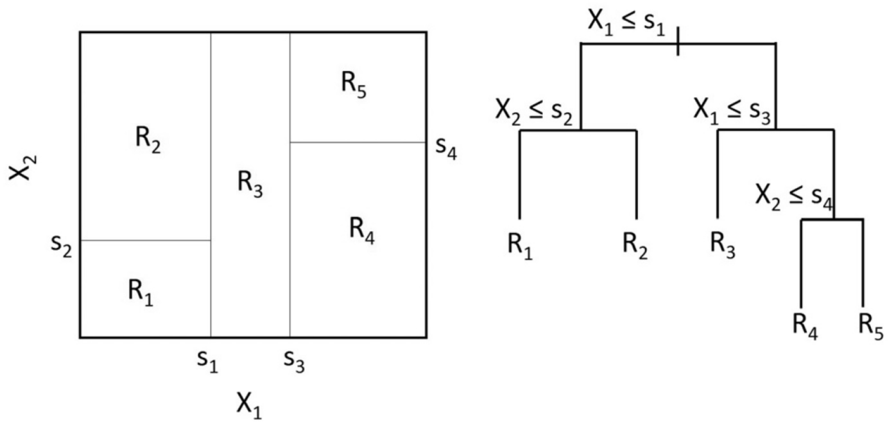
3.1.5 Complete Subset Regression (CSR)

Introduced by Elliott, Gargano, and Timmermann (2013, 2015), this ensemble method combines estimates from all (or several) possible linear regression models keeping the number of predictors fixed. Let p be the total available predictors and $k \leqslant p$ be the number of "selected" predictors (complete subsets). The CSR involves the estimation of $\frac{k!}{(k-p)!k!}$ linear models. Variables when "non-selected" has their coefficients set to zero. The final CSR estimate is the average of all estimates. Thus, subset regression has a shrinkage interpretation since when averaging parameters that sometimes assume zero value, this average generates a shrunken estimates of the coefficients, which can contribute to more accurate forecasts. Due to the high computational cost arising from a large number of predictors, we pre-selected $\widetilde{p}$ predictors based on a ranking of t-statistics in absolute value. This procedure is similar to that used in the target factor model. So, instead of considering all available p predictors, we run the CSR considering pre-selected $\widetilde{p}$ predictors.

3.1.6 Random Forest (RF)

Breiman (2001) introduced the Random Forest, a model that combines several based-tree regressions using bagging. A regression tree is a nonparametric model that approximates an unknown nonlinear function with local predictions via recursive partitioning, as illustrated in Figure 2.

Figure 2: Illustration of the Random Forest with two explanatory variables



Extracted from Medeiros, Vasconcelos, Veiga, and Zilberman (2021).

Formally, a regression tree model can be written as follows:

$$\pi_{it} = \sum_{k=1}^{K} c_k \, \mathcal{I}_k(\boldsymbol{x}_{t-h} \in R_k)$$

where $\mathcal{I}_k(\boldsymbol{x}_{t-h} \in R_k)$ is an indicator function that assumes the value 1 when $\boldsymbol{x}_{t-h}$ belongs to the k-th region R_k, and c_k is the average of π_t in this region. We have to set the minimum number of observations per region. Then, we obtain B trees by implementing a double draw: we draw on the observation dimension using block bootstrap, and we draw which variables are selected to estimate the tree. The idea is that this double draw will ensure the variability of the trees. Let K_b be the number of regions of the b-th tree, $b = 1, \ldots, B$. Lastly, the final forecast is given by the average of the forecasts obtained by each tree evaluated in the original data, that is,

$$\widehat{\pi}_{i,\, t+h\,|\,t} = \frac{1}{B} \sum_{b=1}^{B} \sum_{k=1}^{K_b} \widehat{c}_{k,b} \, \mathcal{I}_{k,b}(\boldsymbol{x}_{t-h} \in R_{k,b})$$

where $R_{k,b}$ is the k-th region of the b-th tree.

3.2 Evaluation: Metrics and Tests

3.2.1 Metrics

We use out-of-sample R-squared ($R^2_{\text{oos},h}$) and mean squared error (MSE) as main metrics to evaluate forecast performance. For each forecast horizon h, the first metric is described by

$$R^2_{\text{oos},h} = 1 - \frac{\sum_{t=1}^{T} \left(\pi_{t+h} - \widehat{\pi}_{t+h\,|\,t}\right)^2}{\sum_{t=1}^{T} \left(\pi_{t+h} - \overline{\pi}_t\right)^2}$$

where $\overline{\pi}_t = \frac{1}{S}\sum_{s=t-S}^{t-1} \pi_s$ is an average computed from the inflation observed in the rolling window ended at t. The higher the R2, the better the model's predictive performance. Notice that R^2_{oos} compares a model and the historical average. If $R^2_{\text{oos}} > 0$, then the forecast performance of the model is superior to the historical average. By construction, the R^2_{oos} of the historical average is zero.

The second metric, used in the Diebold-Mariano test explained following, is defined by

$$\text{MSE}_h = \frac{1}{T}\sum_{t=1}^{T} \left(\pi_{t+h} - \widehat{\pi}_{t+h\,|\,t}\right)^2.$$

The lower the MSE, the better the model's predictive performance. Notice that the MSE is contained in the second term of R^2_{oos} (multiplied by T). As this second term of R^2_{oos} is negative, the ordering of best models in the MSE is the inverse of the ordering in R^2_{oos}.

3.2.2 Test

To pragmatically assess the results, we consider the widely employed test developed by Diebold and Mariano (1995). Let $\widehat{v}_{t+h\,|\,t,m} = y_{t+h} - \widehat{y}_{t+h\,|\,t,m}$ be a forecast error of the model m. Let $g(\cdot)$ be a metric to be applied to $\widehat{v}_{t+h\,|\,t,m}$ (e.g., MSE). The Diebold-Mariano (DM) test statistic is given by

$$d_{m,m'} = \frac{1}{T}\sum_{t=1}^{T} \Big(g\left(\widehat{v}_{t+h\,|\,t,m}\right) - g\left(\widehat{v}_{t+h\,|\,t,m'}\right)\Big)$$

where m' indicates another model, a competitor (i.e., a benchmark model or forecast, for example). We will consider that the normality of DM statistics is likely a trustworthy approximation, including for model-based forecasts.

4 Data and Setup

The data cover the period from Jul/2006 to Jun/2021 – 15 years of monthly data. For aggregate inflation, we employ the monthly inflation measured by IPCA, the official Brazilian price index computed by Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística* – IBGE). For disaggregations, we consider all groups and subgroups of the IPCA. There are nine groups and 19 subgroups throughout the period analyzed. Namely, the nine groups are foods and beverages, housing, household goods, clothing, transport, health and personal care, personal expenses, education, and communications. Subgroups are subdivisions and, in some cases, the group itself. In addition, we use the disaggregation defined by the Central Bank of Brazil (BCB) based on IBGE data. The BCB disaggregation consists of administrated, non-tradables, and tradables items. The use of this last disaggregation is interesting because, in principle, it presents more economic intuition, which can contribute to better forecast performance.

We consider inflation expectations of the Central Bank of Brazil's Focus survey and lags of the predicted variables among the admissible predictors. To forecast a disaggregate, we consider the lags of the other disaggregates of the same disaggregation, which allows capturing potential lagged "cross-effects". The Focus survey has a daily frequency and contains inflation expectations formed by many economic agents (experts) for several horizons (months) ahead. Reflecting the opinion of experts, the Focus may contain inside information that is not available to the econometrician – hence the importance of considering this variable in our information set. We consider the latest available inflation expectation for the horizon of interest when we generate our forecast. Moreover, there are seventy-four other potential predictor variables divided into eleven categories: Prices and Money (12), Commodities Prices (4), Economic Activity (19), Employment (3), Electricity (4), Confidence (2), Finance (9), Credit (3), Government (12), and Exchange and International Transactions (6). In Appendix A, Table A1 presents a description of these variables as well as the transformations implemented to guarantee the stationarity of the series, and Figure A1 contains an illustration of the correlation between these variables.

The reference day to compute our forecasts is the 15th of each month. If this day is not a business day, we consider an immediately previous business day. For the results shown in section 5, we consider two lags for all explanatory variables, including the variables mentioned above, factors in factor models, idiosyncratic components in FarmPredict, and the lags of all disaggregates. The only exception is the factors in the target factor model for which we employ only one (target) factor. As mentioned in subsection 2.3, the main results were generated based on 10-year rolling windows. In this setup, we generate 59 one-period-ahead forecasts, 57 two-period-ahead forecasts, 55 three-period-ahead forecasts, up to 39 twelve-period-ahead forecasts. The decrease in forecasts by horizon is due to the increase in lags used as predictors and the lack of actual values to verify the predictive performance.

The regularization parameters (λ's) of the Ridge, LASSO, and adaLASSO are obtained via Bayesian Information Criterion (BIC). We restrict the number of possible selected variables by the ceiling of $\sqrt{T}$ to enforce discipline. The number K of latent factors in factor models is selected via Bai and Ng (2002) information criterion IC_{p2}. For CSR, we set $\widetilde{p} = 20$ (number of pre-selected predictors) and $p = 4$ (number of selected variables by CSR). For pre-selecting of both target factor and CSR models, we adopt the 5% significance level ($\alpha = 0.05$). In its turn, for the RF models, we allowed the trees to grow until five observations by leaf. We set the proportion of variables selected in each split to 1/3 and the number of bootstrap samples to 500 ($B = 500$). All settings are similar to those adopted by Garcia, Medeiros, and Vasconcelos (2017) and Medeiros, Vasconcelos, Veiga, and Zilberman (2021). Finally, to estimate the empirical Phillips curve, we use the Central Bank of Brazil's economic activity index (IBC-Br) and nominal BRL/USD exchange rate.

5 Results

5.1 Overview of the Results

Table 1 exhibits the results of the out-of-sample R-squared (R^2_{oos}) for all analyzed models, including forecast horizons from one to twelve months ahead and forecasts accumulated over 12 months. The Focus survey (Panel A) outperforms all other competitors considering the first four forecast horizons (short-term), which points to the difficulty of overcoming the survey-based forecasts in the short term. For example, while the survey-based forecasts present an R^2_{oos} of about 74% for the first horizon, the adaLASSO for aggregates (Panel B) and disaggregation into IBGE groups (Panel D) register values of about 69%. For horizons from two to four months, the performance of Focus drops considerably compared to the first horizon, but it still maintains its superiority over all models. However, the performance of adaLASSO employing Central Bank of Brazil (BCB) disaggregation (Panel C) is closer to the Focus expectations between two and four months ahead. In general and not just for shorter horizons, further increasing the level of disaggregation (including no economic intuition) does not guarantee improvements. The poor performance of the disaggregation into IBGE subgroups corroborates this conclusion (see Panel E). However, we will see better employment for this disaggregation in subsection 5.4.

The results change considerably for horizons from five months onwards. The CSR considering aggregates perform relatively well for horizons from four to eight months ahead (mid-term). This result is similar to that found by Garcia, Medeiros, and Vasconcelos (2017). However, the Diebold-Mariano (henceforth, DM) test does not report statistical superiority compared to the Focus survey. From the sixth to the ninth horizons, the factor model and FarmPredict, both considering the BCB disaggregates, perform well, outperforming the Focus survey. The DM test attests to the predictive superiority of the FarmPredict about the Focus for horizons from seven to nine.More isolated, the FarmPredict using inflation disaggregated into IBGE groups was statistically superior to the Focus for the tenth horizon.

Table 1: Out-of-sample R-squared (%) – Survey and all models

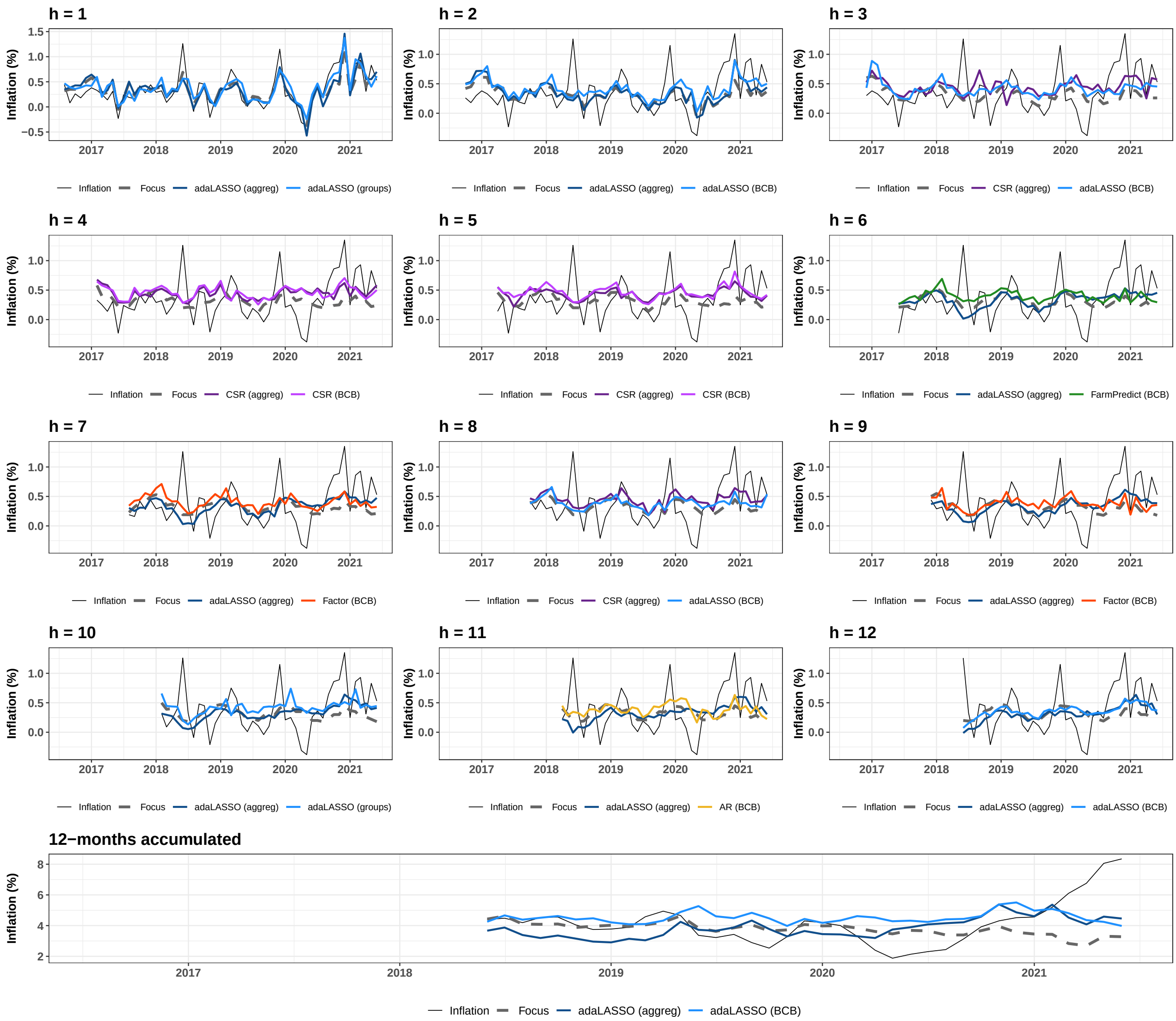
Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	12 m
A. Survey													
Focus	**74.55**	**24.26**	**10.61**	**9.25**	9.19	7.23	1.21	-1.18	-2.34	-1.03	-1.58	-1.74	*51.56*
B. Aggregate inflation													
RW	-7.64	-43.84	-39.78	-72.92	-86.10	-83.48	-76.67	-51.31	-32.75	-43.02	-64.90	-94.71	-332.59
Historical Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AR	15.16	4.33	3.86	-5.77	-8.37	-10.66	-6.44	-1.77	0.11	-2.95	-10.38	-18.73	0.75
HNKPC	66.37	13.38	4.84	-12.34	-5.48	-3.83	-3.12	-5.52	-3.60	-2.60	-2.36	-9.00	46.62
Augmented AR	63.41	9.91	-5.34	-14.89	-20.51	-18.14	-16.87	-13.93	-16.01	-18.45	-20.27	-18.71	7.02
adaLASSO	*69.18*	*23.44*	1.90	2.55	5.32	6.60	10.52	*6.90*	*8.07**	*9.25**	4.08	*6.19*	**60.61**
Factor	68.00	17.62	3.51	3.86	5.06	2.38	-2.56	-4.82	-5.27	-4.03	-0.85	-5.29	44.15
FarmPredict	68.60	15.42	-0.35	2.15	7.97	2.98	-6.67	-0.80	1.13	-1.39	*4.59*	-0.71	45.49
Target Factor	66.79	8.93	-4.01	3.38	-15.75	2.84	*11.34*	-4.44	-20.39	-22.97	-5.73	-21.80	42.49
CSR	36.84	7.69	5.00	6.49	**12.15**	**13.21**	8.77	**11.82**	6.63	-6.17	-4.22	-10.23	31.01
Random Forest	45.10	7.88	-3.88	-5.95	1.51	1.18	-3.43	-3.04	-4.11	-1.46	-0.13	0.64	23.49
C. Disaggregation: tradable, nontradable and monitored prices (BCB)													
AR	17.58	0.73	-6.98	-15.36	-15.55	-12.14	-6.19	-8.10	0.76	1.82	**9.52**	4.76	-4.93
Augmented AR	58.74	7.46	-12.49	-15.66	-18.48	-19.85	-21.29	-17.00	-22.34	-23.88	-13.33	-14.67	-1.27
Ridge	65.65	8.31	1.94	-1.84	0.10	1.09	-3.30	-2.32	-4.94	-5.11	3.47	2.00	20.26
adaLASSO	65.78	*21.30*	*8.51*	*7.28*	4.18	7.69	7.62	4.26	-0.60	-8.01	3.83	**9.96***	*53.13*
Factor	63.72	18.33	*7.90*	4.90	*10.93*	*11.46*	*11.54*	4.81	*9.90**	-1.70	-1.09	-1.70	47.03
FarmPredict	61.76	16.07	3.78	4.24	6.83	7.25	**11.72***	*9.61***	**12.38****	3.92	*5.98*	-0.70	41.19
Target Factor	64.31	13.87	-18.96	-18.09	-8.32	-0.76	-22.71	-9.28	-40.87	-17.60	-15.52	-10.46	35.24
CSR	25.19	8.50	1.56	-2.53	*10.89*	3.78	-2.67	-9.85	-11.72	-2.67	-6.91	-14.05	-1.70
Random Forest	32.75	-0.53	-5.14	-3.22	-2.21	-8.11	-13.34	-13.13	-16.25	-13.25	-5.29	-1.35	4.46
D. Disaggregation: groups (IBGE)													
AR	19.05	-1.50	-5.27	-5.51	-3.62	-2.51	-0.51	-1.28	-4.49	-8.00	-4.09	-3.77	-0.57
Augmented AR	61.66	7.62	-10.30	-7.85	-13.00	-17.85	-18.25	-14.62	-19.68	-21.33	-11.98	-12.04	12.32
Ridge	9.59	-1.34	-1.77	-6.85	-4.10	-3.03	-5.87	-8.09	-7.21	-4.04	-0.86	-4.04	4.32
adaLASSO	*69.30*	16.54	6.30	2.59	3.43	4.21	-3.68	-1.57	4.87	3.92	-4.22	5.29	48.92
Factor	60.49	8.57	5.98	*7.94*	6.84	5.11	-4.50	-9.92	4.45	-2.32	-4.33	0.64	39.19
FarmPredict	58.85	12.90	6.57	6.20	5.77	-0.06	-1.66	-4.45	7.44	**11.30***	-1.47	-1.87	41.26
Target Factor	61.38	14.96	-19.78	1.92	1.74	*10.62*	8.03	1.28	-7.85	-36.39	-10.28	-14.82	49.20
CSR	36.42	5.89	5.36	5.94	8.29	8.27	2.20	0.11	0.76	3.20	-6.19	5.11	26.74
Random Forest	25.91	-2.81	-4.68	-2.68	-4.17	-5.55	-8.19	-15.20	-16.42	-11.61	-2.71	3.05	-0.48
E. Disaggregation: subgroups (IBGE)													
AR	18.79	-3.18	-6.10	-7.71	-7.44	-4.92	-3.37	-3.23	-5.09	-6.77	-5.28	-6.13	4.60
Augmented AR	59.17	3.56	-12.94	-10.54	-17.20	-21.10	-20.52	-15.97	-21.55	-22.70	-14.31	-14.60	10.97
Ridge	-3.64	-3.70	-0.87	-6.23	-7.45	-4.77	-5.63	-6.03	-6.86	-8.29	-2.88	-0.57	-10.67
adaLASSO	55.70	8.33	-2.91	-2.59	-0.58	-3.66	1.79	-0.82	1.72	2.19	-6.40	1.39	39.23
Factor	55.08	2.63	-5.99	-3.18	-3.82	-5.72	-1.54	-1.69	-6.85	0.33	-10.38	-11.51	30.15
FarmPredict	56.48	8.25	-0.93	-0.85	-2.87	-8.12	3.35	1.26	-1.41	*9.00*	-0.74	-10.05	35.12
Target Factor	60.10	6.13	-47.73	6.82	-14.59	4.29	-17.24	-8.30	-15.28	-37.25	-31.10	-14.39	38.43
CSR	38.16	5.96	0.96	-6.52	3.78	0.10	-6.60	-5.81	-7.08	-3.03	-9.25	3.49	15.60
Random Forest	26.73	-6.41	-9.66	-11.97	-5.86	-8.09	-13.53	-23.27	-18.62	-10.61	-3.15	*5.67*	-6.76

Notes: ***, **, and * indicate that a specific model m performed statistically better than the Focus Survey (benchmark) in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}(\widehat{\pi}^{m}_{t+h\,|\,t}) = \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t}) \times \mathbb{H}_1 : \text{MSE}(\widehat{\pi}^{m}_{t+h\,|\,t}) < \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t})$. The highlighted value in blue bold indicates the best model for each horizon in terms of R^2_{oos}, and the values in blue italics indicate the second and third best models.

Finally, considering the long-term (nine months onwards) and 12-month cumulative forecasts, the adaLASSO models based on aggregates and BCB disaggregates performed statistically better than Focus for horizons nine, ten, and twelve. For the eleventh horizon, the AR model considering BCB disaggregates incidentally delivered the best result; however, the DM test did not attest to statistical superiority under the traditional significance levels. Despite presenting the best out-of-sample R-squared in the 12-month cumulative forecasts, adaLASSO with aggregates ($R^2_{\text{oos}} \approx 60.5\%$) and BCB disaggregates ($R^2_{\text{oos}} \approx 53\%$) were not statistically superior to the Focus survey ($R^2_{\text{oos}} \approx 51.5\%$). The p-value of the one-tailed DM test for adaLASSO using aggregates was 0.113. Thus, this model almost performed statistically better than the benchmark Focus on 12-month accumulated inflation expectations considering the more slack significance level (e.g., 0.10). No other model outperforms the Focus survey in forecasting 12-month cumulative inflation.

Figure 3 presents the temporal evolution of actual inflation, Focus survey expectations, and the best both aggregate- and disaggregates-based models. Looking at the projections for $h = 1$, we partially understand why it is not easy to outperform the survey in the very short term. The Focus forecast is very close to the realized value. Furthermore, as we will see in subsection 5.2, the inflation expectation (survey) available when the forecasts are computed is the primary predictor variable in models for the first horizon. The survey carries much information not available to the econometrician, so the result is already expected.

Figure 3: Forecasts by the horizon and 12-month accumulated



Notes: Black solid lines indicate the actual inflation. Gray dashed lines indicate the Focus survey inflation expectations on the 15th day of each month (or the last working day immediately preceding). Blue, purple, orange, green, and solid yellow lines indicate the forecasts generated by adaLASSO, CSR, Factor, FarmPredict, and AR models, respectively.

When analyzing the other horizons ($h \geqslant 2$), it is clear that models and the survey are challenging to predict peaks and valleys. Already at $h = 2$, many errors occur. Moreover, the period from March 2020 was marked by the coronavirus pandemic in Brazil, a very uncertain and challenging time for economic forecasts. One consequence was that, in this period, the Focus survey initially overestimated inflation and afterward systematically underestimated one. Despite the difficulty, some models perform better than the expert forecasts from the end of 2020 onwards for all horizons except the first. For example, adaLASSO using aggregates and BCB disaggregations as well as other models, achieved a great result in forecasting the peak in December 2020 at $h = 2$ when the available inflation expectation was far from the value realized. Thus, other variables besides the available inflation expectation were necessary for the results of the models that performed better than the survey.

5.2 Forecast of Disaggregates and Variable Selection

To interpret the performance of the best models for aggregates and disaggregates, Table 2 displays the R^2_{oos} of adaLASSO and FarmPredict for aggregates, aggregation of BCB groups, and disaggregates themselves. Notice that the adaLASSO for aggregates generally performs numerically better than the adaLASSO aggregating disaggregated forecasts, with statistical superiority at $h = 10$ according to the DM test. On the other hand, the FarmPredict, in general, performs better for aggregation of disaggregated forecasts. In addition to its numerical superiority about the direct employment of the aggregate, there is statistical superiority of the FarmPredict using BCB disaggregates at horizons seven and nine. The FarmPredict with aggregates is statistically superior to FarmPredict with disaggregates only in the first horizon. Before proceeding to the analysis of the forecast performance of the disaggregates, note that aggregating disaggregated forecasts do not generate a "convex combination" of the out-of-sample R-squares. Thus, the R^2_{oos} of the aggregation can even be greater than the out-of-sample R-squares of the disaggregates.

The good performance of the FarmPredict with BCB disaggregations, as well as the result of the adaLASSO using the same disaggregations, is mainly due to the good forecast performance of these methods for the non-tradables price variation. Both FarmPredict and adaLASSO record approximately 50% of R^2_{oos} in predicting this disaggregate for most horizons. Including, the forecast performance presented a tendency for improvement as the forecast horizon increases. Moreover, non-tradables registered an average share of 41.6% in the aggregate index in the projection period, the largest share since both administered and tradables prices had shares of 25.5% and 32.9%, respectively. On the other hand, both methods have difficulty forecasting the variation of administered prices. Their forecasts for this disaggregation perform worse than the historical mean for all horizons from the second one (the R^2_{oos}'s are negatives in all these cases). Finally, adaLASSO registered positive out-of-sample R squares in all horizons except the third and was consistently superior to FarmPredict in forecasting tradables inflation.

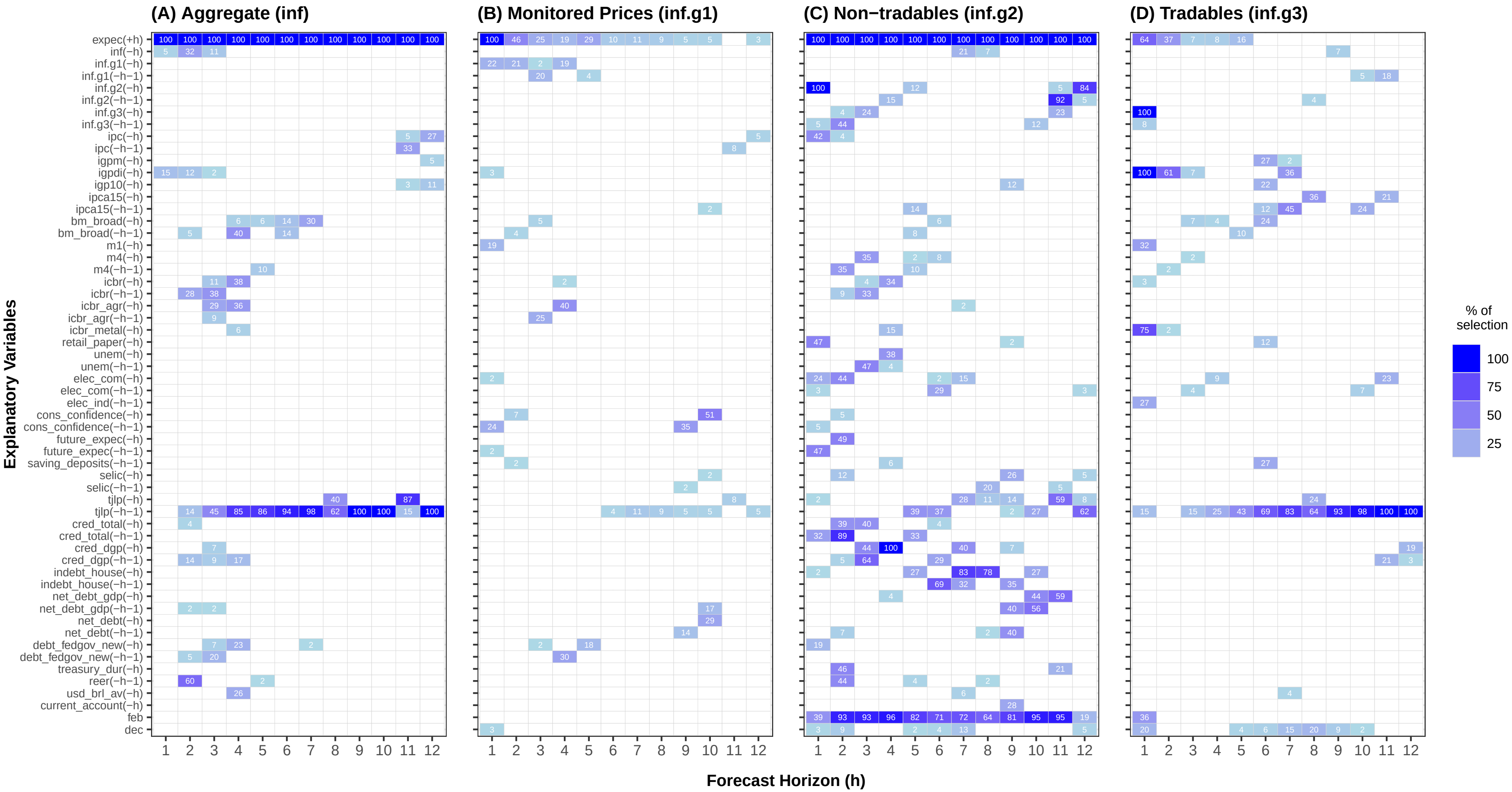
Table 2: Out-of-sample R-squared (%) – BCB disaggregation: adaLASSO and FarmPredict

Predicted variable	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$
						A. adaLASSO						
Aggregate	69.18	23.44	1.90	2.55	5.32	6.60	10.52	6.90	8.07	9.25**	4.08	6.19
Aggregating disaggregates	65.78	21.30	8.51	7.28	4.18	7.69	7.62	4.26	-0.60	-8.01	3.83	9.96
Monitored Prices (inf.g1)	30.92	-5.85	-9.07	-6.57	-5.65	-0.20	-7.17	-2.44	-1.34	-6.48	-9.44	-10.32
Non-tradables (inf.g2)	37.62	49.04	48.21	51.17	50.83	51.58	54.23	55.31	53.55	48.12	49.45	52.26
Tradables (inf.g3)	43.50	3.62	-4.24	0.55	3.05	6.90	9.69	8.95	12.10	6.71	9.23	14.26
						B. FarmPredict						
Aggregate	68.60*	15.42	-0.35	2.15	7.97	2.98	-6.67	-0.80	1.13	-1.39	4.59	-0.71
Aggregating disaggregates	61.76	16.07	3.78	4.24	6.83	7.25	11.72**	9.61	12.38*	3.92	5.98	-0.70
Monitored Prices (inf.g1)	28.88	-7.51	-8.92	-7.82	-7.93	-17.46	-14.83	-11.56	-0.79	-7.65	-16.42	-11.01
Non-tradables (inf.g2)	36.90	49.66	42.22	47.01	47.84	50.65	52.39	58.93	48.26	51.71	54.68	60.12
Tradables (inf.g3)	35.29	-4.75	-5.70	-2.01	0.18	0.51	3.92	-1.58	3.24	3.34	0.33	-9.79

Notes: ***, **, and * indicate that, excluding disaggregates, one adaLASSO (or FarmPredict) performed statistically better than other adaLASSO (or FarmPredict) in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}\big(\widehat{\pi}^{\text{adaLASSO 1}}_{t+h\,|\,t}\big) = \text{MSE}\big(\pi^{\text{adaLASSO 2}}_{t+h\,|\,t}\big) \times \mathbb{H}_1 : \text{MSE}\big(\widehat{\pi}^{\text{adaLASSO 1}}_{t+h\,|\,t}\big) < \text{MSE}\big(\pi^{\text{adaLASSO 2}}_{t+h\,|\,t}\big)$ (or $\mathbb{H}_0 : \text{MSE}\big(\widehat{\pi}^{\text{FarmPredict 1}}_{t+h\,|\,t}\big) = \text{MSE}\big(\pi^{\text{FarmPredict 2}}_{t+h\,|\,t}\big) \times \mathbb{H}_1 : \text{MSE}\big(\widehat{\pi}^{\text{FarmPredict 1}}_{t+h\,|\,t}\big) < \text{MSE}\big(\pi^{\text{FarmPredict 2}}_{t+h\,|\,t}\big)$, where numbers 1 and 2 indicate "aggregate" or "aggregating disaggragates". The average weights of each disaggregation in the aggregate from 07/2016 to 06/2018 (forecast period) were: (i) monitored prices: 25.5%; (ii) non-tradables: 41.6%; (iii) tradables: 32.9%.

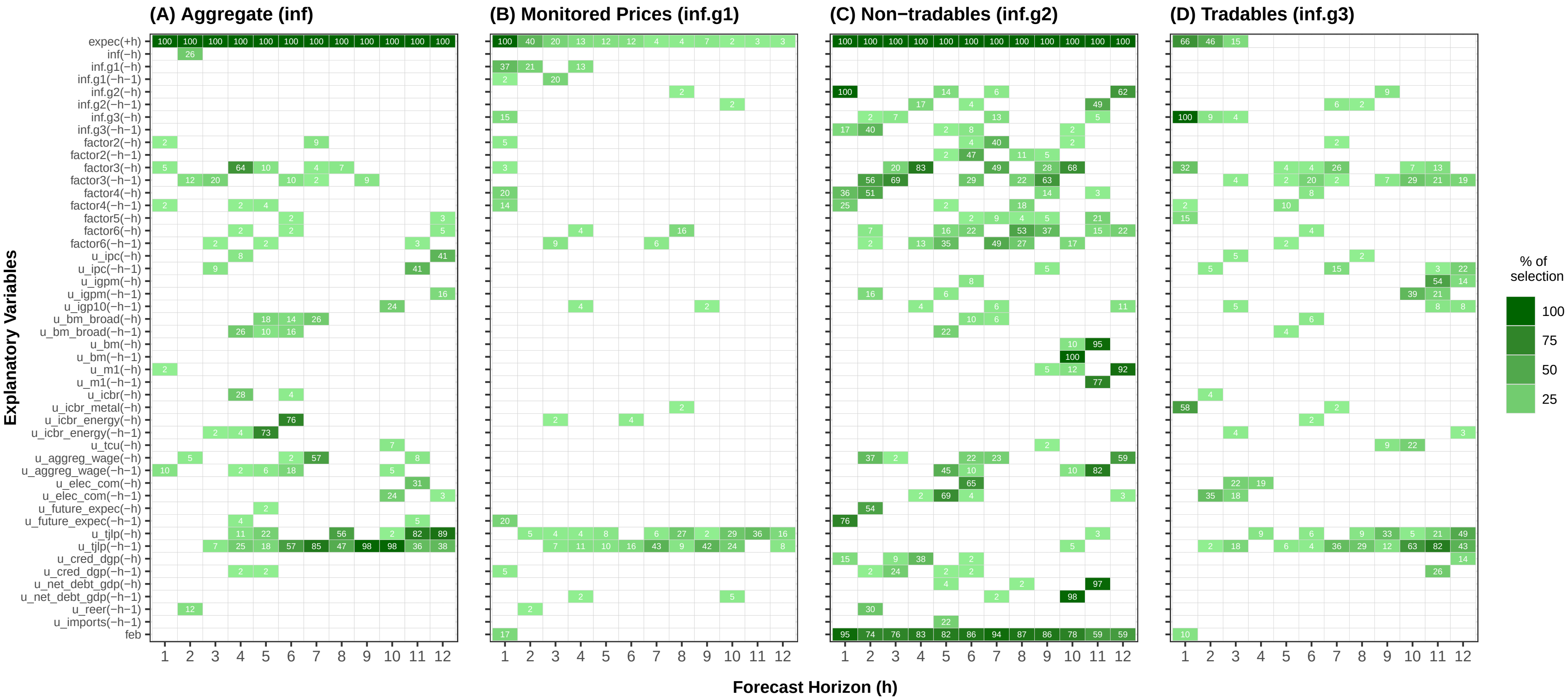
Looking at the forecasting performance of disaggregates leads to two questions. The first is that combining different models forecasting different disaggregations *can improve* (but not necessarily improve) aggregate forecasting via aggregation of disaggregated forecasts. This question is entirely empirical and is explored in subsection 5.4 following. The second question is to explore potential economic intuitions for the results of both aggregated and disaggregated analyses. To compare what is behind the two approaches, the variable selections of the adaLASSO and FarmPredict, both in aggregated and disaggregated cases, are shown in Figures 4 and 5. In each Figure, panel A brings the predictors selected by each method to forecast the aggregate directly. Panels B, C, and D display the predictors selected to predict price variation of administrated, non-tradable, and tradable items. To make the presentation viable, we restricted ourselves to the variables chosen 20% of the time in at least one forecast horizon. Variables definitions are shown in Table A1 in Appendix A. The prefix "u_" that appears in some variables in Figure 5 indicates that the variable had the common factors "discounted" and, therefore, only its idiosyncratic component was left. These variables are represented by $\widehat{u}_j$ in equation (3.3).

Figure 4: adaLASSO selection – aggregate inflation and BCB disaggregates (% of rolling windows)



Notes: We cut out variables that did not exceed 20% of selection at least one forecast horizon. The definitions of the variables are in Table A1.

Figure 5: FarmPredict selection – aggregate inflation and BCB disaggregates (% of rolling windows)



Notes: We cut out variables that did not exceed 20% of selection at least one forecast horizon. The definitions of the variables are in Table A1.

We can summarize the results of the variable selections in the following topics:

1. *High variability in the type of selected predictors*. All groups of variables displayed in Table A1 in Appendix A are represented in the adaLASSO selection. Even controlling for common factors (FarmPredict), most of the groups of variables are still represented. This result shows the importance of considering a broad set of information.

2. *Low participation of economic activity variables*. Variables related to economic activity appear little in the adaLASSO selection and almost do not appear when we control for common factors via FarmPredict. Do other variables contain information on economic activity? When we control for them, do the activity variables cease to be relevant, or do they not matter to forecasting inflation? A very profound answer to these questions is beyond the scope of this work. However, Figure A1 in Appendix A presents a correlation map that indicates that retail-related variables are correlated with monetary base and M1 and M2 money supplies, variables occasionally selected. In addition, there is no indication of any other indirect effect. Thus, the most likely answer is that economic activity variables are of little relevance to forecasting Brazilian inflation on any horizon ahead. The rare exceptions are paper sales (`retail_paper`), at $h = 1$, and unemployment rate (`unem`), at $h = 3$ and $h = 4$, for non-tradables.

3. *Available inflation expectations (survey) are frequently picked*. For all horizons, the inflation expectation (`expec`) is selected 100% of the time for aggregates and non-tradables. In the case of administrated and tradable items, the selection of the expectation decreases as the horizon increases, appearing only in the short term for tradables. It is important to note that the expectation is about aggregate inflation, so it is reasonable that it is not relevant to explain some specific disaggregations.Due to its greater share, non-tradables present a more remarkable similarity with aggregate inflation than the other breakdowns.

4. *Price variables (including commodities) are often chosen*. The variable selection shows that the various Brazilian inflation indices and commodities price variations are relevant to forecast inflation measured by the Brazilian official price index (IPCA). In other words, the various indices and price variations carry much information to forecast Brazilian inflation. In addition, due to past inflationary history, Brazil has several monthly indexes that cover different periods (e.g., days 1 to 30, 11 to 10, and 21 to 30). Undoubtedly, we must consider this information when forecasting Brazilian inflation.

5. *Factors that explain most of the variability of predictors are not always more relevant to forecasting inflation.* Interestingly, the common factor that explains most of the variability of the predictors (`factor1`) is rarely selected to forecast aggregate inflation or any disaggregation. Instead, the following factors up to the seventh are chosen for various horizons, mainly to predict non-tradables, lesser extent for aggregates and tradables, and very little for administrated ones.

6. *Brazilian long-term interest rate has great explanatory power for both aggregate and tradables inflation in the mid- and long-term.*

7. *Non-tradables record the more rich structure of predictors*. The good predictability of non-tradables is potentially related to its admits many predictors. Note that non-tradables have more predictors than aggregate inflation itself. Then this highlights the importance of looking at disaggregates. The relevance of the February dummy for forecasting inflation, for example, only appeared when we considered the forecast of non-tradables inflation. In this case, this dummy is crucial because it captures the variation in education prices, a sector whose contracts are usually redefined in January in Brazil.

5.3 Does the average between models and *available* survey matter?

The forecasts analyzed in this subsection are computed using arithmetic average between forecasts generated by selected models and Focus expectations *available* when the econometrician computes its forecasts, that is,

$$\widehat{\pi}^{\text{ave},\,m}_{t+h\,|\,t} = \frac{\widehat{\pi}^{m}_{t+h\,|\,t} + \pi^{aval.\ \text{Focus}}_{t+h\,|\,t}}{2}, \tag{5.1}$$

where m indicates a model among which we consider adaLASSO, FarmPredict, and CSR for aggregate and disaggregation into BCB and IBGE groups. The idea is to investigate whether the mathematical handling (5.1) leads to improvements in forecast performance. Table 3 displays the results in terms of out-of-sample R-squared (R^2_{oos}). To facilitate the comparison, we keep in Table 3 part of the results shown in Table 1.

The most remarkable result is the significant improvement in the short-term forecasting performance that the average generates for *all* models considering aggregate inflation and disaggregations. For the first horizon, for example, the average between adaLASSO considering the disaggregation into IBGE groups and available Focus expectations generates an R^2_{oos} close to the Focus expectation at reference day (74.50% at Panel G against 74.55% at Panel A). Since overcoming the survey one period ahead is very difficult, this "tie" is a very expressive result. Regarding other short-term forecasts, averages with adaLASSO using aggregates (for two-period-ahead) and adaLASSO using BCB disaggregations (for three-period-ahead) were statistically superior to the Focus survey at the 10% level according to the one-tailed DM test. Even numerical superiority was not achieved without computing the average (see Panels B, D, and F against Panel A for $h = 2, 3, 4$). Now, *all* three best numerical results in terms of R^2_{oos} for each horizon from two to five are obtained via averaging (see Panels C, E, and G).

For longer horizons ($h \leqslant 7$), the best performances continue to be from "pure" models, especially FarmPredict using BCB disaggregation. Nonetheless, the improvement generated for the initial horizons leads the average for adaLASSO employing aggregates to be statisti-

Table 3: Out-of-sample R-squared (%) – Average between Models and (available) Focus

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	12 m
A. Survey													
Focus	**74.55**	24.26	10.61	9.25	9.19	7.23	1.21	-1.18	-2.34	-1.03	-1.58	-1.74	51.56
B. Aggregate inflation													
adaLASSO	69.18	23.44	1.90	2.55	5.32	6.60	*10.52*	6.90	*8.07**	*9.25**	4.08	*6.19*	**60.61**
FarmPredict	68.60	15.42	-0.35	2.15	7.97	2.98	-6.67	-0.80	1.13	-1.39	4.59	-0.71	45.49
CSR	36.84	7.69	5.00	6.49	12.15	**13.21**	*8.77*	**11.82**	6.63	-6.17	-4.22	-10.23	31.01
C. Average between aggregate inflation models and (available) Focus survey													
adaLASSO & Focus	*72.80*	*30.19**	11.52	8.70	9.20	5.10	6.52*	4.27	1.31	4.13	4.95	2.86	*58.64***
FarmPredict & Focus	72.37	25.76	10.04	7.99	11.15	4.05	-2.45	-0.25	-2.51	-1.92	4.01	-1.26	50.44
CSR & Focus	62.48	24.57	*16.86*	*12.38*	*13.09*	*11.14*	8.14	*9.34*	1.36	-2.20	2.15	-3.95	54.24
D. Disaggregation: tradable, nontradable and monitored prices (BCB)													
adaLASSO	65.78	21.30	8.51	7.28	4.18	7.69	7.62	4.26	-0.60	-8.01	3.83	**9.96***	53.13
FarmPredict	61.76	16.07	3.78	4.24	6.83	7.25	**11.72***	*9.61***	**12.38****	3.92	**5.98**	-0.70	41.19
CSR	25.19	8.50	1.56	-2.53	10.89	3.78	-2.67	-9.85	-11.72	-2.67	-6.91	-14.05	-1.70
E. Average between BCB disaggregation models and (available) Focus survey													
adaLASSO & Focus	72.69	**30.25**	*16.59**	*12.29*	8.67	5.18	4.11	2.84	-2.41	-3.41	*5.13**	4.89*	*57.51*
FarmPredict & Focus	70.96	26.28	12.55	9.10	9.34	5.09	6.34	4.61	3.12	2.01	*5.59**	-0.38	50.52
CSR & Focus	59.50	26.80	15.76	8.04	**14.41**	7.72	3.09	-1.75	-5.95	0.87	1.74	-4.42	44.94
F. Disaggregation: groups (IBGE)													
adaLASSO	69.30	16.54	6.30	2.59	3.43	4.21	-3.68	-1.57	4.87	3.92	-4.22	*5.29*	48.92
FarmPredict	58.85	12.90	6.57	6.20	5.77	-0.06	-1.66	-4.45	*7.44*	**11.30***	-1.47	-1.87	41.26
CSR	36.42	5.89	5.36	5.94	8.29	*8.27*	2.20	0.11	0.76	3.20	-6.19	5.11	26.74
G. Average between IBGE groups disaggregation models and (available) Focus survey													
adaLASSO & Focus	*74.50*	*28.79*	15.15	10.17	9.08	4.43	-0.13	0.96	0.82	3.47	0.91	3.91	56.90
FarmPredict & Focus	69.70	25.61	14.92	10.92	10.40	2.26	1.38	-0.82	1.31	*6.07*	1.19	0.15	51.38
CSR & Focus	63.47	23.82	**16.95**	**12.58**	*12.56*	7.80	3.81	3.95	-1.30	2.16	1.21	5.10	54.82

Notes: ***, **, and * indicate that a specific model m performed statistically better than the Focus Survey (benchmark) in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}(\widehat{\pi}^{m}_{t+h\,|\,t}) = \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t}) \times \mathbb{H}_1 : \text{MSE}(\widehat{\pi}^{m}_{t+h\,|\,t}) < \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t})$. The highlighted value in blue bold indicates the best model for each horizon in terms of R^2_{oos}, and the values in blue italics indicate the second and third best models.

cally superior to the survey on forecasting accumulated inflation 12 months ahead. Despite recording only the second-best R^2_{oos} in the 12-month accumulated period, this model was the only statistically superior to the Focus at the moment, according to the DM test. Even though the "pure" adaLASSO using aggregates obtained the highest R^2_{oos}, it was not statistically superior to the Focus (the p-value of the one-tailed DM test was 0.113). The average between adaLASSO using BCB disaggregation and available Focus achieved the third-best result, but it was not statistically superior to the Focus (p-value 0.137). These results suggest that merging models with different performances for different horizons can be interesting to increase the forecast performance for inflation accumulated during the 12 months. We exhibit results in this direction in the following subsection 5.4.

5.4 Model Combinations: disaggregates and *all* models

In this subsection, we examine the performance of two different model combinations. The first combines different models forecasting different disaggregations to obtain the aggregate forecast. Based on the previous predictive performance, we choose the disaggregate forecasts to compute the current aggregate forecast. We adopted the minor out-of-sample squared error (SE) at the previous period as the criterion. Mathematically, for each period t and horizon h,

$$\widehat{\pi}^{\text{comb. (1)}}_{t+h\,|\,t} = \sum_{i=1}^{D} \widetilde{\omega}_{it} \widehat{\pi}^{*}_{i,\,t+h\,|\,t} \qquad \text{s.t.}$$

$$\forall i, \quad \widehat{\pi}^{*}_{i,\,t+h\,|\,t} \,:\, \text{SE}\left(\widehat{\pi}^{*}_{i,\,t+h-1\,|\,t-1}\right) \leqslant \text{SE}\left(\widehat{\pi}^{m}_{i,\,t+h-1\,|\,t-1}\right), \quad \forall m \in \mathcal{M},$$

where star indicates the model that generates the smallest forecasting squared error, $\text{SE}\left(\widehat{x}\right) = (x - \widehat{x})^2$, for a given disaggregation at the immediately previous period. The weights $\widetilde{\omega}_{it}$ are those *available* at the moment in which the aggregate forecast is computed. For each disaggregation i, there are nine models in the set of models $\mathcal{M}$.

Considering *all* models analyzed in this paper, the second model combination picks out the better model in terms of predictive performance in the previous period. We including model combinations based on disaggregates, all averages with the available Focus, and the Focus survey in isolation. Mathematically, for each period t and horizon h, we have

$$\widehat{\pi}^{\text{comb. (2)}}_{t+h\,|\,t} = \widehat{\pi}^{*}_{t+h\,|\,t} \,:\, \text{SE}\left(\widehat{\pi}^{*}_{t+h-1\,|\,t-1}\right) \leqslant \text{SE}\left(\widehat{\pi}^{m}_{t+h-1\,|\,t-1}\right), \ \forall m \in \mathcal{M},$$

where $\mathcal{M}$ is now the set of all 58 models considered in this paper.

The out-of-sample R-squared of combinations and averages between combinations and *available* Focus expectations are exhibited in Table 4. Once again, some results from Table 1 are kept for the sake of ease of exposition. According to the results displayed in Panel E, the combination obtained from the disaggregations into groups performed very well in horizons from two to eight months ahead. However, it was statistically superior to the Focus at the 10% level only in the latter. In the cumulative 12 months period, despite being subtly numerically superior, this combination was not statistically superior to the Focus survey. Considering the mean forecasts obtained by the average between combinations and Focus expectations available to the econometrician (Panel F), we noticed an improvement in short-term ($h \leqslant 5$) predictive performance for all combinations considered. The averaging became the combinations of groups and subgroups statistically superior to the Focus at least the 10% significance level for horizons two, three, four, and eight. The good performance of averaging in the short term led to an improvement of the 12-month accumulated forecasts concerning the case that did not take the average. However, they were still not statistically superior to the Focus in this case.

Table 4: Out-of-sample R-squared (%) – Model Combinations and Average with Focus

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	12 m
A. Survey													
Focus	**74.55**	24.26	10.61	9.25	9.19	7.23	1.21	-1.18	-2.34	-1.03	-1.58	-1.74	51.56
B. Aggregate inflation													
adaLASSO	69.18	23.44	1.90	2.55	5.32	6.60	*10.52*	6.90	*8.07**	*9.25**	4.08	6.19	**60.61**
FarmPredict	68.60	15.42	-0.35	2.15	7.97	2.98	-6.67	-0.80	1.13	-1.39	4.59	-0.71	45.49
CSR	36.84	7.69	5.00	6.49	12.15	**13.21**	8.77	11.82	6.63	-6.17	-4.22	-10.23	31.01
C. Disaggregation: tradable, nontradable and monitored prices (BCB)													
adaLASSO	65.78	21.30	8.51	7.28	4.18	7.69	7.62	4.26	-0.60	-8.01	3.83	*9.96**	53.13
FarmPredict	61.76	16.07	3.78	4.24	6.83	7.25	*11.72**	9.61**	*12.38***	3.92	*5.98*	-0.70	41.19
CSR	25.19	8.50	1.56	-2.53	10.89	3.78	-2.67	-9.85	-11.72	-2.67	-6.91	-14.05	-1.70
D. Disaggregation: groups (IBGE)													
adaLASSO	*69.30*	16.54	6.30	2.59	3.43	4.21	-3.68	-1.57	4.87	3.92	-4.22	5.29	48.92
FarmPredict	58.85	12.90	6.57	6.20	5.77	-0.06	-1.66	-4.45	7.44	*11.30**	-1.47	-1.87	41.26
CSR	36.42	5.89	5.36	5.94	8.29	8.27	2.20	0.11	0.76	3.20	-6.19	5.11	26.74
E. Combining Disaggregated Forecasts													
BCB Combination	63.03	21.70	14.34	-5.63	0.36	5.14	-3.14	5.81	3.41	1.04	-0.33	**21.16**	40.49
Groups Combination	52.42	*31.68*	*22.03*	13.80	*12.41*	*11.22*	5.58	*15.41**	-4.36	-9.09	-0.47	8.49	52.19
Subgroups Combination	45.23	31.51	16.75	12.96	-1.32	8.50	2.88	6.51	-1.62	-8.51	-11.26	7.64	47.00
F. Average between combined disaggregate forecasts and (available) Focus survey													
BCB Comb. & Focus	*72.97*	31.61	19.83*	6.46	7.02	6.74	0.68	5.30	1.02	0.19	3.88	*15.34*	51.80
Groups Comb. & Focus	68.64	*35.84***	*24.08***	*15.71**	**14.25**	*9.58*	7.15	11.93*	-1.55	-2.41	2.39	6.87	56.93
Subgroups Comb. & Focus	65.61	**36.24****	20.80**	*15.66*	7.79	7.89	5.88	7.80	0.38	-2.39	-2.07	6.94	54.04
G. Combining all models and average between combination of all models and (available) Focus survey													
Comb. *all* models	53.83	18.23	21.77	12.68	7.01	-7.91	**12.20**	**26.70*****	**16.79****	**14.18***	*5.65*	7.70	*57.52*
Comb. *all* models & Focus	68.41	30.66	**26.71****	**17.54**	*12.51*	3.81	8.62	*16.35***	7.85	7.37	**10.50**	9.03	*58.08**

Notes: ***, **, and * indicate that a specific model m performed statistically better than the Focus Survey (benchmark) in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}(\hat{\pi}^m_{t+h\,|\,t}) = \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t}) \times \mathbb{H}_1 : \text{MSE}(\hat{\pi}^m_{t+h\,|\,t}) < \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t})$. The highlighted value in blue bold indicates the best model for each horizon in terms of R^2_{oos}, and the values in blue italics indicate the second and third best models.

Finally, the broad combination that considers all models (Panel G) performed very well in the mid-term, obtaining the best R^2_{oos} in horizons from seven to ten among all models considered. This combination was statistically superior to Focus for eighth, ninth, and tenth horizons at significance levels of 1, 5, and 10%, respectively. Despite not having the superiority statistically attested in the accumulated in 12 months, this combination obtained an R^2_{oos} 6 p.p. higher than that recorded by the survey. All short-term results ($h \leqslant 6$) improved when averaging the broad combination with the available Focus. For $h = 3$ and $h = 4$, this average obtained the best R^2_{oos} among all the models, highlighting its statistical superiority at the 5% level about the Focus for the third horizon. The results for longer horizons are reasonable but inferior compared to the pure combination (without averaging). Consequently, the average performed well in the cumulative 12 months: R^2_{oos} of 58.08 against 51.56% for the Focus – this difference was statistically significant at the 10% level according to the DM test. It is important to note that only this model and the average between adaLASSO using aggregate and available Focus were statistically superior to the Focus (the latter at the 5% level).

6 Conclusion

In this paper, we investigate the use of inflation disaggregations to forecast the aggregate via aggregation of disaggregated forecasts. We innovated by considering the prediction of several disaggregates in a data-rich environment (many predictors), which is only possible with machine learning techniques. Such models employing disaggregates proved to be important in the econometrician's toolbox. For many horizons, the aggregation of disaggregated forecasts performed as well as survey-based expectations and model-based forecasts that directly employ the aggregate. In addition, combinations including disaggregated models obtained the best results among all the possibilities explored in the paper, even for short horizons. Such combinations did not be as advantageous or even would not exist if we did not consider the disaggregated models.

References

ALTUG, S., AND C. ÇAKMAKLI (2016): "Forecasting inflation using survey expectations and target inflation: Evidence for Brazil and Turkey," *International Journal of Forecasting*, 32(1), 138–153.

ANG, A., G. BEKAERT, AND M. WEI (2007): "Do macro variables, asset markets, or surveys forecast inflation better?," *Journal of Monetary Economics*, 54(4), 1163–1212.

ARAUJO, G. S., AND W. P. GAGLIANONE (2022): "Machine Learning Methods for Inflation Forecasting in Brazil: new contenders versus classical models," BCB Working Paper Series n. 561.

ARON, J., AND J. MUELLBAUER (2012): "Improving forecasting in an emerging economy, South Africa: Changing trends, long run restrictions and disaggregation," *International Journal of Forecasting*, 28(2), 456–476.

ATHEY, S. (2019): "The Impact of Machine Learning on Economics," in *The Economics of Artificial Intelligence: An Agenda*, chap. 21, pp. 507–547. University of Chicago Press.

BAI, J., AND S. NG (2002): "Determining the number of factors in approximate factor models," *Econometrica*, 70(1), 191–221.

——— (2008): "Forecasting economic time series using targeted predictors," *Journal of Econometrics*, 146(2), 304–317.

BAŃBURA, M., D. LEIVA-LEÓN, AND J.-O. MENZ (2021): "Do inflation expectations improve model-based inflation forecasts?," ECB Working Paper n. 2604.

BAŞTÜRK, N., C. ÇAKMAKLI, S. P. CEYHAN, AND H. K. VAN DIJK (2014): "Posterior-predictive evidence on US inflation using extended new Keynesian Phillips curve models with non-filtered data," *Journal of Applied Econometrics*, 29(7), 1164–1182.

BAYBUZA, I. (2018): "Inflation forecasting using machine learning methods," *Russian Journal of Money and Finance*, 77(4), 42–59.

BENALAL, N., J. L. DIAZ DEL HOYO, B. LANDAU, M. ROMA, AND F. SKUDELNY (2004): "To aggregate or not to aggregate? Euro area inflation forecasting," ECB Working Paper n. 374.

BERMINGHAM, C., AND A. D'AGOSTINO (2014): "Understanding and forecasting aggregate and disaggregate price dynamics," *Empirical Economics*, 46(2), 765–788.

BLANCHARD, O., AND J. GALÍ (2007): "Real wage rigidities and the New Keynesian model," *Journal of Money, Credit and Banking*, 39, 35–65.

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

BRUNEAU, C., O. DE BANDT, A. FLAGEOLLET, AND E. MICHAUX (2007): "Forecasting inflation using economic indicators: the case of France," *Journal of Forecasting*, 26(1), 1–22.

CAPISTRÁN, C., C. CONSTANDSE, AND M. RAMOS-FRANCIA (2010): "Multi-horizon inflation forecasts using disaggregated data," *Economic Modelling*, 27(3), 666–677.

CARLO, T. C., AND E. F. MARÇAL (2016): "Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon," *Applied Economics*, 48(50), 4846–4860.

CHOUDHARY, M. A., AND A. HAIDER (2012): "Neural Network Models for Inflation Forecasting: An Appraisal," *Applied Economics*, 44(20), 2631–2635.

CROUSHORE, D. (2010): "An evaluation of inflation forecasts from surveys using real-time data," *The BE Journal of Macroeconomics*, 10(1).

DEL NEGRO, M., AND S. EUSEPI (2011): "Fitting observed inflation expectations," *Journal of Economic Dynamics and Control*, 35(12), 2105–2131.

DIEBOLD, F. X., AND R. S. MARIANO (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3).

DUARTE, C., AND A. RUA (2007): "Forecasting inflation through a bottom-up approach: How bottom is bottom?," *Economic Modelling*, 24(6), 941–953.

ELLIOTT, G., A. GARGANO, AND A. TIMMERMANN (2013): "Complete subset regressions," *Journal of Econometrics*, 177(2), 357–373.

——— (2015): "Complete subset regressions with large-dimensional sets of predictors," *Journal of Economic Dynamics and Control*, 54, 86–110.

ESPASA, A., AND R. ALBACETE (2007): "Econometric modelling for short-term inflation forecasting in the euro area," *Journal of Forecasting*, 26(5), 303–316.

ESPASA, A., E. SENRA, AND R. ALBACETE (2002): "Forecasting inflation in the European Monetary Union: A disaggregated approach by countries and by sectors," *The European Journal of Finance*, 8(4), 402–421.

FAN, J., R. MASINI, AND M. C. MEDEIROS (2021): "Bridging factor and sparse models," Available at SSRN 3789141.

FAUST, J., AND J. H. WRIGHT (2013): "Forecasting Inflation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2A, chap. 1, pp. 3–56. Elsevier, North Holland.

FLORIDO, G. P. (2021): "Essays on Inflation Forecasting," Ph.D. thesis, Sao Paulo School of Economics, FGV.

FORNI, M., M. HALLIN, M. LIPPI, AND L. REICHLIN (2003): "Do financial variables help forecasting inflation and real activity in the euro area?," *Journal of Monetary Economics*, 50(6), 1243–1255.

FULTON, C., AND K. HUBRICH (2021): "Forecasting US inflation in real time," *Econometrics*, 9(36), 1–20.

GALÍ, J., AND M. GERTLER (1999): "Inflation dynamics: A structural econometric analysis," *Journal of Monetary Economics*, 44(2), 195–222.

GARCIA, M. G., M. C. MEDEIROS, AND G. F. VASCONCELOS (2017): "Real-time inflation forecasting with high-dimensional models: The case of Brazil," *International Journal of Forecasting*, 33(3), 679–693.

GOGAS, P., AND T. PAPADIMITRIOU (2021): "Machine learning in economics and finance," *Computational Economics*, 57(1), 1–4.

HENDRY, D. F., AND K. HUBRICH (2011): "Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate," *Journal of Business & Economic Statistics*, 29(2), 216–227.

HUBRICH, K. (2005): "Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?," *International Journal of Forecasting*, 21(1), 119–136.

IBARRA, R. (2012): "Do disaggregated CPI data improve the accuracy of inflation forecasts?," *Economic Modelling*, 29(4), 1305–1313.

INOUE, A., AND L. KILIAN (2008): "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.

LÜTKEPOHL, H. (1984): "Linear transformations of vector ARMA processes," *Journal of Econometrics*, 26(3), 283–293.

MEDEIROS, M. C., G. VASCONCELOS, AND E. FREITAS (2016): "Forecasting Brazilian inflation with high-dimensional models," *Brazilian Review of Econometrics*, 36(2), 223–254.

MEDEIROS, M. C., G. F. VASCONCELOS, Á. VEIGA, AND E. ZILBERMAN (2021): "Forecasting inflation in a data-rich environment: the benefits of machine learning methods," *Journal of Business & Economic Statistics*, 39(1), 98–119.

MOSER, G., F. RUMLER, AND J. SCHARLER (2007): "Forecasting Austrian Inflation," *Economic Modelling*, 24(3), 470–480.

MOSHIRI, S., AND N. CAMERON (2000): "Neural Network versus Econometric Models in Forecasting Inflation," *Journal of Forecasting*, 19(3), 201–217.

MULLAINATHAN, S., AND J. SPIESS (2017): "Machine learning: an applied econometric approach," *Journal of Economic Perspectives*, 31(2), 87–106.

NAKAMURA, E. (2005): "Inflation forecasting using a neural network," *Economics Letters*, 86(3), 373–378.

STOCK, J. H., AND M. W. WATSON (1999): "Forecasting inflation," *Journal of Monetary Economics*, 44(2), 293–335.

——— (2002): "Macroeconomic forecasting using diffusion indexes," *Journal of Business & Economic Statistics*, 20(2), 147–162.

THOMAS, L. B. (1999): "Survey measures of expected US inflation," *Journal of Economic Perspectives*, 13(4), 125–144.

ZENG, J. (2017): "Forecasting aggregates with disaggregate variables: Does boosting help to select the Most relevant predictors?," *Journal of Forecasting*, 36(1), 74–90.

ZOU, H. (2006): "The Adaptive Lasso and Its Oracle Properties," *Journal of the American Statistical Association*, 101(476), 1418–1429.

Appendices

A Explanatory Variables

Table A1: Explanatory variables description

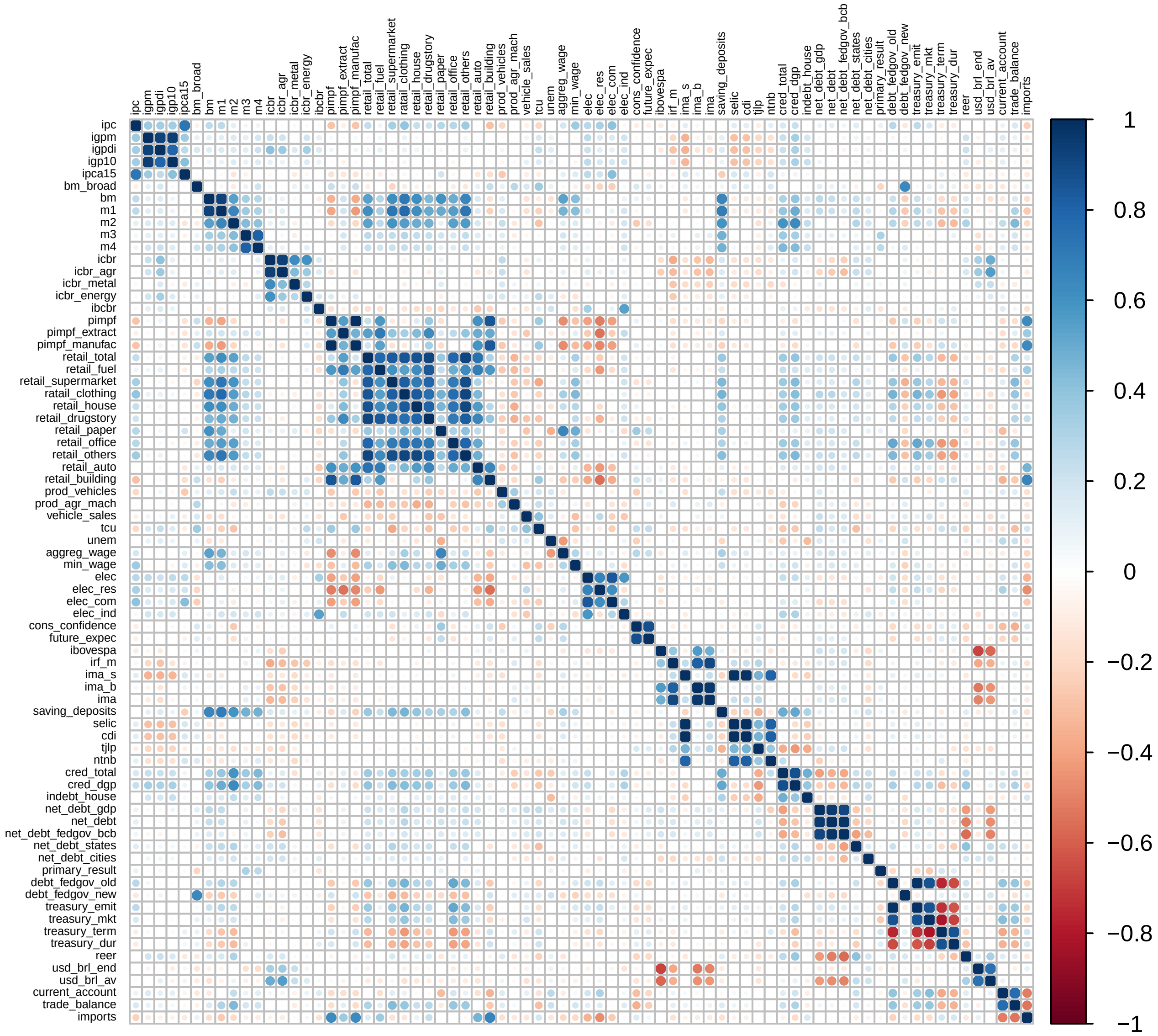
#	Abbreviation	Description	Unit	Source	Lag	Transformation
		A. Prices and Money				
1	inf, inf.g#	IPCA and disaggregations	% per month	IBGE	1	-
2	expec	Inflation expectation	% per month	BCB	0	-
2	ipc	IPC-Br	% per month	FGV	1	-
3	igpm	IGP-M	% per month	FGV	1	-
4	igpdi	IGP-DI	% per month	FGV	1	-
5	igp10	IGP-10	% per month	FGV	1	-
6	ipca15	IPCA-15	% per month	IBGE	1	-
7	bm_broad	Broad Monetary Base – end-of-period balance	index	BCB	2	% change
8	bm	Monetary Base – working day balance average	Index	BCB	2	% change
9	m1	Money supply M1 – working day balance average	Index	BCB	2	% change
10	m2	Money supply M2 – end-of-period balance	Index	BCB	2	% change
11	m3	Money supply M3 – end-of-period balance	Index	BCB	2	% change
12	m4	Money supply M4 – end-of-period balance	Index	BCB	2	% change
		B. Commodities Prices				
13	icbr	Brazilian Commodity index (all)	index	BCB	1	% change
14	icbr_agr	Brazilian Commodity index – agriculture	index	BCB	1	% change
15	icbr_metal	Brazilian Commodity index – metal	index	BCB	1	% change
16	icbr_energy	Brazilian Commodity index – energy	index	BCB	1	% change
		C. Economic Activity				
17	ibcbr	Brazilian IBC-Br Economic Activity index	index	BCB	3	% change
18	pimpf	Industrial Production (general)	index	IBGE	2	% change
19	pimpf_extract	Extractive Industry Production	index	IBGE	2	% change
20	pimpf_manufac	Manufacturing Industry Production	index	IBGE	2	% change
21	retail_total	Retail sales volume – total	index	IBGE	2	% change
22	retail_fuel	Retail sales volume – fuels and oils	index	IBGE	2	% change
23	retail_supermarket	Retail sales volume – supermarkets and food products	index	IBGE	2	% change
24	retail_clothing	Retail sales volume – fabrics, clothing and shoes	index	IBGE	2	% change
25	retail_house	Retail sales volume – furniture and appliances	index	IBGE	2	% change
26	retail_drugstory	Retail sales volume – pharmaceutical and cosmetic articles	index	IBGE	2	% change
27	retail_paper	Retail sales volume – books, newspapers and stationery	index	IBGE	2	% change
28	retail_office	Retail sales volume – office and eletronical equipments	index	IBGE	2	% change
29	retail_others	Retail sales volume – others	index	IBGE	2	% change
30	retail_building	Retail sales volume – building material	index	IBGE	2	% change
31	retail_auto	Retail sales volume – automotive and parts	index	IBGE	2	% change
32	prod_vehicles	Vehicle production – total	units	Anfavea	1	% change
33	prod_agr_mach	Production of agricultural machinery – total	units	Anfavea	1	% change
34	vehicle_sales	Vehicle sales by dealerships – total	units	Fenabrave	1	% change
35	tcu	Total capacity utilization – manufacturing industry	%	FGV	1	first difference
		D. Employment				
36	unem	Unemployment (combination of PME and PNADC)	%	IBGE	3	first difference
37	aggreg_wage	Overall Earnings (broad wage income)	R$ (million)	BCB	2	% change
38	min_wage	Federal Minimum Wage	R$	MTb	0	% change
		E. Electricity				
39	elec	Electricity consumption – total	GWh	Eletrobrás	3	% change
40	elec_res	Electricity consumption – residential	GWh	Eletrobrás	3	% change
41	elec_com	Electricity consumption – commercial	GWh	Eletrobrás	3	% change
42	elec_ind	Electricity consumption – industry	GWh	Eletrobrás	3	% change
		F. Confidence				
43	cons_confidence	Consumer Confidence index	index	Fecomercio	1	% change
44	future_expec	Future expectations index	index	Fecomercio	1	% change
		G. Finance				
45	ibovespa	Ibovespa index	% per month	B3	1	-
46	irf_m	Anbima Market Index of the prefixed federal bonds	index	Anbima	1	% change
47	ima_s	Anbima Market Index of the federal bonds tied to the SELIC rate	index	Anbima	1	% change
48	ima_b	Anbima Market Index of the federal bonds tied to the IPCA index	index	Anbima	1	% change
49	ima	General Anbima Market index	index	Anbima	1	% change
50	saving_deposits	Saving deposits – end-of-period balance	R$ (mil)	BCB	2	% change
51	selic	Selic Basic Interest rate	% per month	BCB	1	–
52	cdi	Cetip DI Interbank Deposits rate	% per month	Cetip	1	–
53	tjlp	TJLP Long-term Interest rate	% per year	BCB	1	–
54	ntnb	3-Year Treasury (real) Rate indexed to the IPCA (NTN-B)	% per year	Anbima	1	–

(continued on next page)

Table A1: Explanatory variables description (cont.)

#	Abbreviation	Description	Unit	Source	Lag	Transformation
		H. Credit				
55	cred_total	Credit outstanding – total	R$ (million)	BCB	2	% change
56	cred_dgp	Credit outstanding as a percentage of GDP	% of GDP	BCB	2	first difference
57	indebt_house	Household debt to income	% of 12m income	BCB	4	first difference
		I. Government				
58	net_debt_gdp	Net public debt (% GDP) -Consolidated public sector	% of GDP	BCB	2	first difference
59	net_debt	Net public debt – Total – Consolidated public sector	R$ (million)	BCB	2	% change
60	net_debt_fedgov_bcb	Net public debt – Federal Government and Central Bank	R$ (million)	BCB	2	% change
61	net_debt_states	Net public debt – State governments	R$ (million)	BCB	2	% change
62	net_debt_cities	Net public debt – Municipal governments	R$ (million)	BCB	2	% change
63	primary_result	Primary result – Consolidated public sector	R$ (million)	BCB	2	% change
64	debt_fedgov_old	Gross general government debt – Method used until 2007	R$ (million)	BCB	2	% change
65	debt_fedgov_new	Gross general government debt – Method used since 2008	R$ (million)	BCB	2	% change
66	treasury_emit	National Treasuary domestic securities – Total issued	R$ (million)	BCB	2	% change
67	treasury_mkt	National Treasuary domestic securities – Total on market	R$ (million)	BCB	2	% change
68	treasury_term	National Treasury securities debt – medium term	months	BCB	2	first difference
69	treasury_dur	National Treasury securities debt – medium duration	months	BCB	2	first difference
		J. Exchange and International Transactions				
70	reer	Real Effective Exchange Rate	R$/other	BIS	2	% change
71	usd_brl_end	USD-BRL rate – end-of-period	R$/US$	BCB	1	% change
72	usd_brl_av	USD-BRL rate – monthly average	R$/US$	BCB	1	% change
73	current_account	Current account – net	US$ (million)	BCB	2	% change
74	trade_balance	Balance on goods and services – net (Brazilian trade balance)	US$ (million)	BCB	2	% change
75	imports	Imports	US$ (million)	BCB	2	% change

Figure A1: Explanatory variables correlation



Notes: Correlations were calculated considering the period from March 2007 to January 2021 according to data availability for all variables.